archives-ouvertes.fr

Mathematics for theoretical physics

Jean Claude Dutailly

Mathematics for theoretical physics

Jean Claude.Dutailly
Paris

September 25, 2012

Abstract

This book intends to give the main definitions and theorems in mathematics which could be useful for workers in theoretical physics. It gives an extensive and precise coverage of the subjects which are addressed, in a consistent and intelligible manner.The first part addresses the Foundations (mathematical logic, set theory, categories), the second Algebra (algebraic strucutes, groups, vector spaces tensors, matrices, Clifford algebra). The third Analysis (general topology, measure theory, Banach Spaces, Spectral theory). The fourth Differential Geometry (derivatives, manifolds, tensorial bundle, pseudo-riemannian manifolds, symplectic manifolds). The fifth Lie Algebras, Lie Groups.and representation theory. The sixth Fiber bundles and jets. The last one Functional Analysis (differential operators, distributions, ODE, PDE, variational calculus). Several signficant new results are presented (distributions over vector bundles, functional derivative, spin bundle and manifolds with boundary).

The purpose of this book is to give a comprehensive collection of precise definitions and results in advanced mathematics, which can be useful to workers in mathematic or physics.

The specificities of this book are :

- it is self contained : any definition or notation used can be found within

- it is precise : any theorem lists the precise conditions which must be met for its use

- it is easy to use : the book proceeds from the simple to the most advanced topics, but in any part the necessary definitions are reminded so that the reader can enter quickly into the subject

- it is comprehensive : it addresses the basic concepts but reaches most of the advanced topics which are required nowodays

- it is pedagogical : the key points and usual misunderstandings are underlined so that the reader can get a strong grasp of the tools which are presented.

The first option is unusual for a book of this kind. Usually a book starts with the assumption that the reader has already some background knowledge. The problem is that nobody has the same background. So a great deal is dedicated to remind some basic stuff, in an abbreviated way, which does not left much scope to their understanding, and is limited to specific cases. In fact, starting

from the very beginning, it has been easy, step by step, to expose each concept in the most general settings. And, by proceeding this way, to extend the scope of many results so that they can be made available to the - unavoidable - special case that the reader may face. Overall it gives a fresh, unified view of the mathematics, but still affordable because it avoids as far as possible the sophisticated language which is fashionable. The goal is that the reader understands clearly and effortlessly, not to prove the extent of the author's knowledge.

The definitions choosen here meet the "generally accepted definitions" in mathematics. However, as they come in many flavors according to the authors and their field of interest, we have striven to take definitions which are both the most general and the most easy to use.

Of course this cannot be achieved with some drawbacks. So many demonstrations are omitted. More precisely the chosen option is the following :

- whenever a demonstration is short, it is given entirely, at least as an example of "how it works"

- when a demonstation is too long and involves either technical or specific conditions, a precise reference to where the demonstation can be found is given. Anyway the theorem is written in accordance with the notations and definitions of this book, and a special attention has been given that they match the reference.

- exceptionnaly, when this is a well known theorem, whose demonstration can be found easily in any book on the subject, there is no reference.

The bibliography is short. Indeed due to the scope which is covered it could be enormous. So it is strictly limited to the works which are referenced in the text, with a priority to the most easily available sources.

This is not mainly a research paper, even if the unification of the concepts is, in many ways, new, but some significant results appear here for the first time, to my knowledge.

- distributions over vector bundles
- a rigorous definitition of functional derivatives
- a manifold with boundary can be defined by a unique function

and several other results about Clifford algebras, spin bundles and differential geometry.

[1]

---

[1] j.c.dutailly@free.fr

CONTENTS

PART 2 : ALGEBRA

PART 3 : ANALYSIS

PART 4 : DIFFERENTIAL GEOMETRY

PART 4 : DIFFERENTIAL GEOMETRY

PART 5 : LIE ALGEBRAS AND LIE GROUPS

PART 6 : FIBER BUNDLES

PART 7 : FUNCTIONAL ANALYSIS

BIBLIOGRAPHY

Part I
PART1 : FOUNDATIONS

In this first part we start with what makes the real foundations of today mathematics : logic, set theory and categories. The two last subsections are natural in this book, and they will be mainly dedicated to a long list of definitions, mandatory to fix the language that is used in the rest of the book. A section about logic seems appropriate, even if it gives just an overview of the topic, because this is a subject that is rarely addressed, except in specialized publications, and should give some matter for reflection, notably to physicists.

1 LOGIC

For a mathematician logic can be addressed from two points of view :

- the conventions and rules that any mathematical text should follow in order to be deemed "right"

- the consistency and limitations of any formal theory using these logical rules.

It is the scope of a branch of mathematics of its own : "mathematical logic"

Indeed logic is not limited to a bylaw for mathematicians : there are also theorems in logic. To produce these theorems one distinguishes the object of the investigation ("language-object" or "theory") and the language used to proceed to the demonstrations in mathematical logic, which is informal (plain english). It seems strange to use a weak form of "logic" to prove results about the more formal theories but it is related to one of the most important feature of any scientific discourse : that it must be perceived and accepted by other workers in the field as "sensible" and "convincing". And in fact there are several schools in logic : some do not accept any nonnumerable construct, or the principle of non contradiction, which makes logic a confusing branch of mathematics. But whatever the interest of exotic lines of reasoning in specific fields, for the vast majority of mathematicians, in their daily work, there is a set of "generally accepted logical principles".

On this topic we follow mainly Kleene where definitions and theorems can be found.

1.1 Propositional logic

Logic can be considered from two points of view : the first ("models") which is focused on telling what are true or false statements, and the second ("demonstration") which strives to build demonstrations from premisses. This distinction is at the heart of many issues in mathematical logic.

1.1.1   Models

Formulas

Definition 1 *An **atom**[2] is any given sentence accepted in the theory.*

The atoms are denoted as latin letters A,B,..

Definition 2 *The logical operators are :*

$\sim$: equivalent
$\Rightarrow$: imply
$\wedge$: and (both)
$\vee$: or (possibly both)
$\neg$: negation
(notation and list depending on the authors)

Definition 3 *A **formula** is any finite sequence of atoms linked by logical operators.*

One can build formulas from other formulas using these operators. A formula is "well-built" (it is deemed acceptable in the theory) if it is constructed according to the previous rules.

Examples : if "$3 + 2 = x$","$\sqrt{5} - 3 > 2$","$x^2 + 2x - 1 = 0$" are atoms then $\left((3 + 2 = x) \wedge \left(x^2 + 2x - 1 = 0\right)\right) \Rightarrow \left(\sqrt{5} - 3 > 2\right)$ is a well built formula.

In building a formula we do not question the meaning or the validity of the atoms (this the job of the theory which is scrutinized) : we only follow rules to build formulas from given atoms.

When building formulas with the operators it is always good to use brackets to delimite the scope of the operators. However there is a rule of precedence (by decreasing order): $\sim > \Rightarrow > \wedge > \vee > \neg$

Truth-tables

The previous rules give only the "grammar" : how to build accepted formulas. But a formula can be well built but meaningless, or can have a meaning only if certain conditions are met. Logic is the way to tell if something is true or false.

Definition 4 *To each atom of a theory is attached a "**truth-table**", with only two values : true (T) or false (F) exclusively.*

Definition 5 *A **model** for a theory is the list of its atoms and their truth-table.*

Definition 6 *A **proposition** is any formula issued from a model*

---

[2]The name of an object is in boldface the first time it appears (in its definition)

The rules telling how the operators work to deduce the truth table of a formula from the tables ot its atoms are the following (A,B are any formula) :

$$\begin{bmatrix} A & B & (A \sim B) & (A \Rightarrow B) & (A \wedge B) & (A \vee B) \\ T & T & T & T & T & T \\ T & F & F & F & F & T \\ F & T & F & T & F & T \\ F & F & T & T & F & F \end{bmatrix} \begin{bmatrix} A & (\neg A) \\ T & F \\ F & T \end{bmatrix}$$

The only non obvious rule is for $\Rightarrow$. It is the only one which provides a full and practical set of rules, but other possibilities are mentioned in quantum physics.

Valid formulas

With these rules the truth-table of any formula can be computed (formulas have only a finite number of atoms).

The formulas which are always true (their truth-table presents only T) are of particular interest.

Definition 7 *A formula A of a model is said to be **valid** if it is always true. It is then denoted $\vDash A$.*

Definition 8 *A formula B is a **valid consequence** of A if $\vDash (A \Rightarrow B)$. This is denoted : $A \vDash B$.*

More generally one writes : $A_1, ..A_m \vDash B$

Valid formulas are crucial in logic. There are two different categories of valid formulas:

- formulas which are always valid, whatever the model : they provide the "model" of propositional calculs in mathematical logic, as they tell how to produce "true" statements without any assumption about the meaning of the formulas.

- formulas which are valid in some model only : they describe the properties assigned to some atoms in the theory which is modelled. So, from the logical point of view, they define the theory itself.

The following formula are always valid in any model (and most of them are of constant use in mathematics). Indeed they are just the traduction of the previous tables.

1. first set (they play a specific role in logic):
$(A \wedge B) \Rightarrow A ; (A \wedge B) \Rightarrow B$
$A \Rightarrow (A \vee B) ; B \Rightarrow (A \vee B)$
$\neg\neg A \Rightarrow A$
$A \Rightarrow (B \Rightarrow A)$
$(A \sim B) \Rightarrow (A \Rightarrow B) ; (A \sim B) \Rightarrow (B \Rightarrow A)$
$(A \Rightarrow B) \Rightarrow ((A \Rightarrow (B \Rightarrow C)) \Rightarrow (A \Rightarrow C))$

$A \Rightarrow (B \Rightarrow (A \wedge B))$
$(A \Rightarrow B) \Rightarrow ((A \Rightarrow \neg B) \Rightarrow \neg A)$
$(A \Rightarrow B) \Rightarrow ((B \Rightarrow A) \Rightarrow (A \sim B))$

2. Others (there are infinitely many others formulas which are always valid) :

$A \Rightarrow A;$
$A \sim A; (A \sim B) \sim (B \sim A); ((A \sim B) \wedge (B \sim C)) \Rightarrow (A \sim C)$
$(A \Rightarrow B) \sim ((\neg A) \Rightarrow (\neg B))$
$\neg A \Rightarrow (A \Rightarrow B)$
$\neg\neg A \sim A; \neg(A \wedge (\neg A)); A \vee (\neg A)$
$\neg(A \vee B) \sim ((\neg A) \wedge (\neg B)); \neg(A \wedge B) \sim ((\neg A) \vee (\neg B)); \neg(A \Rightarrow B) \sim (A \wedge (\neg B))$

Notice that $\vDash A \vee (\neg A)$ meaning that a formula is either true or false is an obvious consequence of the rules which have been set up here.

An example of formula which is valid in a specific model : in a set theory the expressions "$a \in A$", "$A \subset B$" are atoms, they are true or false (but their value is beyond pure logic). And "$((a \in A) \wedge (A \subset B)) \Rightarrow (a \in B)$" is a formula. To say that it is always true expresses a fundamental property of set theory (but we could also postulate that it is not always true, and we would have another set theory).

Theorem 9 *If $\vDash A$ and $\vDash (A \Rightarrow B)$ then : $\vDash B$*

Theorem 10 *$\vDash A \sim B$ iff[3] A and B have same tables.*

Theorem 11 *Duality: Let be E a formula built only with atoms $A_1, .. A_m$, their negation $\neg A_1, .. \neg A_m$, the operators $\vee, \wedge$, and E' the formula deduced from E by substituting $\vee$ with $\wedge, \wedge$ with $\vee, A_i$ with $\neg A_i, \neg A_i$ with A_i then :*
If $\vDash E$ then $\vDash \neg E'$
If $\vDash \neg E$ then $\vDash E'$
With the same procedure for another similar formula F:
If $\vDash E \Rightarrow F$ then $\vDash F' \Rightarrow E'$
If $\vDash E \sim F$ then $\vDash E' \sim F'$

1.1.2   Demonstration

Usually one does not proceed by truth tables but by demonstrations. In a formal theory, axioms, hypotheses and theorems can be written as formulas. A demonstration is a sequence of formulas using logical rules and rules of inference, starting from axioms or hypotheses and ending by the proven result.

In deductive logic a **formula** is always true. They are built according to the following rules by linking formulas with the logical operators above :

i) There is a given set of formulas $(A_1, A_2, ... A_m, ..)$ (possibly infinite) called the **axioms** of the theory

---

[3]We will use often the usual abbreviation "iff" for "if and only if"

ii) There is an inference rule : if A is a formula, and $(A \Rightarrow B)$ is a formula, then (B) is a formula.

iii) Any formula built from other formulas with logical operators and using the "first set" of rules above is a formula

For instance if A,B are formulas, then $((A \wedge B) \Rightarrow A)$ is a formula.

The formulas are listed, line by line. The last line gives a "true" formula which is said to be proven.

Definition 12 *A **demonstration** is a finite sequence of formulas where the last one B is the proven formula, and this is denoted :* $\Vdash B$. *B is **provable**.*

Similarly B is deduced from $A_1, A_2, ..$ is denoted : $A_1, A_2, ..A_m, .. \Vdash B :$. In this picture there are logical rules (the "first set" of formulas and the inference rule) and "non logical" formulas (the axioms). The set of logical rules can vary according to the authors, but is roughly always the same. The critical part is the set of axioms which is specific to the theory which is under review.

Theorem 13 $A_1, A_2, ...A_m \Vdash A_p$ *with 1<p$\leq$ m*

Theorem 14 *If* $A_1, A_2, ...A_m \Vdash B_1, A_1, A_2, ...A_m \Vdash B_2, ...A_1, A_2, ...A_m \Vdash B_p$ *and* $B_1, B_2, ...B_p \Vdash C$ *then* $A_1, A_2, ...A_m \Vdash C$

Theorem 15 *If* $\Vdash (A \Rightarrow B)$ *then* $A \Vdash B$ *and conversely : if* $A \Vdash B$ *then* $\Vdash (A \Rightarrow B)$

1.2 Predicates

In propositional logic there can be an infinite number of atoms (models) or axioms (demonstration) but, in principle, they should be listed prior to any computation. This is clearly a strong limitation. So the previous picture is extended to **predicates**, meaning formulas including variables and functions.

1.2.1 Models with predicates

Predicate

The new elements are : variables, quantizers, and propositional functions.

Definition 16 *A **variable** is a symbol which takes its value in a given collection D (the **domain**).*

They are denoted x,y,z,...It is assumed that the domain D is always the same for all the variables and it is not empty. A variable can appears in different places, with the usual meaning that in this case the same value must be assigned to these variables.

Definition 17 *A **propositional function** is a symbol, with definite places for one or more variables, such that when one replaces each variable by one of their value in the domain, the function becomes a proposition.*

They are denoted : $P(x, y), Q(r), ...$There is a truth-table assigned to the function for all the combinations of variables.

Definition 18 *A **quantizer** is a logical operator acting on the variables.*

They are :
$\forall$: for any value of the variable (in the domain D)
$\exists$: there exists a value of the variable (in the domain D)
A quantizer acts, on one variable only, each time it appears : $\forall x, \exists y, ..$. This variable is then **bound**. A variable which is not bound is **free**. A quantizer cannot act on a previously bound variable (one cannot have $\forall x, \exists x$ in the same formula). As previously it is always good to use different symbols for the variables and brackets to precise the scope of the operators..

Definition 19 *A **predicate** is a sentence comprised of propositions, quantizers preceding variables, and propositional functions linked by logical operators.*

Examples of predicates :
$((\forall x, (x + 3 > z)) \wedge A) \Rightarrow^\neg \left(\exists y, \left(\sqrt{y^2 - 1} = a \right) \right) \vee (z = 0)$
$\forall n ((n > N) \wedge (\exists p, (p + a > n))) \Rightarrow B$
To evaluate a predicate one needs a truth-rule for the quantizers $\forall, \exists$:
- a formula $(\forall x, A(x))$ is T if A(x) is T for all values of x
- a formula $(\exists x, A(x)))$ is T if A(x) has at least one value equal to T
With these rules whenever all the variables in a predicate are bound, this predicate, for the thuth table purpose, becomes a proposition.
Notice that the quantizers act only on variables, not formulas. This is specific to **first order predicates**. In higher orders predicates calculus there are expressions like "$\forall A$", and the theory has significantly different outcomes.

Valid consequence
With these rules it is possible, in principle, to compute the truth table of any predicate.

Definition 20 *A predicate A is **D-valid**, denoted $^D \models A$ if it is valid whatever the value of the free variables in D. It is **valid** if is D-valid whatever the domain D.*

The propositions listed previously in the "first set" are valid for any D.
$\models A \sim B$ iff for any domain D A and B have the same truth-table.

1.2.2 Demonstration with predicates

The same new elements are added : variables, quantizers, propositional functions. Variables and quantizers are defined as above (in the model framework) with the same conditions of use.

A formula is built according to the following rules by linking formulas with the logical operators and quantizers :

i) There is a given set of formulas $(A_1, A_2, ...A_m, ..)$ (possibly infinite) called the **axioms** of the theory

ii) There are three inference rules :

- if A is a formula, and $(A \Rightarrow B)$ is a formula, then (B) is a formula

- If C is a formula where x is not present and A(x) a formula, then :

if $C \Rightarrow A(x)$ is a formula, then $C \Rightarrow \forall x A(x)$ is a formula

if $A(x) \Rightarrow C$ is a formula, then $\exists x A(x) \Rightarrow C$ is a formula

iii) Any formula built from other formulas with logical operators and using the "first set" of rules above plus :

$\forall x A(x) \Rightarrow A(r)$

$A(r) \Rightarrow \exists x A(x)$

where r is free, is a formula

Definition 21 *B is **provable** if there is a finite sequence of formulas where the last one is B, which is denoted :* $\Vdash B$.

B can be deduced from $A_1, A_2, ...A_m$ if B is provable starting with the formulas $A_1, A_2, ...A_m$,and is denoted : $A_1, A_2, ...A_m \Vdash B$

1.3 Formal theories

1.3.1 Definitions

The previous definitions and theorems give a framework to review the logic of formal theories. A formal theory uses a symbolic language in which terms are defined, relations between some of these terms are deemed "true" to express their characteristics, and logical rules are used to evaluate formulas or deduce theorems. There are many refinements and complications but, roughly, the logical rules always come back to some kind of predicates logic as exposed in the previous section. But there are two different points of view : the "models" side and the "demonstration" side : the same theory can be described using a model (model type theory) or axioms and deductions (deductive type).

Models are related to the "semantic" of the theory. Indeed they are based on the assumption that for every atom there is some truth-table that could be exhibited, meaning that there is some "extra-logic" to compute the result. And the non purely logical formulas which are set to be valid (always true in the model) characterize the properties of the objects "modelled" by the theory.

Demonstrations are related to the "syntactic" part of the theory. They deal only with formulas without any concern about their meaning : either they are logical formulas (the first set) or they are axioms, and in both cases they are

assumed to be "true", in the meaning that they are worth to be used in a demonstration. The axioms sum up the non logical part of the system. The axioms on one hand and the logical rules on the other hand are all that is necessary to work.

Both model theories and deductive theories use logical rules (either to compute truth-tables or to list formulas), so they have a common ground. And the non-logical formulas which are valid in a model are the equivalent of the axioms of a deductive theory. So the two points of view are not opposed, but proceed from the two meanings of logic.

In reviewing the logic of a formal theory the main questions that arise are :

- which are the axioms needed to account for the theory (as usual one wants to have as few of them as possible) ?

- can we assert that there is no formula A such that both A and its negation $\neg A$ can be proven ?

- can we prove any valid formula ?

- is it possible to list all the valid formulas of the theory ?

A formal theory of the model type is said to be "**sound**" (or consistent) if only valid formulas can be proven. Conversely a formal theory of the deductive type is said to be "**complete"** if any valid formula can be proven.

1.3.2   Completness of the predicate calculus

Predicate logic (**first order logic**) can be seen as a theory by itself. From a set of atoms, variables and propositional functions one can build formulas by using the logical operators for predicates. There are formulas which are always valid in the propositional calculus, and there are similar formulas in the predicates calculus, whatever the domain D. Starting with these formulas, and using the set of logical rules and the inference rules as above one can build a deductive theory.

The Gödel's completness theorem says that any valid formula can be proven, and conversely that only valid formulas can be proven. So one can write in the first order logic : $\vDash A$ iff $\Vdash A$.

It must be clear that this result, which justifies the apparatus of first order logic, stands only for the formulas (such as those listed above) which are valid in any model : indeed they are the pure logical relations, and do not involve any "non logical" axioms.

A "compactness" theorem by Gödel says in addition that if a formula can be proven from a set of formulas, it can also be proven by a finite set of formulas : there is always a demonstration using a finite number of steps and formulas.

These results are specific to first order logic, and does not hold for higher order of logic (when the quantizers act on formulas and not only on variables).

Thus one can say that mathematical logic (at least under the form of first order propositional calculus) has a strong foundation.

1.3.3   Incompletness theorems

At the beginning of the XX° century mathematicians were looking forward to a set of axioms and logical rules which could give solid foundations to mathematics (the "Hilbert's program"). Two theories are crucial for this purpose : set theory and natural number (arithmetics). Indeed set theory is the language of modern mathematics, and natural numbers are a prerequisite for the rule of inference, and even to define infinity (through cardinality). Such formal theories use the rules of first order logic, but require also additional "non logical" axioms. The axioms required in a formal set theory (such as Zermelo-Frankel's) or in arithmetics (such as Peano's) are well known. There are several systems, more or less equivalent.

A formal theory is said to be **effectively generated** if its set of axioms is a recursively enumerable set. This means that there is a computer program that, in principle, could enumerate all the axioms of the theory. Gödel's first incompleteness theorem states that any effectively generated theory capable of expressing elementary arithmetic cannot be both consistent and complete. In particular, for any consistent, effectively generated formal theory that proves certain basic arithmetic truths, there is an arithmetical statement that is true but not provable in the theory (Kleene p. 250). In fact the "truth" of the statement must be understood as : neither the statement or its negation can be proven. As the statement is true or false the statement itself or its converse is true. All usual theories of arithmetics fall under the scope of this theorem. So one can say that in mathematics the previous result ($\vDash A$ iff $\Vdash A$) does not stand.

This result is not really a surprise : in any formal theory we can build infinitely many predicates, which are "grammatically" correct. To say that there is always a way to prove any such predicate (or its converse) is certainly a crude assumption. It is linked to the possibility to write computer programs to automatically check demonstrations.

1.3.4   Decidable and computable theories

The incompletness theorems are closely related to the concepts of "decidable" and "computable".

In a formal deductive theory computer programs can be written to "formalize" demonstrations (an exemple is "Isabelle" see the Internet), so that they can be made safer.

One can go further and ask if it is possible to design a program such that it could, for any statement of the theory, check if it is valid (model side) or provable (deducible side). If so the theory is said **decidable**.

The answer is yes for the propositional calculus (without predicate), because it is always possible to compute the truth table, but it is no in general for predicates calculus. And it is no for theories modelling arthmetics.

Decidability is an aspect of computability : one looks for a program which could, starting from a large class of inputs, compute an answer which is yes or

no.

Computability is studied through "**Türing machines**", which are schematic computers. A Türing machine is comprised of an input system (a flow of binary data read bit by bit), a program (the computer has p "states", including an "end", and it goes from one state to another according to its present state and the bit that has been read), and an output system (the computer writes a bit). A Türing machine can compute integer functions (the input, output and parameters are integers). One demonstration of the Gödel incompletness theorem shows that there are functions that cannot be computed : notably the function telling, for any given input, in how many steps the computer would stop.

If we look for a program that can give more than a "Yes/No" answer one has the so-called "function problems", which study not only the possibility but the efficiency (in terms of ressources used) of algorithms. The **complexity** of a given problem is measured by the ratio of the number of steps required by a Türing machine to compute the function, to the size in bits of the input (the problem).

2 SET THEORY

2.1 Axiomatic

Set theory was founded by Cantor and Dedekind in early XX° century. The initial set theory was impaired by paradoxes, which are usually the consequences of an inadequate definition of a "set of sets". Several improved versions were proposed, and its most common , formalized by Zermello-Fraenkel, is denoted ZFC when it includes the axiom of choice. For the details see Wikipedia "Zermelo–Fraenkel set theory".

2.1.1 Axioms of ZFC

Some of the axioms listed below are redundant, as they can be deduced from others, depending of the presentation. But it can be useful to know their names :

Axiom 22 *Axiom of extensionality : Two sets are equal (are the same set) if they have the same elements.*

Equality is defined as : $(A = B) \sim ((\forall x \, (x \in A \sim x \in B)) \wedge (\forall x \, (A \in x \sim B \in x)))$

Axiom 23 *Axiom of regularity (also called the Axiom of foundation) : Every non-empty set A contains a member B such that A and B are disjoint sets.*

Axiom 24 *Axiom schema of specification (also called the axiom schema of separation or of restricted comprehension) : If A is a set, and P(x) is any property which may characterize the elements x of A, then there is a subset B of A containing those x in A which satisfy the property.*

The axiom of specification can be used to prove the existence of one unique empty set, denoted $\varnothing$, once the existence of at least one set is established.

Axiom 25 *Axiom of pairing : If A and B are sets, then there exists a set which contains A and B as elements.*

Axiom 26 *Axiom of union : For any set S there is a set A containing every set that is a member of some member of S.*

Axiom 27 *Axiom schema of replacement : If the domain of a definable function f is a set, and f(x) is a set for any x in that domain, then the range of f is a subclass of a set, subject to a restriction needed to avoid paradoxes.*

Axiom 28 *Axiom of infinity : Let S(x) abbreviate $x \cup \{x\}$, where x is some set. Then there exists a set X such that the empty set is a member of X and, whenever a set y is a member of X, then S(y) is also a member of X.*

More colloquially, there exists a set X having infinitely many members.

Axiom 29 *Axiom of power set : For any set A there is a set, called the power set of A whose elements are all the subsets of A.*

Axiom 30 *Well-ordering theorem : For any set X, there is a binary relation R which well-orders X.*

This means R is an order relation on X such that every non empty subset of X has a member which is minimal under R (see below the definition of order relation).

Axiom 31 *The axiom of choice (AC) : Let X be a set whose members are all non-empty. Then there exists a function f from X to the union of the members of X, called a "choice function", such that for all $Y \in X$ one has $f(Y) \in Y$.*

To tell it plainly : if we have a collection (possibly infinite) of sets, its is always possible to choose an element in each set.

The axiom of choice is equivalent to the Well-ordering theorem, given the other 8 axioms. AC is characterized as non constructive because it asserts the existence of a set of chosen elements, but says nothing about how to choose them.

2.1.2 Extensions

There are several axiomatic extensions of ZFC, which strive to incorporate larger structures without the hindrance of "too large sets". Usually they introduce a distinction between "sets" (ordinary sets) and "classes" or "universes" (which are larger but cannot be part of a set). A universe is comprised of sets, but is not a set itself and does not meet the axioms of sets. This precaution precludes the possibility of defining sets by recursion : any set must be defined before it can be used.

von Neumann organizes sets according to a hierarchy based on ordinal numbers : at each step a set can be added only if all its elements are part of a previous step (starting with $\varnothing$). The final step gives the universe.

"New foundation" (Jensen, Holmes) is another system based on a different hierarchy.

We give below the extension used by Kashiwara and Schapira which is typical of these extensions, and will be used later in categories theory.

A **universe** U is an object satisfying the following properties :

1. $\varnothing \in U$
2. $u \in U \Rightarrow u \subset U$
3. $u \in U \Rightarrow \{u\} \in U$ (the set with the unique element u)
4. $u \in U \Rightarrow 2^u \in U$ (the set of all subsets of u)
5. if for each member of the family (see below) $(u_i)_{i \in I}$ of sets $u_i \in U$ then $\cup_{i \in I} u_i \in U$
6. $\mathbb{N} \in U$

A universe is a "collection of sets", with the implicit restriction that all its elements are known (there is no recursive definition) so that the ususal paradoxes are avoided. As a consequence :

7. $u \in U \Rightarrow \cup_{x \in u} x \in U$

8. $u, v \in U \Rightarrow u \times v \in U$

9. $u \subset v \in U \Rightarrow u \in U$

10.if for each member of the family (see below) of sets $(u_i)_{i \in I}$ $u_i \in U$ then

$$\prod_{i \in I} u_i \in U$$

An axiom is added to the ZFC system : for any set x there exists an universe U such that $x \in U$

A set X is **U-small** if there is a bijection between X and a set of U.

2.1.3 Operations on sets

In formal set theories "x belongs to X" : $x \in X$ is an atom (it is always true or false). In "fuzzy logic" it can be neither.

1. From the previous axioms and this atom are defined the following operators on sets:

Definition 32 *The **Union** of the sets A and B, denoted $A \cup B$, is the set of all objects that are a member of A, or B, or both.*

Definition 33 *The **Intersection** of the sets A and B, denoted $A \cap B$, is the set of all objects that are members of both A and B.*

Definition 34 *The **Set difference** of U and A, denoted $U \setminus A$ is the set of all members of U that are not members of A.*

Example : The set difference {1,2,3} \ {2,3,4} is {1} , while, conversely, the set difference {2,3,4} \ {1,2,3} is {4} .

Definition 35 *A **subset** of a set A is a set B such that all its elements belong to A*

Definition 36 *The **complement** of a subset A with respect to a set U is the set difference $U \setminus A$*

If the choice of U is clear from the context, the notation A^c will be used. Another notation is $\complement_U^A = A^c$

Definition 37 *The **Symmetric difference** of the sets A and B, denoted $A \triangle B = (A \cup B) \setminus (A \cap B)$ is the set of all objects that are a member of exactly one of A and B (elements which are in one of the sets, but not in both).*

Definition 38 *The **Cartesian product** of A and B, denoted A x B, is the set whose members are all possible ordered pairs (a,b) where a is a member of A and b is a member of B.*

The cartesian product of sets can be extended to an infinite number of sets (see below)

Definition 39 *The **Power set** of a set A is the set whose members are all possible subsets of A. It is denoted* 2^A.

Theorem 40 *Union and intersection are associative and distributive*

$$A \cup (B \cup C) = (A \cup B) \cup C$$
$$A \cap (B \cap C) = (A \cap B) \cap C$$
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Theorem 41 *Symmetric difference is commutative, associative and distributive with respect to intersection.*

$$\complement_B^{A \cup B} = (A \cup B)^c = A^c \cap B^c, (A \cap B)^c = A^c \cup B^c$$

Remark : there are more sophisticated operators involving an infinite number of sets (see Measure).

2.2   Maps

2.2.1   Definitions

Definition 42 *A **map** f from a set E to a set F, denoted* $f : E \to F :: y = f(x)$ *is a relation which associates to each element x of E one element y=f(x) of F.*
*x in f(x) is called the **argument**, f(x) is the **value** of f for the argument x.*
*E is the **domain** of f, F is the **codomain** of f.*
The set $f(E) = \{y = f(x), x \in E\}$ *is the **range** (or **image**) of f.*
*The **graph** of f is the set of all ordered pairs* $\{(x, f(x)), x \in E\}$.

(formally one can define the map as the set of pairs (x,f(x)))
We will usually reserve the name "**function**" when the codomain is a field $(\mathbb{R}, \mathbb{C})$.

Definition 43 *The **preimage** (or **inverse image**) of a subset B⊂ F of the map f:E→ F is the subset denoted* $f^{-1}(B) \subset E$ *such that* $\forall x \in f^{-1}(B) : f(x) \in B$

It is usually denoted : $f^{-1}(B) = \{x \in E : f(x) \in B\}$.
Notice that it is not necessary for f to have an inverse map.

Definition 44 *The **restriction** f_A of a map $f : E \to F$ to a subset A⊂ E is the map :* $f_A : A \to F :: \forall x \in A : f_A(x) = f(x)$

Definition 45 *An **embedding** of a subset A of a set E in E is a map $\imath : A \to E$ such that* $\forall x \in A : \imath(x) = x$.

Definition 46 *A **retraction** of a set E on a subset A of E is a map : $\rho : E \to A$ such that : $\forall x \in A, \rho(x) = x$. Then A is said to be a **retract** of E*

Retraction is the converse of an embedding. Usually embedding and retraction maps are morphisms : they preserve the mathematical structures of both A and E, and x could be seen indifferently as an element of A or an element of E.

Example : the embedding of a vector subspace in a vector space.

Definition 47 *The **characteristic function** (or indicator funtion) of the subset A of the set E is the function denoted : $1_A : E \to \{0, 1\}$ with $1_A(x) = 1$ if $x \in A, 1_A(x) = 0$ if $x \notin A$.*

Definition 48 *A set H of maps $f : E \to F$ is said to **separate** E if : $\forall x, y \in E, x \neq y, \exists f \in H : f(x) \neq f(y)$*

Definition 49 *If E,K are sets, F a set of maps : $f : E \to K$ the **evaluation** map at $x \in E$ is the map : $\widehat{x} : F \to K :: \widehat{x}(f) = f(x)$*

This definition, which seems a bit convoluted, is met often with different names.

Definition 50 *Let I be a set, the **Kronecker** function is the function : $\delta : I \times I \to \{0, 1\} :: \delta(i, j) = 1$ if i=j, $\delta(i, j) = 0$ if i≠ j*

When I is a set of indices it is usually denoted $\delta_j^i = \delta(i, j)$ or δ_{ij}.

2. The two following theorems are a consequence of the axioms of the set theory:

Theorem 51 *There is a set, denoted F^E , of all maps with domain E and codomain F*

Theorem 52 *There is a unique map Id_E over a set E, called the **identity**, such that : $Id_E : E \to E :: x = Id_E(x)$*

Maps of several variables

A map f of several variables $(x_1, x_2, ... x_p)$ is just a map with domain the cartesian products of several sets $E_1 \times E_2 ... \times E_p$

From a map $f : E_1 \times E_2 \to F$ one can define a map with one variable by keeping x_1 constant, that we will denote $f(x_1, \cdot) : E_2 \to F$

Definition 53 *The **canonical projection** of $E_1 \times E_2 ... \times E_p$ onto E_k is the map $\pi_k : E_1 \times E_2 ... \times E_p \to E_k :: \pi_k(x_1, x_2, ... x_p) = x_k$*

Definition 54 *A map $f : E \times E \to F$ is **symmetric** if $\forall x_1 \in E, \forall x_2 \in E ::$ $f(x_1, x_2) = f(x_2, x_1)$*

Injective, surjective maps

Definition 55 *A map is **onto** (or **surjective**) if its range is equal to its codomain*

For each element $y \in F$ of the codomain there is at least one element $x \in E$ of the domain such that : $y = f(x)$

Definition 56 *A map is **one-to-one** (or **injective**) if each element of the codomain is mapped at most by one element of the domain*

$$(\forall y \in F : f(x) = f(x') \Rightarrow x = x') \Leftrightarrow (\forall x \neq x' \in E : f(x) \neq f(x'))$$

Definition 57 *A map is **bijective** (or one-one and onto) if it is both onto and one-to-one. If so there is an **inverse map***
$f^{-1} : F \to E :: x = f^{-1}(y) : y = f(x)$

2.2.2 Composition of maps

Definition 58 *The **composition**, denoted $g \circ f$, of the maps $f : E \to F, g : F \to G$ is the map :*
$$g \circ f : E \to G :: x \in E \xrightarrow{f} y = f(x) \in F \xrightarrow{g} z = g(y) = g \circ f(x) \in G$$

Theorem 59 *The composition of maps is always associative : $(f \circ g) \circ h = f \circ (g \circ h)$*

Theorem 60 *The composition of a map $f : E \to E$ with the identity gives f : $f \circ Id_E = Id_E \circ f = f$*

Definition 61 *The **inverse** of a map $f : E \to F$ for the composition of maps is a map denoted $f^{-1} : F \to E$ such that : $f \circ f^{-1} = Id_E, f^{-1} \circ f = Id_F$*

Theorem 62 *A bijective map has an **inverse** map for the composition*

Definition 63 *If the codomain of the map f is included in its domain, the **n-iterated** map of f is the map $f^n = f \circ f... \circ f$ (n times)*

Definition 64 *A map f is said **idempotent** if $f^2 = f \circ f = f$.*

Definition 65 *A map f such that $f^2 = Id$ is an **involution**. If its range is strictly included in its codomain it is a **projection** :*

$$f : E \to F :: f \circ f = Id, f(E) \neq F$$

2.2.3 Sequence

Definition 66 *A **family of elements** of a set E is a map from a set I, called the **index set**, to the set E*

Definition 67 *A **sequence** in the set E is a family of elements of E indexed on the set of natural numbers $\mathbb{N}$.*

Notation 68 $(x_i)_{i \in I} \in E^I$ *is a family of elements of E indexed on I*
$(x_n)_{n \in \mathbb{N}} \in E^{\mathbb{N}}$ *is a sequence of elements in the set E*

Notice that if X is a subset of E then a sequence *in* X is a map $x : \mathbb{N} \to X$

Definition 69 *A **subfamily** of a family of elements is the restriction of the family to a subset of the index set*
*A **subsequence** is the restriction of a sequence to an infinite subset of $\mathbb{N}$.*

The concept of sequence has been generalized to "nets" (Wilansky p.39). A directed set (or a directed preorder or a filtered set) is a non empty set A together with a reflexive and transitive binary relation $\leq$ (that is, a preorder), with the additional property that every pair of elements has an upper bound: in other words, for any a and b in A there must exist c in A with a $\leq$ c and b $\leq$ c. A net is then a map with domain a directed set. $\mathbb{N}$ is a directed set and a sequence is a net.

Definition 70 *On a set E on which an addition has been defined, the **series** (S_n) is the sequence : $S_n = \sum_{p=0}^{n} x_p$ where $(x_n)_{n \in \mathbb{N}} \in E^{\mathbb{N}}$ is a sequence.*

2.2.4 Family of sets

Definition 71 *A **family of sets** $(E_i)_{i \in I}$, over a set E is a map from a set I to the power set of E*

For each argument i E_i is a subset of E : $F : I \to 2^E :: F(i) = E_i$.
The axiom of choice tells that for any family of sets $(E_i)_{i \in I}$ there is a map $f : I \to E$ which associates an element f(i) of E_i to each value i of the index : $\exists f : I \to E :: f(i) \in E_i$
If the sets E_i are not previously defined as subsets of E (they are not related), following the previous axioms of the enlarged set theory, they must belong to a universe U, and then the set $E = \cup_{i \in I} E_i$ also belongs to U and all the E_i are subsets of E.

Definition 72 *The **cartesian product** $E = \prod_{i \in I} E_i$ of a family of sets is the set of all maps : $f : I \to \cup_{i \in I} E_i$ such that $\forall i \in I : f(i) \in E_i$. The elements $f(i)$ are the **components** of f.*

This is the extension of the previous definition to a possibly infinite number of sets.

Definition 73 *A **partition** of a set E is a family* $(E_i)_{i \in I}$, *of sets over E such that :*

   $\forall i : E_i \neq \varnothing$
   $\forall i, j : E_i \cap E_i = \varnothing$
   $\cup_{i \in I} E_i = E$

Definition 74 *A **refinement** $(A_j)_{j \in J}$ of a partition $(E_i)_{i \in I}$ over E is a partition of E such that :* $\forall j \in J, \exists i \in I : A_j \subset E_i$

Definition 75 *A **family of filters** over a set E is a family $(F_i)_{i \in I}$ over E such that :*

   $\forall i : F_i \neq \varnothing$
   $\forall i, j : \exists k \in I : F_k \subset F_i \cap F_j$

For instance the **Fréchet filter** is the family over $\mathbb{N}$ defined by : $F_n = \{p \in \mathbb{N} : p \geq n\}$

2.3   Binary relations

2.3.1   Definitions

Definition 76 *A **binary relation** R on a set E is a 2 variables propositional function :* $R : E \times E \to \{T, F\}$

Definition 77 *A binary relation R on the set E is :*
   reflexive *if :* $\forall x \in E : R(x, x) = T$
   symmetric *if :* $\forall x, y \in E : R(x, y) \sim R(y, x)$
   antisymmetric *if :* $\forall x, y \in E : (R(x, y) \wedge R(y, x)) \Rightarrow x = y$
   transitive *if :* $\forall x, y, z \in E : (R(x, y) \wedge R(y, z)) \Rightarrow R(x, z)$
   total *if* $\vDash \forall x \in E, \forall y \in E, (R(x, y) \vee R(y, x))$

2.3.2   Equivalence relation

Definition 78 *An **equivalence relation** is a binary relation which is reflexive, symmetric and transitive*

It will be usually denoted by $\sim$

Definition 79 *If R is an equivalence relation on the set E,*
   *- the **class of equivalence** of an element x$\in$ E is the subset denoted* $[x]$ *of elements y$\in$ E such that y $\sim$ x .*
   *- the **quotient set** denoted E/R is the partition of E whose elements are the classes of equivalence of E.*

Theorem 80 *There is a natural bijection from the set of all possible equivalence relations on E to the set of all partitions of E.*

So, if E is finite set with n elements, the number of possible equivalence relations on E equals the number of distinct partitions of E, which is the nth **Bell number** : $B_n = \frac{1}{e} \sum_{k=0}^{\infty} \frac{k^n}{k!}$

Example : for any map $f : E \rightarrow F$ the relation $x \sim y$ if f(x)=f(y) is an equivalence relation.

2.3.3 Order relation

Definition 81 *An **order relation** is a binary relation which is reflexive, anti-symmetric and transitive. If the relation is total it is a **total ordering** , if not the relation is a **partial ordering** (or preorder).*

In a partial ordering there are couples (x,y) for which R(x,y) is not defined.

Example : the relation $\leq$ is a total ordering over $\mathbb{R}$, but is only a partial ordering over $\mathbb{R}^2$

An antisymmetric relation gives 2 dual binary relations ("greater or equal than" and "smaller or equal than").

Bounds

Definition 82 *An **upper bound** of a subset A of E is an element of E which is greater than all the elements of A*

Definition 83 *A **lower bound** of a subset A of E is an element of E which is smaller than all the elements of A*

Definition 84 *A **bounded subset** A of E is a subset which has both an upper bound and a lower bound.*

Definition 85 *A **maximum** of a subset A of E is an element of A which is also an upper bound for A*

$m = \max A \Leftrightarrow m \in E, \forall x \in A : m \geq x$

Definition 86 *A **minimum** of a subset A of E is an element of A which is also a lower bound for A*

$m = \min A \Leftrightarrow m \in E, \forall x \in A : m \leq x$
Maximum and minimum, if they exist, are unique.

Definition 87 *If the set of upper bounds has a minimum, this element is unique and is called the **least upper bound** or supremum*

denoted : $s = \sup A = \min\{m \in E : \forall x \in E : m \geq x\}$

Definition 88 *If the set of lower bounds has a maximum, this element is unique and is called the **greatest lower bound** or infimum.*

denoted $:s = \inf A = \max\{m \in E : \forall x \in E : m \leq x\}$

Theorem 89 *Over $\mathbb{R}$ any non empty subset which has an upper bound has a least upper bound, and any non empty subset which has a lower bound has a greatest lower bound,*

If $f : E \to \mathbb{R}$ is a real valued function, a maximum of f is an element M of E such that f(M) is a maximum of f(E), and a mimimum of f is an element m of E such that f(m) is a minimum of f(E)

Axiom 90 *Zorn lemna : if E is a set with a partial ordering, such that any subset for which the order is total has a least upper bound, then E has also a maximum.*

The Zorn lemna is equivalent to the axiom of choice.

Definition 91 *A set is **well-ordered** if it is totally ordered and if any non empty subset has a minimum. Equivalently if there is no infinite decreasing sequence.*

It is then possible to associate each element with an ordinal number (see below). The axiom of choice is equivalent to the statement that every set can be well-ordered.

As a consequence let I be any set. Thus for any finite subset J of I it is possible to order the elements of J and one can write $J=\{j_1, j_2, ...j_n\}$ with n=card(J).

Definition 92 *A **lattice** is a partially ordered set (also called a poset) in which any two elements have a unique supremum (the least upper bound, called their join) and an infimum (greatest lower bound, called their meet).*

Example : For any set A, the collection of all subsets of A can be ordered via subset inclusion to obtain a lattice bounded by A itself and the null set. Set intersection and union interpret meet and join, respectively.

Definition 93 *A **monotone** map $f : E \to F$ between sets E,F endowed with an ordering is a map which preserves the ordering:*

$\forall x, y \in E, x \leq_E y \Rightarrow f(x) \leq_F f(y)$
The converse of such a map is an **order-reflecting** map :
$\forall x, y \in E, f(x) \leq_F f(y) \Rightarrow x \leq_E y$

2.3.4   Cardinality

Theorem 94 *Bernstein (Schwartz I p.23) For any two sets E,F, either there is an injective map $f : E \to F$,or there is an injective map $g : F \to E$. If there is an injective map $f : E \to F$ and an injective map $: g : F \to E$ then there is a bijective map $:\varphi : E \to F, \varphi^{-1} : F \to E$*

Cardinal numbers

The binary relation between sets E,F : "there is a bijection between E and F" is an equivalence relation.

Definition 95 *Two sets have the same **cardinality** if there is a bijection between them.*

The cardinality of a set is represented by a **cardinal number**. It will be denoted card(E) or $\#E$.

The cardinal of $\varnothing$ is 0.

The cardinal of any finite set is the number of its elements.

The cardinal of the set of natural numbers $\mathbb{N}$, of algebric numbers $\mathbb{Z}$ and of rational numbers $\mathbb{Q}$ is $\aleph_0$ (aleph null: hebraïc letter).

The cardinal of the set of the subsets of E (its power set 2^E) is $2^{card(E)}$

The cardinal of $\mathbb{R}$ (and $\mathbb{C}$, and more generally $\mathbb{R}^n, n \in \mathbb{N}$) is $c = 2^{\aleph_0}$, called the **cardinality of the continuum**

It can be proven that : $c^{\aleph_0} = c, c^c = 2^c$

Infinite cardinals

The binary relation between sets E,F : "there is an injection from E to F" is an ordering relation. The cardinality of E is said to be smaller than the cardinality of F if there is no injection from F to E. So it is possible to order the classes of equivalence = the cardinal numbers.

Definition 96 *A set is **finite** if its cardinality is smaller than $\aleph_0$*

*A set is **countably infinite** if its cardinality is equal to $\aleph_0$*

*A set is **uncountable** if its cardinality is greater than $\aleph_0$.*

The cardinals equal or greater than $\aleph_0$ are the **transfinite cardinal numbers**. The **continuum hypothesis** is the assumption that there is no cardinal number between $\aleph_0$ and $2^{\aleph_0}$. Depending of the formal system used for set theory it can be an axiom (as in ZFC), to be added or not to the system (Cohen 1963), or an hypothesis (to be proven true or false) .

Theorem 97 *A set E is infinite iff there is bijective map between E and a subset of E distinct of E.*

This theorem has different interpretations (the "Dedekind infinite") according the the set theory used.

2.3.5 Ordinality

Definition

Cardinality is the number of elements of a set. Ordinality is related to the possibility to order them in an increasing sequence.

Definition 98 *Two totally ordered sets E and F are of the same **order type** (or ordinality) if there is a bijection $f : E \to F$ such that f and f^{-1} are order preserving maps.*

The relation "E and F are of the same order type" is an equivalence relation.

Ordinal numbers

The ordinality of a totally ordered set is represented by an **ordinal number**.

The sets ot the same order type have the same cardinality but the converse is not always true.

For finite sets the ordinal number is equal to the cardinal number.

For infinite sets, the **transfinite ordinal numbers** are not the same as the transfinite cardinal numbers.

The order type of the natural integers $\mathbb{N}$ is the first transfinite ordinal number, denoted ω which can be identified with $\aleph_0$

The next ordinal number following the transfinite ordinal number α is denoted $\alpha + 1$.

Whereas there is only one countably infinite cardinal, namely $\aleph_0$ itself, there are uncountably many countably infinite ordinals, namely

$\aleph_0$, $\aleph_0 + 1$, ..., $\aleph_0 \cdot 2$, $\aleph_0 \cdot 2 + 1$, ..., $\aleph_0 2$, ..., $\aleph_0 3$, ..., $\aleph_0 \aleph_0$, ..., $\aleph_0 \aleph_0 \aleph_0$, ...

Here addition and multiplication are not commutative: in particular $1 + \aleph_0$ is $\aleph_0$ rather than $\aleph_0 + 1$ and likewise, $2 \cdot \aleph_0$ is $\aleph_0$ rather than $\aleph_0 \cdot 2$. The set of all countable ordinals constitutes the first uncountable ordinal ω_1, which is identified with the next cardinal after $\aleph_0$.

The order type of the rational numbers $\mathbb{Q}$ is the transfinite ordinal number denoted η. Any countable totally ordered set can be mapped injectively into the rational numbers in an order-preserving way.

Transfinite induction

Transfinite induction is the following logical rule of inference (which is always valid):

Axiom 99 *For any well-ordered set, any property that passes from the set of ordinals smaller than a given ordinal α to α itself, is true of all ordinals : if $P(\alpha)$ is true whenever $P(\beta)$ is true for all $\beta < \alpha$, then $P(\alpha)$ is true for all α.*

3 CATEGORIES

Categories is now a mandatory part of advanced mathematics, at an almost egal footing as the set theory. However it is one of the most abstract mathematical theories. It requires the minimum of properties from its objects, so it provides a nice language to describe many usual mathematical objects in a unifying way. It is also a powerful tool in some specialized fields. But the drawback is that it leads quickly to very convoluted and abstract constructions when dealing with precise subjects, that border mathematical pedantism, without much added value. So, in this book, we use it when, and only when, it is really helpful and the presentation is limited to the main definitions and principles, in short to the vocabulary needed to understand what lies behind the language.

On this topic we follow mainly Lane and Kashiwara.

3.1 Categories

In mathematics, whenever a set is endowed with some structure, there are some maps, meeting properties matching those of the structure of the set, which are of special interest : the continuous maps with topological space, the linear maps with vector space,...The basic idea is to consider packages, called categories, including both sets and their related maps.

All the definitions and results presented here, which are quite general, can be found in Kashirawa-Shapira or Lane

3.1.1 Definitions

Definition 100 *A **category** C consists of the following data:*
*- a set $Ob(C)$ of **objects***
*- for each ordered pair (X,Y) of objects of Ob(C), a set of **morphisms** hom(X,Y) from the domain X to the codomain Y:*
$$\forall (X,Y) \in Ob(C) \times Ob(C), \exists \hom(X,Y) = \{f, dom\,(f) = X, codom\,(f) = Y\}$$
- a function ∘ called composition between morphisms :
$$\circ : \hom(X,Y) \times \hom(Y,Z) \to \hom(X,Z)$$
which must satisfy the following conditions :
Associativity
$$f \in \hom(X,Y), g \in \hom(Y,Z), h \in \hom(Z,T) \Rightarrow (f \circ g) \circ h = f \circ (g \circ h)$$
Existence of an identity morphism for each object
$$\forall X \in Ob(C)\ , \exists id_X \in \hom(X,X) : \forall f \in \hom(X,Y) : f \circ id_X = f, \forall g \in \hom(Y,X) : id_X \circ g = g$$

If Ob(C) is a set of a universe U (therefore all the objects belong also to U), and if for all objects the set $\hom(A,B)$ is isomorphic to a set of U then the category is said to be a "U-small category". Here "isomorphic" means that there is a bijective map which is also a morphism.

Remarks :

i) When it is necessary to identify the category one denotes $\hom_C(X, Y)$ for $\hom(X, Y)$

ii) The use of "universe" is necessary as in categories it is easy to run into the problems of "too large sets".

iii) To be consistent with some definitions one shall assume that the set of morphisms from one object A to another object B can be empty.

iv) A morphism is not necessarily a map $f : X \to X'$. Let U be a universe of sets (the sets are known), C the category defined as : objects = sets in U, morphisms : $\hom_C(X, Y) = \{X \sqsubseteq Y\}$ meaning the logical proposition $X \sqsubseteq Y$ which is either true of false. One can check that it meets the conditions to define a category.

As such, the definition of a category brings nothing new to the usual axioms and definitions of set theory. The concept of category is useful when all the objects are endowed with some specific structure and the morphisms are the specific maps related to this structure: we have the category of "sets", "vector spaces", "manifolds",..It is similar to set theory : one can use many properties of sets without telling what are the elements of the set. The term "morphism" refer to the specific maps used in the definition of the category.

The concept of morphism is made precise in the language of categories, but, as a rule, we will always reserve the name morphism for maps between sets endowed with similar structures which "conserve" these structures. And similarly isomorphism for bijective morphism.

Examples

1. For a given universe U the category U-set is the category with objects the sets of U and morphisms any map between sets of Ob(U-set). It is necessary to fix a universe because there is no "Set of sets".

2. 0 is the empty category with no objects or morphisms

3. The category $\mathfrak{V}$ of "vector spaces over a field K" : the objects are vector spaces, the morphisms linear maps

4. The category of "topological spaces" (often denoted "Top") : the objects are topological spaces, the morphisms continuous maps

5. The category of "smooth manifolds" : the objects are smooth manifolds, the morphisms smooth maps

Notice that the morphisms must meet the axioms (so one has to prove that the composition of linear maps is a linear map). The manifolds and differentiable maps are not a category as a manifold can be continuous but not differentiable. The vector spaces over $\mathbb{R}$ (resp.$\mathbb{C}$) are categories but the vector spaces (over any field) are not a category as the product of a R-linear map and a C-linear map is not a C-linear map.

6. A **monoid** is a category with one unique object and a single morphism (the identity).. It is similar to a set M, a binary relation MxM associative with unitary element (semi group).

7. A **simplicial category** has objects indexed on ordinal numbers and morphisms are order preserving maps.

More generally the category of ordered sets with objects = ordered sets belonging to a universe, morphisms = order preserving maps.

3.1.2 Additional definitions about categories

Definition 101 *A **subcategory** C' of the category C has for objects $Ob(C') \subset Ob(C)$ and for $X, Y \in C'$, $\hom_{C'}(X, Y) \subset \hom_C(X, Y)$*
*A subcategory is **full** if $\hom_{C'}(X, Y) = \hom_C(X, Y)$*

Definition 102 *If C is a category, the **opposite category**, denoted C*, has the same objects as C and for morphisms :*
$\hom_{C^*}(X, Y) = \hom_C(Y, X)$
with the composition :
$f \in \hom_{C^*}(X, Y), g \in \hom_{C^*}(Y, Z) : g \circ^* f = f \circ g$

Definition 103 *A category is*
   *- **discrete** if all the morphisms are the identity morphisms*
   *- **finite** it the set of objects and the set of morphisms are finite*
   *- **connected** if it is non empty and for any pair X,Y d'objects there is a finite sequence d'objects $X_0 = X, X_1, ..X_{n-1}, X_n = Y$ such that $\forall i \in [0, n-1]$ at least one of the sets $\hom(X_i, X_{i+1}), \hom(X_{i+1}, X_i)$ is non empty.*

Definition 104 *If $(C_i)_{i \in I}$ is a family of categories indexed by the set I*
   *the **product category** $\prod_{i \in I} C_i$ has*

   - for objects : $Ob\left(\prod_{i \in I} C_i\right) = \prod_{i \in I} Ob(C_i)$

   - for morphisms : $\hom_{\prod_{i \in I} C_i}\left(\prod_{j \in I} X_j, \prod_{j \in I} Y_j\right) = \prod_{j \in I} \hom_{C_j}(X_j, Y_j)$

   *the **disjoint union category** $\sqcup_{i \in I} C_i$ has*
   - for objects : $Ob(\sqcup C_i) = \{(X_i, i), i \in I, X_i \in Ob(C_i)\}$
   - for morphisms : $\hom_{\sqcup C_i}((X_j, j), (Y_k, k)) = \hom_{C_j}(X_j, Y_j)$ *if j=k; $=\varnothing$ if* $j \neq k$

Definition 105 *A **pointed category** is a category with the following properties :*
   - each object X is a set and there is a unique $x \in X$ (called base point) which is singled : let $x = \imath(X)$
   - there are morphisms which preserve x : $\exists f \in \hom(X, Y) : \imath(Y) = f(\imath(X))$

Example : the category of vector spaces over a field K with a basis and linear maps which preserve the basis.

3.1.3   Initial and terminal objects

Definition 106 *An object I is **initial** in the category C if $\forall X \in Ob(C), \# \hom(I, X) = 1$*

   meaning that there is only one morphism going from I to X

Definition 107 *An object T is **terminal** in the category C if $\forall X \in Ob(C), \# \hom(X, T) = 1$*

   meaning that there is only one morphism going from X to T

Definition 108 *An object is **null** (or zero object) in the category C if it is both initial and terminal.*

   It is usually denoted 0. So if there is a null object, $\forall X, Y$ there is a morphism $X \to Y$ given by the composition : $X \to 0 \to Y$
   In the category of groups the null object is the group 1, comprised of the unity.
   Example : define the pointed category of n dimensional vector spaces over a field K, with an identified basis:
   - objects : E any n dimensional vector space over a field K, with a singled basis $(e_i)_{i=1}^{n}$
   - morphisms: $\hom(E, F) = L(E; F)$ (there is always a linear map F : $f(e_i) = f_i$)
   All the objects are null : the morphisms from E to F such that $f(e_i) = f_i$ are unique

3.1.4   Morphisms

Basic definitions
   1. The following definitions generalize, in the language of categories, concepts which have been around for a long time for strctures such as vector spaces, topological spaces,...

Definition 109 *An **endomorphism** is a morphism in a category with domain = codomain : $f \in \hom(X, X)$*

Definition 110 *If $f \in \hom(X, Y), g \in \hom(Y, X)$ such that : $f \circ g = Id_Y$ then f is the **left-inverse** of g, and g is the **right-inverse** of f*

Definition 111 *A morphism $f \in \hom(X, Y)$ is an **isomorphism** if there exists $g \in \hom(Y, X)$ such that $f \circ g = Id_Y, g \circ f = Id_X$*

   Then the two objects X,Y are said to be isomorphic and we is usually denoted : $X \simeq Y$

Definition 112 *An **automorphism** is an endomorphism which is also an isomorphism*

Definition 113 *A category is a **groupoid** if all its morphisms are isomorphisms*

2. The following definitions are specific to categories.

Definition 114 *Two morphisms in a category are **parallel** if they have same domain and same codomain. They are denoted : $f, g : X \rightrightarrows Y$*

Definition 115 *A **monomorphism** $f \in \hom(X, Y)$ is a morphism such that for any pair of parallel morphisms :*
   $g_1, g_2 \in \hom(Z, X) : f \circ g_1 = f \circ g_2 \Rightarrow g_1 = g_2$

Which can be interpreted as f has a left-inverse and so is an injective morphism

Definition 116 *An **epimorphism** $f \in \hom(X, Y)$ is a morphism such that for any pair of parallel morphisms :*
   $g_1, g_2 \in \hom(Y, Z) : g_1 \circ f = g_2 \circ f \Rightarrow g_1 = g_2$

Which can be interpreted as f has a right-inverse and so is a surjective morphism

Theorem 117 *If $f \in \hom(X, Y), g \in \hom(Y, Z)$ and f,g are monomorphisms (resp.epimorphisms, isomorphisms) then $g \circ f$ is a monomorphism (resp.epimorphism, isomorphism)*

Theorem 118 *The morphisms of a category C are a category denoted $\hom(C)$*
   - Its objects are the morphisms in $C : Ob(\hom(C)) = \{\hom_C(X, Y), X, Y \in Ob(C)\}$
   - Its morphisms are the maps u,v such that :
   $\forall X, Y, X', Y \in Ob(C), \forall f \in \hom(X, Y), g \in \hom(X', Y) : u \in \hom(X, X'), v \in \hom(Y, Y') : v \circ f = g \circ u$

The maps u,v must share the general characteristics of the maps in C

Diagrams
   Category theory uses diagrams quite often, to describe, by arrows and symbols, morphisms or maps between sets. A diagram is **commutative** if any path following the arrows is well defined (in terms of morphisms).
   Example : the following diagram is commutative :

$$X \xrightarrow{f} Y$$
$$u \downarrow \quad \downarrow v$$
$$Z \xrightarrow{g} T$$

means :
$g \circ u = v \circ f$

36

Exact sequence

Used quite often with a very abstract definition, which gives, in plain language:

Definition 119 *For a family* $(X_p)_{p \leq n}$ *of objects of a category C and of morphisms* $f_p \in \hom_C (X_p, X_{p+1})$

the sequence : $X_0 \xrightarrow{f_0} X_1..X_p \xrightarrow{f_p} X_{p+1}.... \xrightarrow{f_{n-1}} X_n$ *is* **exact** *if* $f_p(X_p) = \ker(f_{p+1})$

An exact sequence is also called a **complex**. It can be infinite.

That requires to give some meaning to ker. In the usual cases ker may be understood as the subset :

if the X_p are groups : $\ker f_p = \left\{ x \in X_p, f_p(x) = 1_{X_{p+1}} \right\}$ so $f_p \circ f_{p-1} = 1$

if the X_p are vector spaces : $\ker f_p = \left\{ x \in X_p, f_p(x) = 0_{X_{p+1}} \right\}$ so $f_p \circ f_{p-1} = 0$

Definition 120 *A* **short exact sequence** *in a category C is :* $X \xrightarrow{f} Y \xrightarrow{g} Z$ *where :* $f \in \hom_C(X,Y)$ *is a monomorphism (injective)* , $g \in \hom_C(Y,Z)$ *is an epimorphism (surjective), equivalently iff* $f \circ g$ *is an isomorphism.*

Then Y is, in some way, the product of Z and f(X)

it is usally written :

$0 \to X \xrightarrow{f} Y \xrightarrow{g} Z \to 0$ for abelian groups or vector spaces

$1 \to X \xrightarrow{f} Y \xrightarrow{g} Z \to 1$ for the other groups

A short exact sequence $X \xrightarrow{f} Y \xrightarrow{g} Z$ **splits** if either :

$\exists t \in \hom_C(Y,X) :: t \circ f = Id_X$

$$0 \quad \to \quad X \quad \xrightarrow{f} \quad Y \quad \xrightarrow{g} \quad Z$$
$$\xleftarrow{t}$$

of $\exists u \in \hom_C(Z,Y) :: g \circ u = Id_Z$

$$0 \quad \to \quad X \quad \xrightarrow{f} \quad Y \quad \xrightarrow{g} \quad Z$$
$$\xleftarrow{u}$$

then :

- for abelian groups or vector spaces : $Y = X \oplus Z$

- for other groups (semi direct product) : $Y = X \ltimes Z$

3.2 Functors

Functors are roughly maps between categories. They are used to import structures from a category C to a category C', using a general procedure so that some properties can be extended immediately. Example : the functor which associates to each vector space its tensorial algebra. There are more elaborate examples in the following parts.

3.2.1 Functors

Definition 121 *A **functor** (a covariant functor) F between the categories C and C' is :*

a map $F_o : Ob(C) \to Ob(C')$

maps $F_m : \hom(C) \to \hom(C') :: f \in \hom_C(X,Y) \to F_m(f) \in \hom_{C'}(F_o(X), F_o(Y))$

such that

$F_m(Id_X) = Id_{F_o(X)}$

$F_m(g \circ f) = F_m(g) \circ F_m(f)$

Definition 122 *A **contravariant functor** F between the categories C and C' is :*

a map $F_o : Ob(C) \to Ob(C')$

maps $F_m : \hom(C) \to \hom(C') :: f \in \hom_C(X,Y) \to F_m(f) \in \hom_{C'}(F_o(X), F_o(Y))$

such that

$F_m(Id_X) = Id_{F_o(X)}$

$F_m(g \circ f) = F_m(f) \circ F_m(g)$

Notation 123 *$F : C \mapsto C'$ (with the arrow $\mapsto$) is a functor F between the categories C,C'*

Example : the functor which associes to each vector space its dual and to each linear map its transpose is a functor from the category of vector spaces over a field K to itself.

So a contravariant functor is a covariant functor $C^* \mapsto C'^*$.

A functor F induces a functor : $F^* : C^* \mapsto C'^*$

A functor $F : C \mapsto Set$ is said to be **forgetful** (the underlying structure in C is lost).

Definition 124 *A **constant functor** denoted $\Delta_X : I \mapsto C$ between the categories I,C, where $X \in Ob(C)$ is the functor :*

$\forall i \in Ob(I) : (\Delta_X)_o(i) = X$

$\forall i,j \in Ob(I), \forall f \in \hom_I(i,j) : (\Delta_X)_m(f) = Id_X$

Composition of functors

Functors can be composed :

$F : C \mapsto C', F' : C' \mapsto C''$

$F \circ F' : C \mapsto C'' :: (F \circ F')_o = F_o \circ F'_o; (F \circ F')_m = F_m \circ F'_m$

The composition of functors is associative whenever it is defined.

Definition 125 *A functor F is **faithful** if $F_m : \hom_C(X,Y) \to \hom_{C'}(F_o(X), F_o(Y))$ is injective*

Definition 126 *A functor F is **full** if $F_m : \hom_C(X,Y) \to \hom_{C'}(F_o(X), F_o(Y))$ is surjective*

Definition 127 *A functor F is **fully faithful** if $F_m : \hom_C(X,Y) \to \hom_{C'}(F_o(X), F_o(Y))$ is bijective*

These 3 properties are closed by composition of functors.

Theorem 128 *If $F : C \mapsto C'$ is faithful, and if F(f) with $f \in \hom_C (X, Y)$ is an epimorphism (resp. a monomorphism) then f is an epimorphism (resp. a monomorphism)*

Product of functors

One defines naturally the product of functors. A **bifunctor** $F : C \times C' \mapsto C''$ is a functor defined over the product CxC', so that for any fixed $X \in C, X' \in C'$ $F(X, .), F(., X')$ are functors

If C,C' are categories, CxC' their product, the right and left **projections** are functors defined obviously :
$L_o (X \times X') = X; R_o (X, X') = X'$
$L_m (f \times f') = f; R_m (f, f') = f'$

They have the universal property : whatever the category D, the functors $F : D \mapsto C; F' : D \mapsto C'$ there is a unique functor $G : D \mapsto C \times C'$ such that $L \circ G = F, R \circ G = F'$

3.2.2   Natural transformation

A natural transformation is a map between functors. The concept is mostly used to give the conditions that a map must meet to be consistent with structures over two categories.

Definition 129 *Let F,G be two functors from the categories C to C'. A **natural transformation** ϕ (also called a morphism of functors) denoted $\phi : F \hookrightarrow G$ is a map : $\phi : Ob(C) \to \hom_{C'} (Ob(C'), Ob(C'))$ such that the following diagram commutes :*

$$
\begin{array}{cccc}
\text{C} & & \text{C'} & \\
\ulcorner \quad \quad \urcorner & \ulcorner & & \urcorner \\
\text{X} & F_o (X) & \overset{\phi(X)}{\to} & G_o(X) \\
\downarrow & \downarrow & & \downarrow \\
f \quad \downarrow & \downarrow F_m (f) & & \downarrow G_m (f) \\
\downarrow & \downarrow & & \downarrow \\
\text{Y} & F_o(Y) & \overset{\phi(Y)}{\to} & G_o(Y)
\end{array}
$$

$\forall X, Y \in Ob (C), \forall f \in \hom_C (X, Y) :$
$G_m (f) \circ \phi (X) = \phi (Y) \circ F_m (f) \in \hom_{C'} (F_o (X), G_o (Y))$
$F_m (f) \in \hom_{C'} (F_o (X), F_o(Y))$
$G_m (f) \in \hom_{C'} (G_o (X), G_o(Y))$
$\phi (X) \in \hom_{C'} (F_o (X), G_o(X))$
$\phi (Y) \in \hom_{C'} (F_o (Y), G_o(Y))$
The components of the transformation are the maps $\phi (X), \phi (Y)$

If $\forall X \in Ob(C)$ $\phi(X)$ is inversible then the functors are said to be **equivalent**.

Natural transformations can be composed in the obvious way. Thus :

Theorem 130 *The set of functors from a category C to a category C' is itself a category denoted $Fc(C, C')$. Its objects are $Ob(Fc(C, C'))$ any functor $F : C \mapsto C'$ and its morphisms are natural transformations : $\hom(F_1, F_2) = \{\phi : F_1 \hookrightarrow F_2\}$*

3.2.3 Yoneda lemna

(Kashirawa p.23)

Let U be a universe, C a category such that all its objects belong to U, and USet the category of all sets belonging to U and their morphisms.

Let :

Y be the category of contravariant functors $C \mapsto USet$

Y* be the category of contravariant functors $C \mapsto USet^*$

h_C be the functor : $h_C : C \mapsto Y$ defined by : $h_C(X) = \hom_C(-, X)$. To an object X of C it associes all the morphisms of C whith codomain X and domain any set of U.

k_C be the functor : $k_C : C \mapsto Y^*$ defined by : $k_C(X) = \hom_C(F, -)$. To an object X of C it associes all the morphisms of C whith domain X and codomain any set of U.

So : $Y = Fc(C^*, USet), Y^* = Fc(C^*, USet^*)$

Theorem 131 *Yoneda Lemna*

i) For F$\in$ Y, $X \in C$: $\hom_Y(h_C(X), F) \simeq F(X)$

ii) For G$\in Y^*$, $X \in C$: $\hom_{Y^*}(k_C(X), G) \simeq G(X)$

Moreover these isomorphisms are functorial with respect to X,F,G : they define isomorphisms of functors from C*xY to USet and from Y**xC to USet.

Theorem 132 *The two functors h_C, k_C are fully faithful*

These abstracts definitions are the basis of the theory of representation of a category. For instance, if G is a group and E a vector space, a representation of G over E is a map f to the set L(E;E) of linear maps on E such that f(gh)=f(g)f(h). The group structure of G is transfered into E through an endomorphism over E. The inversible endomorphisms over a vector space have a group structure with composition law, so a vector space can be included in the category of groups with these morphisms. What the Yoneda lemma says is that to represent G we need to consider a larger category (of sets) and find a set E and a map from G to morphisms over E.

Theorem 133 *A contravariant functor $F : C \mapsto Uset$ is representable if there are an object X of C, called a representative of F, and an isomorphism $h_C(X) \hookrightarrow F$*

Theorem 134 *A covavariant functor $F : C \mapsto Uset$ is representable if there are an object X of C, called a representative of F, and an isomorphism $k_C(X) \hookrightarrow F$*

3.2.4 Universal functors

Many objects in mathematics are defined through an "universal property" (tensor product, Clifford algebra,...) which can be restated in the language of categories. It gives the following.

Let : $F : C \mapsto C'$ be a functor and X' an object of C'

1. An **initial morphism** from X' to F is a pair $(A, \phi) \in Ob\,(C) \times \hom_{C'}(X', F_o\,(A))$ such that :

$\forall X \in Ob\,(C)\,, f \in \hom_{C'}(X', F_o\,(X))\,, \exists g \in \hom_C\,(A, X) : f = F_m\,(g) \circ \phi$

The key point is that g must be unique, then A is unique up to isomorphism

$$
\begin{array}{ccc}
\mathbf{X'} \to\to \phi \to F_o(A) & \quad & \mathbf{A} \\
\searrow \qquad \downarrow & & \downarrow \\
\quad f \qquad \downarrow F_m(g) & & \downarrow g \\
\searrow \quad \downarrow & & \downarrow \\
F_o(X) & & X
\end{array}
$$

2. A **terminal morphism** from X' to F is a pair $(A, \phi) \in Ob\,(C) \times \hom_{C'}(F_o\,(A), X')$ such that :

$\forall X \in Ob(X), f \in \hom_{C'}(F(X); X')\,, \exists g \in \hom_C\,(X, A) : f = \phi \circ F_m\,(g)$

The key point is that g must be unique, then A is unique up to isomorphism

$$
\begin{array}{ccc}
X & F_o(X) & \\
\downarrow & \downarrow \quad \searrow & \\
g \quad F_m(g) \downarrow & \quad f & \\
\downarrow & \downarrow \quad \searrow & \\
\mathbf{A} & F_o(A) \to \phi \to \mathbf{X'} &
\end{array}
$$

3. **Universal morphism** usually refers to initial morphism.

Part II
PART2 : ALGEBRA

Given a set, the theory of sets provides only a limited number of tools. To go further one adds "mathematical structures" on sets, meaning operations, special collection of sets, maps...which become the playing ground of mathematicians.

Algebra is the branch of mathematics that deals with structures defined by operations between elements of sets. An algebraic structure consists of one or more sets closed under one or more operations, satisfying some axioms. The same set can be given different algebraic structures. Abstract algebra is primarily the study of algebraic structures and their properties.

To differentiate algebra from other branches of mathematics, one can say that in algebra there is no concepts of limits or "proximity" such that are defined by topology.

We will give a long list of definitions of all the basic objects of common use, and more detailed (but still schematic) study of groups (there is a part dedicated to Lie groups and Lie algebras) and a detailed study of vector spaces and Clifford algebras, as they are fundamental for the rest of the book.

4 USUAL ALGEBRAIC STRUCTURES

We list here the most common algebraic structures, mostly their definitions. Groups and vector spaces will be reviewed in more details in the next sections.

4.0.5 Operations

Definition 135 *An **operation** over a set A is a map : $\cdot : A \times A \to A :: x \cdot y = z$*
It is :
*- **associative** if $\forall x, y, z \in A : (x \cdot y) \cdot z = x \cdot (y \cdot z)$*
*- **commutative** if $\forall x, y \in A : x \cdot y = y \cdot x$*

Definition 136 *An element e of a set A is an **identity element** for the operation $\cdot$ if : $\forall x \in A : e \cdot x = x \cdot e = x$*
An element x of a set A is a :
*- **right-inverse** of y for the operation $\cdot$ if : $y \cdot x = e$*
*- **left-inverse** of y for the operation $\cdot$ if : $x \cdot y = e$*
*- is **invertible** if it has a right-inverse and a left-inverse (which then are necessarily equal and called **inverse**)*

Definition 137 *If there are two operations denoted $+$ and $*$ on the same set A, then $*$ is **distributive** over (say also distributes over) $+$ if: $\forall x, y, z \in A : x * (y + z) = (x * y) + (x * z), (y + z) * x = (y * x) + (z * x)$*

Definition 138 *An operation $\cdot$ on a set A is said to be **closed** in a subset B of A if $\forall x, y \in B : x \cdot y \in B$*

If E and F are sets endowed with the operations $\cdot, *$ the product set ExF is endowed with an operation in the obvious way :
$(x, x') \,\hat{}\, (y, y') = (x \cdot y, x' * y')$

4.1 From Monoid to fields

Definition 139 *A **monoid** is a set endowed with an associative operation for which it has an identity element*

but its elements have not necessarily an inverse.
Classical monoids :
$\mathbb{N}$: natural integers with addition
$\mathbb{Z}$: the algebraic integers with multiplication
the square nxn matrices with multiplication

4.1.1 Group

Definition 140 *A **group** $(G, \cdot)$ is a set endowed G with an associative operation $\cdot$, for which there is an identity element and every element has an inverse.*

Theorem 141 *In a group, the identity element is unique. The inverse of an element is unique.*

Definition 142 *A commutative (or **abelian**) group is a group with a commutative operation*

Notation 143 *$+$ denotes the operation in a commutative group*
0 denotes the identity element in a commutative group
- x denotes the inverse of x in a commutative group
1 (or 1_G) denotes the identity element in a (non commutative) group G
x^{-1} denotes the inverse of x in a a (non commutative) group G

Classical groups (see the list of classical linear groups in "Lie groups"):
$\mathbb{Z}$: the algebraic integers with addition
$\mathbb{Z}/k\mathbb{Z}$:the algebraic integers multiples of $k \in \mathbb{Z}$ with addition
the mxp matrices with addition
$\mathbb{Q}$: rational numbers with addition and multiplication
$\mathbb{R}$: real numbers with addition and multiplication
$\mathbb{C}$: complex numbers with addition and multiplication
The trivial group is the group denoted $\{1\}$ with only one element.
A group G is a category, with Ob=the unique element G and morphisms $\hom(G, G)$

4.1.2 Ring

Definition 144 *A **ring** is a set endowed with two operations : one called addition, denoted + for which it is an abelian group, the other denoted · for which it is a monoid, and · is distributive over +.*

Remark : some authors do not require the existence of an identity element for · and then call unital ring a ring with an identity element for · .

If 0=1 (the identity element for + is also the identity element for ·) the ring has only one element, said 1 and is called a trivial ring.

Classical rings :

$\mathbb{Z}$: the algebraic integers with addition and multiplication

the square nxn matrices with addition and multiplication

Ideals

They are important structures, which exist in more elaborate ways on other algebraic structures. So it is good to understand the concept in this simple form.

Definition 145 *A **right-ideal** of a ring E is a subset R of E such that :*

R is a subgroup of E for addition and $\forall a \in R, \forall x \in E : x \cdot a \in R$

*A **left-ideal** of a ring E is a subset L of E such that :*

L is a subgroup of E for addition and $\forall a \in L, \forall x \in E : a \cdot x \in L$

*A **two-sided ideal** (or simply an **ideal**) is a subset which is both a right-ideal and a left-ideal.*

Definition 146 *For any element a of the ring E :*

*the **principal right-ideal** is the right-ideal :$R = \{x \cdot a, x \in E\}$*

*the **principal left-ideal** is the left-ideal :$L = \{a \cdot x, x \in E\}$*

Division ring :

Definition 147 *A **division ring** is a ring for which any element other than 0 has an inverse for the second operation ·.*

The difference between a division ring and a field (below) is that · is not necessarily commutative.

Theorem 148 *Any finite division ring is also a field.*

Examples of division rings : the square invertible matrices, quaternions

Quaternions :

This is a division ring, usually denoted H, built over the real numbers, using 3 special "numbers" i,j,k (similar to the i of complex numbers) with the multiplication table :

$$\begin{bmatrix} 1\backslash 2 & i & j & k \\ i & -1 & k & -j \\ j & -k & -1 & i \\ k & j & -i & -1 \end{bmatrix}$$

$i^2 = j^2 = k^2 = -1, ij = k = -ji, jk = i = -kj, ki = j = -ik,$

Quaternions numbers are written as : $x = a + bi + cj + dk$ with $a, b, c, d \in \mathbb{R}$. Addition and multiplication are processed as usual for a,b,c,d and as the table above for i,j,k. So multiplication is *not* commutative.

$\mathbb{R}, \mathbb{C}$ can be considered as subsets of H (with b=c=d=0 or c=d=0 respectively).

The "real part" of a quaternion number is :$\operatorname{Re}(a + bi + cj + dk) = a$ so $\operatorname{Re}(xy) = \operatorname{Re}(yx)$

The "conjugate" of a quaternion is : $\overline{a + bi + cj + dk} = a - bi - cj - dk$ so $\operatorname{Re}(x\overline{y}) = \operatorname{Re}(y\overline{x})$, $x\overline{x} = a^2 + b^2 + c^2 + d^2 = \|x\|_{R^4}^2$

4.1.3 Field

Definition

Definition 149 *A **field** is a set with two operations (+ addition and x multiplication) which is an abelian group for +, the non zero elements are an abelian group for x, and multiplication is distributive over addition.*

A field is a commutative division ring.

Remark : an older usage did not require the multiplication to be commutative, and distinguished commutative fields and non commutative fields. It seems now that fields=commutative fields only. "Old" non commutative fields are now called division rings.

Classical fields :

$\mathbb{Q}$: rational numbers with addition and multiplication

$\mathbb{R}$: real numbers with addition and multiplication

$\mathbb{C}$: complex numbers with addition and multiplication

Algebraic numbers : real numbers which are the root of a one variable polynomial equation with integers coefficients

$x \in A \Leftrightarrow \exists n, (q_k)_{k=1}^{n-1}, q_k \in \mathbb{Q} : x^n + \sum_{k=0}^{n-1} q_k x^k = 0$

$\mathbb{Q} \subset A \subset \mathbb{R}$

For $a \in A, a \notin \mathbb{Q}$, define $A^*(a) = \left\{ x \in \mathbb{R} : \exists (q_k)_{k=1}^{n-1}, q_k \in Q : x = \sum_{k=0}^{n-1} q_k a^k \right\}$ then $A^*(a)$ is a field. It is also a n dimensional vector space over the field $\mathbb{Q}$

Characteristic :

Definition 150 *The **characteristic** of a field is the smallest integer n such that 1+1+...+1 (n times)=0. If there is no such number the field is said to be of characteristic 0 .*

All finite fields (with only a finite number of elements), also called "Gallois fields", have a finite characteristic which is a prime number.

Fields of characteristic 2 are the boolean algebra of computers.

Polynomials

1. Polynomials can be defined on a field (they can also be defined on a ring but we will not use them) :

Definition 151 *A **polynomial** of degree n with p variables on a field K is a function :*

$P : K^p \to K :: P(X_1, ..., X_p) = \sum a_{i_1...i_k} X_1^{i_1}...X_p^{i_p}, \sum_{j=1}^{p} i_j \leq n, a_{i_1...i_p} \in K$

*If $\sum_{j=1}^{p} i_j = n$ the polynomial is said to be **homogeneous**.*

Theorem 152 *The set of polynomials of degree n with p variables over a field K has the structure of a finite dimensional vector space over K denoted usually $K_n[X_1, ...X_p]$*

The set of polynomials of any degree with k variables has the structure of a commutative ring, with pointwise multiplication, denoted usually $K[X_1, ...X_p]$. So it is a (infinite dimensional) commutative algebra.

Definition 153 *A field is **algebraically closed** if any polynomial equation (with 1 variable) has at least one solution :*

$\forall a_0, ..a_n \in K, \exists x \in K : P(x) = a_n x^n + a_{n-1} x^{n-1} + ... + a_1 x + a_{.0} = 0$

$\mathbb{R}$ is not algebraically closed, but $\mathbb{C}$ is closed (this is the main motive to introduce $\mathbb{C}$).

Anticipating on the following, this generalization of a classic theorem.

Theorem 154 *Homogeneous functions theorem (Kolar p.213): Any smooth function $f : \prod_{i=1}^{n} E_i \to \mathbb{R}$ where $E_i, i = 1..n$ are finite dimensional real vector spaces, such that : $\exists a_i > 0, b \in \mathbb{R}, \forall k \in \mathbb{R} : f(k^{a_1} x_1, .., k^{a_n} x_n) = k^b f(x_1, .., x_n)$ is the sum of polynomials of degree d_i in x_i satisfying the relation : $b = \sum_{i=1}^{n} d_i a_i$. If there is no such non negative integer d_i then f=0.*

Complex numbers

This is the algebraic extension $\mathbb{C}$ of the real numbers $\mathbb{R}$. The fundamental theorem of algebra says that any polynomial equation has a solution over $\mathbb{C}$.

Complex numbers are written : $z = a + ib$ with $a, b \in \mathbb{R}, i^2 = -1$

The **real part** of a complex number is :$\mathrm{Re}(a + bi) = a$ and the **imaginary part** is $\mathrm{Im}(a + ib) = b$.

The **conjugate** of a complex number is : $\overline{a + bi} = a - bi$

So there are the useful identities :

$\mathrm{Re}(zz') = \mathrm{Re}(z)\mathrm{Re}(z') - \mathrm{Im}(z)\mathrm{Im}(z')$

$\mathrm{Im}(zz') = \mathrm{Re}(z)\mathrm{Im}(z') + \mathrm{Im}(z)\mathrm{Re}(z')$

$\mathrm{Re}(z) = \mathrm{Re}(\overline{z}) ; \mathrm{Im}(z) = -\mathrm{Im}(\overline{z})$

$\mathrm{Re}(z\overline{z'}) = \mathrm{Re}(z)\mathrm{Re}(z') + \mathrm{Im}(z)\mathrm{Im}(z')$

$\mathrm{Im}(z\overline{z'}) = \mathrm{Re}(z)\mathrm{Im}(z') - \mathrm{Im}(z)\mathrm{Re}(z')$

The **module** of a complex number is : $|a + ib| = \sqrt{a^2 + b^2}$ and $z\overline{z} = |z|^2$

The infinite sum : $\sum_{n=0}^{\infty} \frac{z^n}{n!} = \exp z$ always converges and defines the **exponential** function. The cos and sin functions can be defined as : $\exp z = |z|(\cos\theta + i\sin\theta)$ thus any complex number can be written as : $z = |z|(\cos\theta + i\sin\theta) = |z|e^{i\theta}, \theta \in [0, \pi]$ and $\exp(z_1 + z_2) = \exp z_1 \exp z_2$.

The set denoted $\mathrm{SU}(1) = \{z \in \mathbb{C} : |z| = 1\} = \{e^{i\theta}, \theta \in [0, 2\pi]\}$ is frequently used.

A formula which can be useful. Let be $z = a + ib$ then the complex numbers $\alpha + i\beta$ such that $(\alpha + i\beta)^2 = z$ are :

$$\alpha + i\beta = \pm\frac{1}{\sqrt{2}}\left(\sqrt{a + |z|} + i\frac{b}{\sqrt{a + |z|}}\right) = \pm\frac{1}{\sqrt{2}\sqrt{a+|z|}}(a + |z| + ib) = \pm\frac{z + |z|}{\sqrt{2}\sqrt{a+|z|}}$$

4.2   From vector spaces to algebras

4.2.1   Vector space

(or linear space)

Affine spaces are considered in the section vector spaces.

Definition 155 *A **vector space E over a field K** is a set with two operations : addition denoted + for which it is an abelian group, and multiplication by a scalar : $K \times E \to E$ which is distributive over addition.*

The elements of vector spaces are **vectors**. And the elements of the field K are **scalars**.

Remark : a module over a ring R is a set with the same operations as above. The properties are not the same. Definitions and names differ according to the authors.

4.2.2   Algebra

Algebra is a structure that is very common. It has 3 operations.

Definition

Definition 156 *An **algebra** $(A, \cdot)$ **over a field K** is a set A which is a vector space over K, endowed with an additional internal operation $\cdot : A \times A \rightarrow A$ with the following properties :*

    $\cdot$ is associative : $\forall X, Y, Z \in A : X \cdot (Y \cdot Z) = (X \cdot Y) \cdot Z$

    $\cdot$ is distributive over addition : $\forall X, Y, Z \in A : X \cdot (Y + Z) = X \cdot Y + X \cdot Z; (Y + Z) \cdot X = Y \cdot X + Z \cdot X$

    $\cdot$ is compatible with scalar multiplication : $\forall X, Y \in A, \forall \lambda, \mu \in K : (\lambda X) \cdot (\mu Y) = (\lambda \mu) X \cdot Y$

    *If there is an identity element I for $\cdot$ the algebra is said to be **unital**.*

    Remark : some authors do not require $\cdot$ to be associative

    An algebra A can be made unital by the extension : $A \rightarrow \widetilde{A} = K \oplus A = \{(k, X)\}, I = (1, 0), (k, X) = k1 + X$

Definition 157 *A **subalgebra** of an algebra $(A, \cdot)$ is a subset B of A which is also an algebra for the same operations*

    So it must be closed for the operations of the algebra.

Examples :

  quaternions

    square matrices over a field

    polynomials over a field

    linear endomorphisms over a vector space (with composition)

    Clifford algebra (see specific section)

Ideal

Definition 158 *A **right-ideal** of an algebra $(A, \cdot)$ is a vector subspace R of A such that : $\forall a \in R, \forall x \in E : x \cdot a \in R$*

    *A **left-ideal** of an algebra $(A, \cdot)$ is a vector subspace L of A: $\forall a \in L, \forall x \in E : a \cdot x \in L$*

    *A **two-sided ideal** (or simply an **ideal**) is a subset which is both a right-ideal and a left-ideal.*

Definition 159 *An algebra $(A, \cdot)$ is **simple** if the only two-sided ideals are 0 and A*

Derivation

Definition 160 *A **derivation** over an algebra $(A, \cdot)$ is a linear map : $D : A \rightarrow A$ such that*

    $\forall u, v \in A : D(u \cdot v) = (Du) \cdot v + u \cdot (Dv)$

(we have a relation similar to the Leibniz rule for the derivative of the product of two scalar functions)

Commutant

Definition 161 *The **commutant,** denoted S', of a subset S of an algebra $(A, \cdot)$, is the set of all elements in A which commute with all the elements of S for the operation $\cdot$.*

Theorem 162 *(Thill p.63-64) A commutant is a subalgebra, containing I if A is unital.*
$$S \subset T \Rightarrow T' \subset S'$$

For any subset S, the elements of S commute with each others iff $S \subset S'$
S' is the centralizer (see Groups below) of S for the internal operation $\cdot$.

Definition 163 *The **second commutant** of a subset of an algebra $(A, \cdot)$, is the commutant denoted S" of the commutant S' of S*

Theorem 164 *(Thill p.64)*
$$S \subset S"$$
$$S' = (S")'$$
$$S \subset T \Rightarrow (S')' \subset (T')'$$
$$X, X^{-1} \in A \Rightarrow X^{-1} \in (X)"$$

Projection and reflexion

Definition 165 *A **projection** in an algebra $(A, \cdot)$ is a an element X of A such that : $X \cdot X = X$*

Definition 166 *Two projections X,Y of an algebra $(A, \cdot)$ are said to be **orthogonal** if $X \cdot Y = 0$ (then $Y \cdot X = 0$)*

Definition 167 *Two projections X,Y of a unital algebra $(A, \cdot)$ are said to be **complementary** if X+Y=I*

Definition 168 *A **reflexion** of a unital algebra $(A, \cdot)$ is an element X of A such that $X = X^{-1}$*

Theorem 169 *If X is a reflexion of a unital algebra $(A, \cdot)$ then there are two complementary projections such that X=P-Q*

Definition 170 *An element X of an algebra $(A, \cdot)$ is **nilpotent** if $X \cdot X = 0$*

*-algebra

*-algebras (say star algebra) are endowed with an additional operation similar to conjugation-transpose of matrix algebras.

Definition 171 *A ***-algebra** is an algebra $(A, \cdot)$ over a field K, endowed with an involution :* $* : A \to A$ *such that :*
$\forall X, Y \in A, \lambda \in K :$
$(X + Y)^* = X^* + Y^*$
$(X \cdot Y)^* = Y^* \cdot X^*$
$(\lambda X)^* = \overline{\lambda} X^*$ *(if the field K is $\mathbb{C}$)*
$(X^*)^* = X$

Definition 172 *The **adjoint** of an element X of a *-algebra is X^**

Definition 173 *A subset S of a *-algebra is **stable** if it contains all its adjoints :* $X \in S \Rightarrow X^* \in S$

The commutant S' of a stable subset S is stable

Definition 174 *A ***-subalgebra** B of A is a stable subalgebra :* $B^* \sqsubseteq B$

Definition 175 *An element X of a *-algebra $(A, \cdot)$ is said to be :*
normal *if X·X*=X*·X,*
self-adjoint *(or hermitian) if X=X**
anti self-adjoint *(or antihermitian) if X=-X**
unitary *if X·X*=X*·X=I*

(All this terms are consistent with those used for matrices where * is the transpose-conjugation).

Theorem 176 *A *-algebra is commutative iff each element is normal*

If the *algebra A is over $\mathbb{C}$ then :
i) Any element X in A can be written : $X = Y + iZ$ with Y,Z self-adjoint : $Y = \frac{1}{2}(X + X^*), Z = \frac{1}{2i}(X - X^*)$
ii) The subset of self-adjoint elements in A is a real vector space, real form of the vector space A.

4.2.3   Lie Algebra

There is a section dedicated to Lie algebras in the part Lie Groups.

Definition 177 *A **Lie algebra** over a field K is a vector space A over K endowed with a bilinear map called **bracket** :*$[] : A \times A \to A$
$\forall X, Y, Z \in A, \forall \lambda, \mu \in K : [\lambda X + \mu Y, Z] = \lambda [X, Z] + \mu [Y, Z]$
such that :
$[X, Y] = -[Y, X]$
$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ *(Jacobi identities)*

Notice that a Lie algebra is not an algebra, because the bracket is not associative. But any algebra $(A, \cdot)$ becomes a Lie algebra with the bracket : $[X, Y] = X \cdot Y - Y \cdot X$. This is the case for the linear endomorphisms over a vector space.

4.2.4   Algebraic structures and categories

If the sets E and F are endowed with the same algebraic structure a map $f : E \rightarrow F$ is a **morphism** (also called homomorphism) if f preserves the structure = the image of the result of any operation between elements of E is the result of the same operation in F between the images of the elements of E.

Groups : $\forall x, y \in E : f(x * y) = f(x) \cdot f(y)$

Ring : $\forall x, y, z \in E : f((x + y) * z) = f(x) \cdot f(z) + f(y) \cdot f(z)$

Vector space : $\forall x, y \in E, \lambda, \mu \in K : f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$

Algebra : $\forall x, y \in A, \lambda, \mu \in K :: f(x * y) = f(x) \cdot f(y) ; f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$

Lie algebra : $\forall X, Y \in E : f([X, Y]_E) = [f(X), f(Y)]_F$

If f is bijective then f is an isomorphism

If E=F then f is an endomorphism

If f is an endomorphism and an isomorphism it is an automorphism

All these concepts are consistent with the morphisms defined in the category theory.

There are many definitions of "homomorphisms", implemented for various mathematical objects. As far as only algebraic properties are involved we will stick to the universal and clear concept of morphism.

There are the categories of Groups, Rings, Fields, Vector Spaces, Algebras over a field K.

5 GROUPS

We see here mostly general definitions about groups, and an overview of the finite groups. Topological groups and Lie groups are studied in a dedicated part.

5.1 Definitions

Definition 178 *A **group** $(G, \cdot)$ is a set endowed G with an associative operation $\cdot$, for which there is an identity element and every element has an inverse.*

In a group, the identity element is unique. The inverse of an element is unique.

Definition 179 *A commutative (or **abelian**) group is a group whith a commutative operation*

Definition 180 *A **subgroup** of the group $(G, \cdot)$ is a subset A of G which is also a group for $\cdot$*

So : $1_G \in A, \forall x, y \in A : x \cdot y \in A, x^{-1} \in A$

5.1.1 Involution

Definition 181 *An **involution** on a group $(G, \cdot)$ is a map : $* : G \to G$ such that :*
$\forall g, h \in G : (g^*)^* = g; (g \cdot h)^* = h^* \cdot g^*; (1)^* = 1$

$\Rightarrow (g^{-1})^* = (g^*)^{-1}$
A group endowed with an involution is said to be an involutive group.
Any group has the involution : $(g)^* = g^{-1}$ but there are others
Example : $(\mathbb{C}, \times)$ with $(z)^* = \overline{z}$

5.1.2 Morphisms

Definition 182 *If $(G, \cdot)$ and $(G\prime, *)$ are groups a **morphism** (or homomorphism) is a map : $f : G \to G'$ such that :*
$\forall x, y \in G : f(x \cdot y) = f(x) * f(y) \; ; f(1_G) = 1_{G'}$

$\Rightarrow f(x^{-1}) = f(x)^{-1}$
The set of such morphisms f is denoted $\hom(G, G')$

The **category of groups** has objects = groups and morphisms = homomorphisms.

Definition 183 *The **kernel** of a morphism $f \in \hom(G, G')$ is the set :* $\ker f = \{g \in G : f(g) = 1_{G'}\}$

5.1.3 Translations

Definition 184 *The **left-translation** by $a \in (G, \cdot)$ is the map : $L_a : G \to G ::$* $L_a x = a \cdot x$

*The **right-translation** by $a \in (G, \cdot)$ is the map : $R_a : G \to G :: R_a x = x \cdot a$*

So : $L_x y = x \cdot y = R_y x$. Translations are bijective maps.

Definition 185 *The **conjugation** with respect to $a \in (G, \cdot)$ is the map : $Conj_a :$* $G \to G :: Conj_a x = a \cdot x \cdot a^{-1}$

$Conj_a x = L_a \circ R_{a^{-1}}(x) = R_{a^{-1}} \circ L_a(x)$

Definition 186 *The **commutator** of two elements $x,y \in (G, \cdot)$ is : $[x, y] = x^{-1} \cdot y^{-1} \cdot x \cdot y$*

It is 0 (or 1) for abelian groups.

It is sometimes useful (to compute the derivatives for instance) to consider the operation $\cdot$ as a map with two variables :

$x \cdot y = M(x, y)$ with the property $M(M(x, y), z) = M(x, M(y, z))$

5.1.4 Centralizer

Definition 187 *The **normalizer** of a subset A of a group $(G, \cdot)$ is the set :* $N_A = \{x \in G : Conj_x(A) = A\}$

*The **centralizer** of a subset A of a group $(G, \cdot)$ is the set Z_A of elements of G which commute with the elements of A*

*The **center** Z_G of G is the centralizer of G*

$Z_A = \{x \in G : \forall a \in A : ax = xa\}$

Z_A is a subgroup of G.

5.1.5 Quotient sets

Cosets are similar to ideals.

Definition 188 *For a subgroup H of a group $(G, \cdot)$ and $a \in G$*

*The **right coset** of a (with respect to H) is the set : $H \cdot a = \{h \cdot a, h \in H\}$*

*The **left coset** of a (with respect to H) is the set : $a \cdot H = \{a \cdot h, h \in H\}$*

The left and right cosets of H may or may not be equal.

Definition 189 *A subgroup of a group $(G, \cdot)$ is a **normal subgroup** if its right-coset is equal to its left coset*

Then for all g in G, gH = Hg, and $\forall x \in G : x \cdot H \cdot x^{-1} \in H$.

If G is abelian any subgroup is normal.

Theorem 190 *The kernel of a morphism $f \in \hom(G, G')$ is a normal subgroup. Conversely any normal subgroup is the kernel of some morphism.*

Definition 191 *A group $(G, \cdot)$ is **simple** if the only normal subgroups are 1 and G itself.*

Theorem 192 *The left-cosets (resp.right-cosets) of any subgroup H form a partition of G*

that is, the union of all left cosets is equal to G and two left cosets are either equal or have an empty intersection.

So a subgroup defines an equivalence relation

Definition 193 *The **quotient set** G/H of a subgroup H of a group $(G, \cdot)$ is the set $G/\sim$ of classes of equivalence : $x \sim y \Leftrightarrow \exists h \in H : x = y \cdot h$*
*The **quotient set** $H\backslash G$ of a subgroup H of a group $(G, \cdot)$ is the set $G/\sim$ of classes of equivalence : $x \sim y \Leftrightarrow \exists h \in H : x = h \cdot y$*

It is useful to characterize these quotient sets.
The projections give the classes of equivalences denoted $[x]$:
$\pi_L : G \to G/H : \pi_L(x) = [x]_L = \{y \in G : \exists h \in H : x = y \cdot h\} = x \cdot H$
$\pi_R : G \to H\backslash G : \pi_R(x) = [x]_R = \{y \in G : \exists h \in H : x = h \cdot y\} = H \cdot x$
Then :
$x \in H \Rightarrow \pi_L(x) = \pi_R(x) = [x] = 1$
Because the classes of equivalence define a partition of G, by the Zorn lemma one can pick one element in each class. So we have two families :
For G/H : $(\lambda_i)_{i \in I}$: $\lambda_i \in G$: $[\lambda_i]_L = \lambda_i \cdot H, \forall i, j$: $[\lambda_i]_L \cap [\lambda_j]_L = \varnothing, \cup_{i \in I} [\lambda_i]_L = G$
For $H\backslash G$: $(\rho_j)_{j \in J}$: $\rho_j \in G$: $[\rho_j]_R = H \cdot \rho_j \forall i, j$: $[\rho_i]_R \cap [\rho_j]_R = \varnothing, \cup_{j \in J} [\rho_j]_R = G$
Define the maps :
$\phi_L : G \to (\lambda_i)_{i \in I} : \phi_L(x) = \lambda_i :: \pi_L(x) = [\lambda_i]_L$
$\phi_R : G \to (\rho_j)_{j \in J} : \phi_R(x) = \rho_j :: \pi_R(x) = [\rho_j]_R$
Then any $x \in G$ can be written as :$x = \phi_L(x) \cdot h$ or $x = h' \cdot \phi_R(x)$ for unique h,h'$\in H$

Theorem 194 *$G/H=H\backslash G$ iff H is a normal subgroup. If so then $G/H=H\backslash G$ is a group and the sequence $1 \to H \to G \to G/H \to 1$ is exact (in the category of groups, with 1=trivial group with only one element). The projection $G \to G/H$ is a morphism with kernel H.*

There is a similar relation of equivalence with conjugation:

Theorem 195 *The relation : $x \sim y \Leftrightarrow x = y \cdot x \cdot y^{-1} \Leftrightarrow x \cdot y = y \cdot x$ is an equivalence relation over $(G, \cdot)$ which defines a partition of $G : G = \cup_{p \in P} G_p, p \neq q : G_p \cap G_q = \varnothing$. Each subset G_p of G is a **conjugation class**. If G is commutative there is only one subset, G itself*

(as any element commutes with ist powers x^n the conjugation class of x contains at least its powers,including the unity element).

5.1.6 Semi-direct product of groups

Any subgroup H defines a partition and from there any element of the group can be written uniquely as a product of an element of H and an element of a family $(\lambda_i)_{i \in I}$ or $(\rho_j)_{j \in J}$. It these families are themselves a subgroup then G can be written as the product of two subgroups. More precisely :

Theorem 196 *Let $(G, \cdot)$ be a group, N a normal subgroup and H a subgroup of G. The following statements are equivalent:*

i) $G = N \cdot H$ and $N \cap H = \{1\}$.

ii) $G = H \cdot N$ and $N \cap H = \{1\}$.

iii) Every element of G can be written as a unique product of an element of N and an element of H.

iv) Every element of G can be written as a unique product of an element of H and an element of N.

v) There exists a morphism $G \to H$ which is the identity on H and whose kernel is N.

*If one of these statement holds, then G is said to be **semidirect product** of N and H.*

One says also that G splits over N.

If a group is simple its only normal subgroups are trivial, thus it cannot be decomposed in the semi-product of two other groups. Simple groups are the basic bricks from which other groups can be built.

5.1.7 Generators

Definition 197 *A set of **generators** of a group $(G, \cdot)$ is a $(x_i)_{i \in I}$ a family of elements of G indexed on an ordered set I such that any element of G can be written uniquely as the product of a finite ordered subfamily J of $(x_i)_{i \in I}$*

$\forall g \in G, \exists J = \{j_1, ... j_n, ..\} \subset I, : g = x_{j_1} \cdot x_{j_2} ... \cdot x_{j_n} ..$

The **rank** of a group is the cardinality of the smallest set of its generators (if any). A group is free if it has a finite family of generators.

5.1.8 Action of a group

Maps involving a group and a set can have special properties, which deserve definitions because they are frequently used.

Definition 198 *A **left-action** of a group $(G, \cdot)$ on a set E is a map : $\lambda : G \times E \to E$ such that :*

$\forall x \in E, \forall g, g' \in G : \lambda(g, \lambda(g', x)) = \lambda(g \cdot g', x) ; \lambda(1, x) = x$

*A **right-action** of a group $(G, \cdot)$ on a set E is a map : $\rho : E \times G \to E$ such that :*

$\forall x \in E, \forall g, g' \in G : \rho(\rho(x, g'), g)) = \rho(x, g' \cdot g) ; \rho(x, 1) = x$

Notice that left, right is related to the place of g.

Any subgroup H of G defines left and right actions by restriction of the map to H.

Any subgroup H of G defines left and right actions on G itself in the obvious way.

As a consequence of the definition :

$$\lambda\left(g^{-1},x\right) = \lambda\left(g,x\right)^{-1} ; \rho\left(x,g^{-1}\right) = \rho\left(x,g\right)^{-1}$$

All the following definitions are easily adjusted for a right action.

Definition 199 *The **orbit** of the action through a∈ G of the left-action λ of a group $(G,\cdot)$ on a set E is the subset of E denoted $G(a) = \{\lambda\left(g,a\right), g \in G\}$*

The relation $y \in G\left(x\right)$ is an equivalence relation between x,y. The classes of equivalence form a partition of G called the **orbits** of the action (an orbit = the subset of elements of E which can be deduced from each other by the action).

The orbits of the left action of a subgroup H on G are the right cosets defined above.

Definition 200 *A left-action of a group $(G,\cdot)$ on a set E is*
    transitive *if ; $\forall x, y \in E, \exists g \in G : y = \lambda\left(g,x\right)$. If so E is called an **homogeneous space**.*
    free *if : $\lambda\left(g,x\right) = x \Rightarrow g = 1$*
    effective *if : $\forall x : \lambda(g,x) = \lambda(h,x) => g = h$*

Definition 201 *A subset F of E is **invariant** by the left-action λ of a group $(G,\cdot)$ on E if : $\forall x \in F, \forall g \in G : \lambda\left(g,x\right) \in F$.*

F is invariant iff it is the union of a collection of orbits. The minimal non empty invariant sets are the orbits.

Definition 202 *The **stabilizer** of an element a∈ E with respect to the left-action λ of a group $(G,\cdot)$ on E is the subset of $G : A(a) = \{g \in G : \lambda\left(g,a\right) = a\}$*

It is a subgroup of G also called the **isotropy subgroup** (with respect to a). If the action is free the map : $A : E \rightarrow G$ is bijective.

Definition 203 *Two set E,F are **equivariant** under the left actions $\lambda_1 : G \times E \rightarrow E, \lambda_2 : G \times F \rightarrow F$ of a group $(G,\cdot)$ if there is a map : $f : E \rightarrow F$ such that : $\forall x \in E, \forall g \in G : f\left(\lambda_1\left(g,x\right)\right) = \lambda_2\left(g, f\left(x\right)\right)$*

Then f is a natural tranformation for the functors λ_1, λ_2

So if E=F the set is equivariant under the action if : $\forall x \in E, \forall g \in G : f\left(\lambda\left(g,x\right)\right) = \lambda\left(g, f\left(x\right)\right)$

5.2 Finite groups

A finite group is a group which has a finite number of elements. So, for a finite group, one can dress the multiplication table, and one can guess that there a not too many ways to build such a table : mathematicians have strive for years to establish a classification of finite groups.

5.2.1 Classification of finite groups

1. Order:

Definition 204 *The **order** of a finite group is the number of its elements. The order of an element **a** of a finite group is the smallest positive integer number k with $a^k = 1$, where 1 is the identity element of the group.*

Theorem 205 *(Lagrange's theorem) The order of a subgroup of a finite group G divides the order of G.*
*The order of an element **a** of a finite group divides the order of that group.*

Theorem 206 *If n is the square of a prime, then there are exactly two possible (up to isomorphism) types of group of order n, both of which are abelian.*

2. Cyclic groups :

Definition 207 *A group is **cyclic** if it is generated by an element :* $G = \{a^p, p \in \mathbb{N}\}$.

A cyclic group always has at most countably many elements and is commutative. For every positive integer n there is exactly one cyclic group (up to isomorphism) whose order is n, and there is exactly one infinite cyclic group (the integers under addition). Hence, the cyclic groups are the simplest groups and they are completely classified. They are usually denoted $\mathbb{Z}/p\mathbb{Z}$: the algebric number multiple of p with addition.

3. All *simple* finite groups have been classified (the proof covers thousands of pages). Up to isomorphisms there are 4 classes :
   - the cyclic groups with prime order : any group of prime order is cyclic and simple.
   - the alternating groups of degree at least 5;
   - the simple Lie groups
   - the 26 sporadic simple groups.

5.2.2 Symmetric groups

Symmetric groups are the key to the study of permutations.

Definitions
  1. Permutation:

Definition 208 *A **permutation** of a finite set E is a bijective map : $p : E \rightarrow E$.*

  With the composition law the set of permutations of E is a group. As all sets with the same cardinality are in bijection, their group of permutations are isomorphics. Therefore it is convenient, for the purpose of the study of permutations, to consider the set (1,2,....n) of integers.

Notation 209 $\mathfrak{S}(n)$ *is the group of permutation of a set of n elements, called the **symmetric group** of order n*

  An element s of $\mathfrak{S}(n)$ can be represented as a table with 2 rows : the first row is the integers 1,2..n, the second row takes the elements s(1),s(2),...s(n).
  $\mathfrak{S}(n)$ is a finite group with n! elements. Its subgroups are permutations groups. It is abelian iff n<2.
  Remark : one always consider two elements of E as distinct, even if it happens that, for other reasons, they are indicible. For instance take the set $\{1, 1, 2, 3\}$ with cardinality 4. The two first elements are considered as distinct : indeed in abstract set theory nothing can tell us that two elements are not distinct, so we have 4 objects $\{a, b, c, d\}$ that are numbered as $\{1, 2, 3, 4\}$
  2. Transposition:

Definition 210 *A **transposition** is a permutation which exchanges two elements and keep inchanged all the others.*

  A transposition can be written as a couple (a,b) of the two numbers which are transposed.
  Any permutation can be written as the composition of transpositions. However this decomposition is not unique, but the parity of the number p of transpositions necessary to write a given permutation does not depend of the decomposition. The **signature** of a permutation is the number $(-1)^p = \pm 1$. A permutation is even if its signature is +1, odd if its signature is -1. The product of two even permutations is even, the product of two odd permutations is even, and all other products are odd.
  The set of all even permutations is called the **alternating group** A_n (also denoted $\mathfrak{A}_n$). It is a normal subgroup of $\mathfrak{S}(n)$, and for $n \geq 2$ it has $n!/2$ elements. The group $\mathfrak{S}(n)$ is the semidirect product of A_n and any subgroup generated by a single transposition.

Young diagrams
  For any partition of (1,2,...n) in p subsets, the permutations of $\mathfrak{S}(n)$ which preserve globally each of the subset of the partition constitute a class of conjugation.

Example : the 3 permutations $(1,2,3,4,5),(2,1,4,3,5),(1,2,5,3,4)$, preserve the subsets $(1,2),(3,4,5)$ and belong to the same class of conjugation.

A class of conjugation is defined by p integers $\lambda_1 \leq \lambda_2... \leq \lambda_p$ such that $\sum_{i=1}^p \lambda_i = n$ and a partition of (1,2,...n) in p subsets $(i_1,..i_{\lambda_k})$ containing each λ_k elements taken in (1,2,..n). The number S(n,p) of different partitions of n in p subsets is a function of n, which is tabulated (this is the Stirling number of second kind).

Given such a partition denoted λ, as above, a **Young diagram** is a table with p rows i=1,2,...p of λ_k cells each, placed below each other, left centered. Any permutation of $\mathfrak{S}(n)$ obtained by filling such a table with distinct numbers 1,2,...n is called a **Young tableau**. The standard (or canonical) tableau is obtained in the natural manner by filling the cells from the left to the right in each row, and next to the row below with the ordered numbers 1,2,...n.

Given a Young tableau, two permutations belong to the same class of conjugation if they have the same elements in each row (but not necessarily in the same cells).

A Young diagram has also q columns, of decreasing sizes $\mu_j, j = 1...q$. with : $\sum_{i=1}^p \lambda_i = \sum_{j=1}^q \mu_j = n; n \geq \mu_j \geq \mu_{j+1} \geq 1$

If a diagram is read columns by columns one gets another diagram, called the **conjugate** of λ.

5.2.3 Symmetric polynomials

Definition 211 *A map of n variables over a set $E : f : E^n \to F$ is **symmetric** in its variables if it is invariant for any permutation of the n variables : $\forall \sigma \in \mathfrak{S}(n), f\left(x_{\sigma(1)},...,x_{\sigma(n)}\right) = f(x_1,...,x_n)$*

The set $\mathrm{S}_d[X_1,...X_n]$ of symmetric polynomials of n variables and degree d has the structure of a finite dimensional vector space. These polynomials must be homogeneous :

$P(x_1,..x_n) = \sum a_{i_1...i_p} x_1^{i_1}...x_n^{i_n}, \sum_{j=1}^n i_j = d, a_{i_1...i_p} \in F, X_i \in F$

The set $\mathrm{S}[X_1,...X_n]$ of symmetric polynomials of n variables and any degree has the structure of a graded commutative algebra with the multiplication of functions.

Basis of the space of symmetric polynomials

A basis of the vector space $\mathrm{S}[X_1,...X_n]$ is a set of symmetric polynomials of n variables. Their elements can be labelled by a partition λ of d : $\lambda = (\lambda_1 \geq \lambda_2... \geq \lambda_n \geq 0), \sum_{j=1}^n \lambda_j = d$. The most usual bases are the following.

1. Monomials : the basic monomial is $x_1^{\lambda_1} \cdot x_2^{\lambda_2}... \cdot x_n^{\lambda_n}$. The symmetric polynomial of degree d associated to the partition λ is $H_\lambda = \sum_{\sigma \in \mathfrak{S}(n)} x_{\sigma(1)}^{\lambda_1} \cdot x_{\sigma(2)}^{\lambda_2}... \cdot x_{\sigma(n)}^{\lambda_n}$ and a basis of $\mathrm{S}_d[X_1,...X_n]$ is a set of H_λ for each partition λ.

2. Elementary symmetric polynomials : the p elementary symmetric polynomial is : $E_p = \sum_{\{i_1,...i_p\}} x_{i_1} \cdot x_{i_2}.. \cdot x_{i_p}$ where the sum is for all ordered combinations of p indices taken in (1,2,...n): $1 \leq i_1 < i_2.. < i_n \leq n$. It is a symmetric polynomial of degree p. The product of two such polynomials $E_p \cdot E_q$

is still a symmetric polynomial of degree p+q. So any partition λ defines a polynomial : $H_\lambda = \prod_\lambda E_{\lambda_1}...E_{\lambda_q} \in S_d[x_1,...x_n]$ and a basis is a set of H_λ for all partitions λ. There is the identity : $\prod_n^{i=1}(1 + x_i t) = \sum_{j=0}^\infty E_j t^j$

3. Schur polynomials :the Schur polynomial for a partition λ is defined by : $S_\lambda = \det\left[x_j^{\lambda_i+n-i}\right]_{n\times n}/\Delta$ where : $\Delta = \prod_{i<j}(x_i - x_j)$.is the discriminant of a set of n variables.

There is the identiity : $\det\left[\frac{1}{1-x_iy_j}\right] = \left(\prod_{i<j}(x_i - x_j)\right)\left(\prod_{i<j}(y_i - y_j)\right)/\prod_{i,j}(1 - x_iy_j)$

5.2.4   Combinatorics

Combinatorics is the study of finite structures, and involves counting the number of such structures. We will just recall basic results in enumerative combinatorics and signatures.

Enumerative combinatorics
Enumerative combinatorics deals with problems such as "how many ways to select n objects among x ? or many ways to group in n packets x objects ?..."
1. Many enumerative problems can be modelled as following :
Find the number of maps :$f : N \to X$ where N is a set with n elements, X a set with x elements and meeting one of the conditions : f injective, f surjective, or no condition. Moreover any two maps f,f' :
i) are always distinct (no condition)
or are deemed equivalent (counted only once) if
ii) Up to a permutation of X : $f \sim f' : \exists s_X \in \mathfrak{S}(x) : f'(N) = s_X f(N)$
iii) Up to a permutation of N : $f \sim f' : \exists s_N \in \mathfrak{S}(n) : f'(N) = f(s_N N)$
iv) Up to permutations of N and X : $f \sim f' : \exists s \in \mathfrak{S}(x), s_N \in \mathfrak{S}(n) : f'(N) = s_X f(s_N N)$
These conditions can be paired in 12 ways.
2. Injective maps from N to X:
i) No condition : this is the number of sequences of n distinct elements of X without repetitions. The formula is : $\frac{x!}{(n-x)!}$
ii) Up to a permutation of X : 1 si $n \le x$,0 if n>x
iii) Up to a permutation of N : this is the number of subsets of n elements of X, the **binomial coefficient** : $C_x^n = \frac{x!}{n!(x-n)!} = \binom{x}{n}$. If n>x the result is 0.
iv) Up to permutations of N and X : 1 si $n \le x$ 0 if n>x
3. Surjective maps f from N to X:
i) No condition : the result is $x!S(n,x)$ where S(n,x), called the Stirling number of the second kind, is the number of ways to partition a set of n elements in k subsets (no simple formula).
ii) Up to a permutation of X : the result is the Stirling number of the second kind S(n,x).

iii) Up to a permutation of N: the result is : C_{x-1}^{n-1}

iv) Up to permutations of N and X : this is the the number $p_x(n)$ of partitions of n in x non zero integers : $\lambda_1 \geq \lambda_2... \geq \lambda_x > 0 : \lambda_1 + \lambda_2 + ..\lambda_x = n$

4. No restriction on f :

i) No condition : the result is x^n

ii) Up to a permutation of X : the result is $\sum_{k=0}^{x} S(n,k)$ where S(n,k) is the Stirling number of second kind

iii) Up to a permutation of N : the result is : $C_{x+n-1}^{n} = \binom{x+n-1}{x}$

iv) Up to permutations of N and X : the result is : $p_x(n+x)$ where $p_k(n)$ is the number of partitions of n in k integers : $\lambda_1 \geq \lambda_2... \geq \lambda_k : \lambda_1 + \lambda_2 + ..\lambda_k = n$

5. The number of distributions of n (distinguishable) elements over r (distinguishable) containers, each containing exactly k_i elements, is given by the **multinomial coefficients** :

$\binom{n}{k_1 k_2..k_r} = \frac{n!}{k_1!k_2!..k_r!}$

They are the coefficients of the polynomial : $(x_1 + x_2 + ... + x_r)^n$

6. Stirling's approximation of n! : $n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$

The gamma function : $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt : n! = \Gamma(n+1)$

Signatures

1. To compute the signature of any permutation, the basic rule is that the parity of any permutation of integers $(a_1, a_2, ..., a_p)$ (consecutive or not) is equal to the number of inversions in the permutation = the number of times that a given number a_i comes before another number a_{i+r} which is smaller than $a_i : a_{i+r} < a_i$

Example : $(3, 5, 1, 8)$

take 3 : $> 1 \to +1$

take 5 : $> 1 \to +1$

take 1 : $\to 0$

take 8 : $\to 0$

take the sum : 1+1=2 $\to$ signature $(-1)^2 = 1$

2. It is most useful to define the function :

Notation 212 ϵ is the function at n variables : $\epsilon : I^n \to \{-1,0,1\}$ where I is a set of n integers, defined by :

$\epsilon(i_1,...,i_n) = 0$ if there are two indices which are identical : i_k, i_l ,$k \neq l$ such that : $i_k = i_l$

$\epsilon(i_1,...,i_n) =$ the signature of the permutation of the integers $(i_1,...,i_n)$ it they are all distinct

So $\epsilon(3,5,1,8) = 1; \epsilon(3,5,5,8) = 0$

Notation 213 $\epsilon(\sigma)$ where $\sigma \in \mathfrak{S}(n)$ is the signature of the permutation σ

3. Basic formulas :

reverse order : $\epsilon(a_p, a_{p-1}, ..., a_1) = \epsilon(a_1, a_2, ..., a_p)(-1)^{\frac{p(p-1)}{2}}$

inversion of two numbers : $\epsilon(a_1, a_2, .a_j...a_i.., a_p) = \epsilon(a_1, a_2, .a_i...a_j.., a_p)\epsilon(a_i, a_j)$

inversion of one number : $\epsilon(i, 1, 2, 3, ..i-1, i+1, ...p) = (-1)^{i-1}$

6 VECTOR SPACES

Vector spaces should are well known structures. However it is necessary to have clear and precise definitions of the many objects which are involved. Furthermore in this section we do our best to give definitions and theorems which are valid whatever the field K, and for infinite dimensional vector spaces (as they are in the many applications).

6.1 Vector spaces

6.1.1 Vector space

Definition 214 *A **vector space E over a field K** is a set with two operations : addition denoted + for which it is an abelian group, and multiplication by a scalar (an element of K) : $K \times E \to E$ which is distributive over addition.*

So : $\forall x, y \in E, \lambda, \mu \in K$:
$\lambda x + \mu y \in E$,
$\lambda (x + y) = (x + y) \lambda = \lambda x + \lambda y$

Elements of a vector space are called **vectors**. When necessary (and only when necessary) vectors will be denoted with an upper arrow : $\overrightarrow{u}$

Warning ! a vector space structure is defined with respect to a given field (see below for real and complex vector spaces)

6.1.2 Basis

Definition 215 *A family of vectors $(v_i)_{i \in I}$ of a vector space over a field K, indexed on a finite set I, is **linearly independant** if :*
$\forall (x_i)_{i \in I}, x_i \in K : \sum_{i \in I} x_i v_i = 0 \Rightarrow x_i = 0$

Definition 216 *A family of vectors $(v_i)_{i \in I}$ of a vector space, indexed on a set I (finite of infinite) is **free** if any finite subfamily is linearly independant.*

Definition 217 *A **basis** of a vector space E is a free family of vectors which generates E.*

Thus for a basis $(e_i)_{i \in I} : \forall v \in E, \exists J \subset I, \#J < \infty, \exists (x_i)_{i \in J} \in K^J : v = \sum_{i \in J} x_i e_i$

Warning! These complications are needed because without topology there is no clear definition of the infinite sum of vectors. This implies that for any vector *at most a finite number of components are non zero* (but there can be an infinite number of vectors in the basis). So usually "Hilbertian bases" are not bases in this general meaning, because vectors can have infinitely many non zero components.

The method to define a basis is a common trick in algebra. To define some property on a family indexed on an infinite set I, without any tool to compute

operations on an infinite number of arguments, one says that the property is valid on I if it is valid on all the finite subsets J of I. In analysis there is another way, by using the limit of a sequence and thus the sum of an infinite number of arguments.

Theorem 218 *Any vector space has a basis*

(this theorem requires the axiom of choice).

Theorem 219 *The set of indices of bases of a vector space have all the same cardinality, which is the **dimension** of the vector space.*

If K is a field, the set K^n is a vector space of dimension n, and its canonical basis are the vectors $\varepsilon_i = (0, 0, ..0, 1, 0, ...0)$.

6.1.3   Vector subspaces

Definition 220 *A **vector subspace** of a vector space E over a field K is a subset F of E such that the operations in E are algebraically closed in F :*

$\forall u, v \in F, \forall k, k' \in K : ku + k'u' \in F$
the operations (+,x) being the operations as defined in E.

Linear span

Definition 221 *The **linear span** of the subset S of a vector space E is the intersection of all the vector subspaces of E which contains S.*

Notation 222 *Span(S) is the linear span of the subset S of a vector space*

Span(S) is a vector subspace of E, which contains any *finite* linear combination of vectors of S.

Direct sum
This concept is important, and it is essential to understand fully its signficance.

Definition 223 *The sum of a family $(E_i)_{i \in I}$ of vector subspaces of E is the linear span of $(E_i)_{i \in I}$*

So any vector of the sum is the sum of at most a finite number of vectors of some of the E_i

Definition 224 *The sum of a family $(E_i)_{i \in I}$ of vector subspaces of E is **direct** and denoted $\oplus_{i \in I} E_i$ if for any finite subfamily J of I :*
$\sum_{i \in J} v_i = \sum_{i \in J} w_i, v_i, w_i \in E_i \Rightarrow v_i = w_i$

The sum is direct iff the E_i have no common vector but 0 : $\forall j \in I, E_j \cap \left(\sum_{i \in I-j} E_i \right) = \overrightarrow{0}$

Or equivalently the sum is direct iff the decomposition over each E_i is unique : $\forall v \in E, \exists J \subset I, \#J < \infty, \exists v_j$ unique $\in E_j : v = \sum_{j \in J} v_j$

If the sum is direct the projections are the maps : $\pi_i : \oplus_{j \in I} E_j \to E_i$

Warning!

i) If $\oplus_{i \in I} E_i = E$ the sum is direct iff the decomposition of any vector of E with respect to the E_i is unique, but this does not entail that there is a unique collection of subspaces E_i for which we have such a decomposition. Indeed take any basis : the decomposition with respect to each vector subspace generated by the vectors of the basis is unique, but with another basis we have another unique decomposition.

ii) If F is a vector subspace of E there is always a unique subset G of E such that G=F^c but G is not a vector subspace (because 0 must be both in F and G for them to be vector spaces). Meanwhile there are always vector subspaces G such that : $E = F \oplus G$ but G is not unique. A way to define uniquely G is by using a bilinear form, then G is the orthogonal complement (see below) and the projection is the orthogonal projection.

Example : Let $(e_i)_{i=1..n}$ be a basis of a n dimensional vector space E. Take F the vector subspace generated by the first p e_i and G the vector subspace generated by the last n-p e_i . Obviously $E = F \oplus G$. But $G'_a = \{w = a(u+v), u \in G, v \in F\}$ for any fixed $a \in K$ is such that : $E = F \oplus G'_a$

Product of vector spaces

These are obvious objects, but with subtle points.

1. Product of two vector spaces

Theorem 225 *If E,F are vectors spaces over the same field K, the product set ExF can be endowed with the structure of a vector space over K with the operations :* $(u,v) + (u',v') = (u+u', v+v') ; k(u,v) = (ku, kv) ; 0 = (0,0)$

The subsets of ExF : E'=(u,0), F'=(0,v) are vector subspaces of ExF and we have ExF=E'$\oplus$F'.

Conversely, if E_1, E_2 are *vector subspaces* of E such that $E = E_1 \oplus E_2$ then to each vector of E can be associated its unique pair (u,v)$\in E_1 \times E_2$. Define $E'_1 = (u,0), E'_2 = (0,v)$ which are vector subspaces of $E_1 \times E_2$ and $E_1 \times E_2 = E'_1 \oplus E'_2$ but $E'_1 \oplus E'_2 \simeq E$. So in this case one can see the direct sum as the product $E_1 \times E_2 \simeq E_1 \oplus E_2$

In the converse, it is mandatory that $E = E_1 \oplus E_2$. Indeed take ExE, the product is well defined, but not the direct sum (it would be just E).

In a somewhat pedantic way : a vector subspace E_1 of a vector space E splits in E if : $E = E_1 \oplus E_2$ and E$\simeq E_1 \times E_2$ (Lang p.6)

2. Infinite product of vector spaces

This can be generalized to any product of vector spaces $(F_i)_{i \in I}$ over the same field where I is finite. If I is infinite this is a bit more complicated : first one must assume that all the vector spaces F_i belong to some universe.

One defines : $E_T = \cup_{i \in I} F_i$ (see set theory). Using the axiom of choice there are maps : $C : I \to E_T :: C(i) = u_i \in F_i$

One restricts E_T to the subset E of E_T comprised of elements such that only finitely many u_i are non zero. E can be endowed with the structure of a vector space and $E = \prod_{i \in I} F_i$

The identity $E = \oplus_{i \in I} E_i$ with $E_i = \{u_j = 0, j \neq i \in I\}$ does not hold any longer : it would be E_T.

But if the F_i are vector subspaces of some $E = \oplus_{i \in I} F_i$ which have only 0 as common element on can still write $\prod_{i \in I} F_i \simeq \oplus_{i \in I} F_i$

Quotient space

Definition 226 *The **quotient space**, denoted E/F, of a vector space E by any of its vector subspace F is the quotient set $E/\sim$ by the relation of equivalence : $x, y \in E : x - y \in F \Leftrightarrow x \equiv y \pmod{F}$*

It is a vector space on the same field.

The class [0] contains the vectors of F.

The mapping $E \to E/F$ that associates to $x \in E$ its class of equivalence [x] , called the quotient map, is a natural epimorphism, whose kernel is F. This relationship is summarized by the short exact sequence

$0 \to F \to E \to E/F \to 0$

The dimension of E/F is sometimes called the codimension. For finite dimensional vector spaces : dim(E/F)=dim(E) - dim(F)

If $E = F \oplus F'$ then E/F is isomorphic to F'

Graded vector spaces

Definition 227 *A I-**graded vector space** is a vector space E endowed with a family of filters $(E_i)_{i \in I}$ such that each E_i is a vector subspace of E and $E = \oplus_{i \in I} E_i$. A vector of E which belongs to a single E_i is said to be an **homogeneous element**.*

A linear map between two I-graded vector spaces f:E→F is called a graded linear map if it preserves the grading of homogeneous elements: $\forall i \in I : f(E_i) \subset F_i$

Usually the family is indexed on $\mathbb{N}$ and then the family is decreasing : $E_{n+1} \subset E_n$. The simplest example is $E_n =$ the vector subspace generated by the vectors $(e_i)_{i \geq n}$ of a basis. The graded space is $grE = \oplus_{n \in \mathbb{N}} E_n / E_{n+1}$

Cone

Definition 228 *A **cone** with apex a in a real vector space E is a non empty subset C of E such that : $\forall k \geq 0, u \in C \Rightarrow k(u - a) \in C$*

A cone C is proper if $C \cap (-C) = 0$. Then there is an order relation on E by : $X \geq Y \Leftrightarrow X - Y \in C$ thus :

$$X \geq Y \Rightarrow X + Z \geq Y + Z, k \geq 0 : kX \geq kY$$

Definition 229 *A **vectorial lattice** is a real vector space E endowed with an order relation for which it is a lattice :*

$\forall x, y \in E, \exists \sup(x, y), \inf(x, y)$
$x \leq y \Rightarrow \forall z \in E : x + z \leq y + z$
$x \geq 0, k \geq 0 \Rightarrow kx \geq 0$
On a vectorial lattice :
- the cone with apex a is the set : $C_a = \{v \in E : a \geq v\}$
- the sets :
$x_+ = \sup(x, 0); x_- = \sup(-x, 0), |x| = x_+ + x_-$
$a \leq b : [a, b] = \{x \in E : a \leq x \leq b\}$

6.2 Linear maps

6.2.1 Definitions

Definition 230 *A **linear map** is a morphism between vector spaces over the same field K :*

$f \in L(E; F) \Leftrightarrow f : E \to F :: \forall a, b \in K, \forall \overrightarrow{u}, \overrightarrow{v} \in E : g(a\overrightarrow{u} + b\overrightarrow{v}) = ag(\overrightarrow{u}) + bg(\overrightarrow{v}) \in F$

Warning ! To be fully consistent, the vector spaces E and F *must be defined over the same field K*. So if E is a real vector space and F a complex vector space we will not consider as a linear map a map such that : f(u+v)=f(u)+f(v), f(ku)=kf(u) for any k real. This complication is necessary to keep simple the more important definition of linear map. It will be of importance when K=$\mathbb{C}$.

If E=F then f is an **endomorphism.**

Theorem 231 *The composition of linear map between vector spaces over the same field is still a linear map, so vector spaces over a field K with linear maps define a category.*

Theorem 232 *The set of linear maps from a vector space to a vector space on the same field K is a vector space over K*

Theorem 233 *If a linear map is bijective then its inverse is a linear map and f is an **isomorphism.***

Definition 234 *Two vector spaces over the same field are isomorphic if there is an isomorphism between them.*

Theorem 235 *Two vector spaces over the same field are isomorphic iff they have the same dimension*

We will usually denote $E \simeq F$ if the two vector spaces E,F are isomorphic.

Theorem 236 *The set of endomorphisms of a vector space E, endowed with the composition law, is a unital algebra on the same field.*

Definition 237 *An endomorphism which is also an isomorphism is called an **automorphism**.*

Theorem 238 *The set of automorphisms of a vector space E, endowed with the composition law, is a group denoted GL(E).*

Notation 239 *L(E;F) with a semi-colon (;) before the codomain F is the set of linear maps* $\hom(E, F)$.
   GL(E;F) is the subset of invertible linear maps
   GL(E) is the set of automorphisms over the vector space E

Definition 240 *A linear endomorphism such that its k iterated, for some k>0 is null is said to be nilpotent :*

   $f \in L(E; E) : f \circ f \circ .. \circ f = (f)^k = 0$
   Let $(e_i)_{i \in I}$ be a basis of E over the field K, consider the set K^I of all maps from I to K : $\tau : I \to K :: \tau(i) = x_i \in K$ .Take the subset K_0^I of K^I such that only a finite number of $x_i \neq 0$. This a vector space over K.

   For any basis $(e_i)_{i \in I}$ there is a map : $\tau_e : E \to K_0^I :: \tau_e(i) = x_i$. This map is linear and bijective. So E is isomorphic to the vector space K_0^I. This property is fundamental in that whenever only linear operations over finite dimensional vector spaces are involved it is equivalent to consider the vector space K^n with a given basis. This is the implementation of a general method using the category theory : K_0^I is an object in the category of vector spaces over K. So if there is a functor acting on this category we can see how it works on K_0^I and the result can be extended to other vector spaces.

Definition 241 *If E,F are two complex vector spaces, an **antilinear map** is a map $f : E \to F$ such that :*
   $\forall u, v \in E, z \in \mathbb{C} : f(u + v) = f(u) + f(v) ; f(zu) = \overline{z} f(u)$

   Such a map is linear when z is limited to a real scalar.

6.2.2   Matrix of a linear map

(see the "Matrices" section below for more)
   Let $L \in L(E; F)$, E n dimensional, F p dimensional vector spaces with basis $(e_i)_{i=1}^n , (f_j)_{j=1}^p$ respectively
   The matrix of L in these bases is the matrix M , with p rows and n columns : row i, column j : $[M]_{ij}$ such that :
   $L(e_i) = \sum_{j=1}^p M_{ij} f_j$
   So that :
   $\forall u = \sum_{i=1}^n u_i e_i \in E : L(u) = \sum_{j=1}^p \sum_{i=1}^n (M_{ij} u_i) f_j$

$$\begin{bmatrix} v_1 \\ ... \\ v_p \end{bmatrix} = \begin{bmatrix} M_{11} & .. & M_{1n} \\ .. & .. & .. \\ M_{p1} & .. & M_{pn} \end{bmatrix} \begin{bmatrix} u_1 \\ .. \\ u_n \end{bmatrix} \Leftrightarrow v = L(u)$$

or with the vectors represented as column matrices : $[L(u)] = [M][u]$

The matrix of the composed map $L \circ L'$ is the product of the matrices MxM' (the dimensions must be consistent).

The matrix is square is dim(E)=dim(F). f is an isomorphism iff M is invertible (det(M) non zero).

Theorem 242 *A **change of basis** in a vector space is an endomorphism. Its matrix P has for columns the components of the new basis expressed in the old basis :* $\overrightarrow{e_i} \to \overrightarrow{E}_i = \sum_{j=1}^{n} P_{ij} \overrightarrow{e}_j$.*The new components U_i of a vector u are given by :* $[U] = [P]^{-1}[u]$

Proof. $\overrightarrow{u} = \sum_{i=1}^{n} u_i \overrightarrow{e_i} = \sum_{i=1}^{n} U_i \overrightarrow{E_i} \Leftrightarrow [u] = [P][U] \Leftrightarrow [U] = [P]^{-1}[u]$ ■

Theorem 243 *If a change of basis both in E and F the matrix of the map $L \in L(E;F)$ in the new bases becomes :* $[M'] = [Q]^{-1}[M][P]$

Proof. $\overrightarrow{f_i} \to \overrightarrow{F_i} = \sum_{j=1}^{p} Q_{ij} \overrightarrow{f_j}$

$\overrightarrow{v} = \sum_{i=1}^{p} v_i \overrightarrow{f_i} = \sum_{i=1}^{p} V_i \overrightarrow{F_i} \Leftrightarrow [v] = [Q][V] \Leftrightarrow [V] = [Q]^{-1}[v]$

$[v] = [M][u] = [Q][V] = [M][P][U] \Rightarrow [V] = [Q]^{-1}[M][P][U]$

$[p,1] = [p,p] \times [p,n] \times [n,n] \times [n,1]$

$\Rightarrow [M'] = [Q]^{-1}[M][P]$ ■

If L is an endomorphism then P=Q, and $[M'] = [P]^{-1}[M][P] \Rightarrow \det M' = \det M$

An obvious, but most convenient, result : a vector subspace F of E is generated by a basis of r vectors f_j, expressed in a basis e_i of E by a nxr matrix $[A]$: $u \in F \Leftrightarrow u = \sum_{j=1}^{r} x_j f_j = \sum_{i=1}^{n} \sum_{j=1}^{r} x_i A_{ji} e_j = \sum_{i=1}^{n} u_i e_i$

so : $u \in F \Leftrightarrow \exists [x] : [u] = [A][x]$

6.2.3   Eigen values

Definition

Definition 244 *An **eigen vector** of the endomorphism $f \in L(E;E)$ with **eigen value** $\lambda \in K$ is a vector $u \neq 0$ such that $f(u) = \lambda u$*

Warning !

i) An eigen vector is non zero, but an eigen value can be zero.

ii) A linear map may have or not eigen values.

iii) the eigen value must belong to the field K

Fundamental theorems

Theorem 245 *The eigenvectors of an endomorphism* $f \in L(E; E)$ *with the same eigenvalue* λ, *form, with the vector 0, a vector subspace* E_λ *of E called an* **eigenspace**.

Theorem 246 *The eigenvectors corresponding to different eigenvalues are linearly independent*

Theorem 247 *If* u, λ *are eigen vector and eigen value of f, then, for k>0, u and* λ^k *are eigen vector and eigen value of* $(\circ f)^k$ *(k-iterated map)*

So f is nilpotent if its only eigen values are 0.

Theorem 248 *f is injective iff it has no zero eigen value.*

If E is finite dimensional, the eigen value and vectors are the eigen value and vectors of its matrix in any basis (see Matrices)

If E is infinite dimensional the definition stands but the main concept is a bit different : the spectrum of f is the set of scalars λ such that $(f - \lambda Id)$ has no bounded inverse. So an eigenvalue belongs to the spectrum but the converse is not true (see Banach spaces).

6.2.4 Rank of a linear map

Rank

Theorem 249 *The range f(E) of a linear map* $f \in L(E; F)$ *is a vector subspace of the codomain F.*

The **rank** *rank(f) of f is the dimension of f(E)* $\subset F$ *and rank(f)= dim f(E)* $\leq$ *dim(F)*

$f \in L(E; F)$ *is surjective iff f(E)=F, or equivalently if rank(f)=dimE*

Proof. f is surjective iff $\forall v \in F, \exists u \in E : f(u) = v \Leftrightarrow \dim f(E) = \dim F = rank(f)$ ∎

So the map : $\widetilde{f} : E \rightarrow f(E)$ is a linear surjective map L(E;f(E))

Kernel

Theorem 250 *The* **kernel,** *denoted* $\ker(f)$, *of a linear map* $f \in L(E; F)$ *is the set :* $\ker(f) = \{u \in E : f(u) = 0_F\}$.

It is a vector subspace of its domain E and

$\dim \ker(f) \leq \dim E$ *and if* $\dim \ker(f) = \dim E$ *then f=0*

f is injective if ker(f)=0

Proof. f is injective iff $\forall u_1, u_2 \in E : f(u_1) = f(u_2) \Rightarrow u_1 = u_2 \Leftrightarrow \ker(f) = 0_E$
∎

So with the quotient space E/ker(f) the map : $\widehat{f} : E/\ker f \to F$ is a linear injective map L(E/ker(f);F) (two vectors giving the same result are deemed equivalent).

Isomorphism

Theorem 251 *If $f \in L(E; F)$ then rank(f)$\leq \min(\dim E, \dim F)$ and f is an isomorphism iff rank(f)=dim(E)=dim(F)*

Proof. $g : E/\ker f \to f(E)$ is a linear bijective map, that is an isomorphism and we can write : $f(E) \simeq E/\ker(f)$

The two vector spaces have the same dimension thus :
dim(E/ker(f)) = dim E - dimker(f) = dimf(E) =rank(f)
rank(f)$\leq \min(\dim E, \dim F)$ and f is an isomorphism iff rank(f)=dim(E)=dim(F)
∎

To sum up

A linear map $f \in L(E; F)$ falls in one of the three following cases :
i) f is surjective : f(E)=F :
$\mathrm{rank}(f) = \dim f(E) = \dim F = \dim E - \dim \ker f \leq \dim E$ (F is "smaller" or equal to E)
In finite dimensions with dim(E)=n, dim(F)=p the matrix of f is $[f]_{n \times p}, p \leq n$

There is a linear bijection from E/ker(f) to F
ii) f is injective : ker(f)=0
dim E= dimf(E) =rank(f)$\leq \dim F$ (E is "smaller" or equal to F)
In finite dimensions with dim(E)=n, dim(F)=p the matrix of f is $[f]_{n \times p}, n \leq p$

There is a linear bijection from E to f(E)
iii) f is bijective : f(E)=F ,ker(f)=0, dimE=dimF=rank(f)
In finite dimensions with dim(E)=dimF=n, the matrix of f is square $[f]_{n \times n}$
and det $[f] \neq 0$

6.2.5   Multilinear maps

Definition 252 *A **r multilinear map** is a map : $f : E_1 \times E_2 \times ..E_r \to F$, where $(E_i)_{i=1}^r$ is a family of r vector spaces, and F a vector space, all over the same field K, which is linear with respect to each variable*

So :
$\forall u_i, v_i \in E_i, k_i \in K :$
$f(k_1 u_1, k_2 u_2, ..., k_r u_r) = k_1 k_2 ... k_r f(u_1, u_2, ... u_r)$
$f(u_1, u_2, ., u_i + v_i, .., u_r) = f(u_1, u_2, ... u_i, ..., u_r) + f(u_1, u_2, ... v_i, ..., u_r)$

Notation 253 $L^r(E_1, E_2..., E_r; F)$ *is the set of r-linear maps from* $E_1 \times E_2... \times E_r$ *to F*

$L^r(E; F)$ *is the set of r-linear map from* E^r *to F*

Warning ! $E_1 \times E_2$ can be endowed with the structure of a vector space. A *linear* map $f : E_1 \times E_2 \to F$ is such that :

$\forall (u_1, u_2) \in E_1 \times E_2 : (u_1, u_2) = (u_1, 0) + (0, u_2)$ so $f(u_1, u_2) = f(u_1, 0) + f(0, u_2)$

that can be written : $f(u_1, u_2) = f_1(u_1) + f_2(u_2)$ with $f_1 \in L(E_1; F), f_2 \in L(E_2; F)$

So : $L(E_1 \times E_2; F) \simeq L(E_1; F) \oplus L(E_2; F)$

Theorem 254 *The space* $L^r(E; F) \equiv L(E; L(E; ....L(E; F))$

Proof. For $f \in L^2(E, E; F)$ and u fixed $f_u : E \to F :: f_u(v) = f(u, v)$ is a linear map.

Conversely a map : $g \in L(E; L(E; F)) :: g(u) \in L(E; F)$ is equivalent to a bilinear map : $f(u, v) = g(u)(v)$ ∎

For E n dimensional and F p dimensional the components of the bilinear map f reads :

$f \in L^2(E; F) : f(u, v) = \sum_{i,j=1}^{n} u_i v_j f(e_i, e_j)$ with : $f(e_i, e_j) = \sum_{k=1}^{p} (F_{kij}) f_k, F_{kij} \in K$

A bilinear map cannot be represented by a single matrix if F is not unidimensional (meaning if F is not K). It is a tensor.

Definition 255 *A r-linear map* $f \in L^r(E; F)$ *is :*

symmetric *if :* $\forall u_i \in E, i = 1...r, \sigma \in \mathfrak{S}(r) : f(u_1, u_2, ..., u_r) = f(u_{\sigma(1)}, u_{\sigma(2)}, ..., u_{\sigma(r)})$

antisymmetric *if :* $\forall u_i \in E, i = 1...r, \sigma \in \mathfrak{S}_r : f(u_1, u_2, ..., u_r) = \epsilon(\sigma) f(u_{\sigma(1)}, u_{\sigma(2)}, ..., u_{\sigma(r)})$

6.2.6 Dual of a vector space

Linear form

A field K is endowed with the structure of a 1-dimensional vector space over itself in the obvious way, so one can consider morphisms from a vector space E to K.

Definition 256 *A **linear form** on a vector space E on the field K is a linear map valued in K*

A linear form can be seen as a linear function with argument a vector of E and value in the field K :$\varpi(u) = k$

Warning ! A linear form must be valued in the same field as E. A "linear form on a complex vector space and valued in $\mathbb{R}$" cannot be defined without a real structure on E.

Dual of a vector space

Definition 257 *The **algebraic dual** of a vector space is the set of its linear form, which has the structure of a vector space on the same field*

Notation 258 *E^* is the algebraic dual of the vector space E*

The vectors of the dual $(K^n)^*$ are usually represented as 1xn matrices (row matrices).

Theorem 259 *A vector space and its algebraic dual are isomorphic iff they are finite dimensional.*

This important point deserves some comments.

i) Consider first a finite finite n dimensional vector space E.

For each basis $(e_i)_{i=1}^n$ the **dual basis** $(e^i)_{i=1}^n$ of the dual E^* is defined by the condition : $e^i(e_j) = \delta_j^i$. where δ_j^i is the **Kronecker'symbol** =1 if i=j,=0 if not. These conditions define uniquely a basis of the dual, which is indexed on the same set I.

The map : $L : E \to E^* : L\left(\sum_{i\in I} u_i e_i\right) = \sum_{i\in I} u_i e^i$ is an isomorphism.

In a change of basis in E with matrix P (which has for columns the components of the new basis expressed in the old basis) :

$\overrightarrow{e_i} \to E_i = \sum_{j=1}^n P_{ij} e_j$, the dual basis changes as : $e^i \to E^i = \sum_{j=1}^n Q_{ij} e^j$ with $[Q] = [P]^{-1}$

Warning! This isomorphism is not canonical, even in finite dimensions, in that it depends of the choice of the basis.

$f : E \to E^* : u = \sum_{i=1}^n u_i e_i \to f(u) = \sum_{i=1}^n u_i e^i$

In another basis f(u) will not have the same simple components. In general *there is no natural transformation which is an isomorphism between a vector space and its dual*, even finite dimensional. So to define an isomorphism one uses a bilinear form (when there is one).

ii) Consider now an infinite dimensional vector space E over the field K.

Then $\dim(E^*) > \dim(E)$. For infinite dimensional vector spaces the algebraic dual E^* is a *larger set* then E.

Indeed if E has the basis $(e_i)_{i\in I}$ there is a map : $\tau_e : E \to K_0^I :: \tau_e(i) = x_i$ giving the components of a vector, in the set K_0^I of maps $I \to K$ such that only a finite number of components is non zero and $K_0^I \simeq E$. But any map : $\lambda : I \to K$ gives a linear map $\sum_{i\in I} \lambda(i) x_i$ which is well defined because only a finite number of terms are non zero, whatever the vector, and can represent a vector of the dual. So the dual $E^* \simeq K^I$ which is larger than K_0^I.

The condition $\forall i, j \in I : e^i(e_j) = \delta_j^i$ still defines a family $(e^i)_{i\in I}$ of linearly independant vectors of the dual E* but this is not a basis of E*. However there is always a basis of the dual, that we can denote $(e^i)_{i\in I'}$ with #I' > #I and one can require that $\forall i, j \in I : e^i(e_j) = \delta_j^i$

For infinite dimensional vector spaces one considers usually the topological dual which is the set of continuous forms over E. If E is finite dimensional the algebraic dual is the same as the topological dual.

Definition 260 *The **double dual** E^{**} of a vector space is the algebraic dual of E*. The double dual E^{**} is isomorphic to E iff E is finite dimensional*

There is a natural homomorphism ϕ from E into the double dual E^{**}, defined by the evaluation map : $(\phi(u))(\varpi) = \varpi(u)$ for all $v \in E, \varpi \in E^*$. This map ϕ is always injective so $E \sqsubseteq (E^*)^*$; it is an isomorphism if and only if E is finite-dimensional, and if so then E$\simeq$E**.

Definition 261 *The **annihiliator** S^{T} of a vector subspace S of E is the set :* $S^{\mathsf{T}} = \{\varphi \in E^* : \forall u \in S : \varphi(u) = 0\}.$

It is a vector subspace of E*. $E^{\mathsf{T}} = 0$; $S^{\mathsf{T}} + S'^{\mathsf{T}} \subset (S \cap S')^{\mathsf{T}}$

Transpose of a linear map

Theorem 262 *If E,F are vector spaces on the same field, $\forall f \in L(E; F)$ there is a unique map, called the (algebraic) **transpose** (called also dual) and denoted $f^t \in L(F^*; E^*)$ such that : $\forall \varpi \in F^* : f^t(\varpi) = \varpi \circ f$*

The relation $^t : L(E; F) \to L(F^*; E^*)$ is injective (whence the unicity) but not surjective (because E**$\neq E$ if E is infinite dimensional).

The functor which associes to each vector space its dual and to each linear map its transpose is a functor from the category of vector spaces over a field K to itself.

If the linear map f is represented by the matrix A with respect to two bases of E and F, then f^t is represented by the *same* matrix with respect to the dual bases of F* and E*. Alternatively, as f is represented by A acting on the left on column vectors, f^t is represented by the same matrix acting on the right on row vectors. So if vectors are always represented as matrix columns the matrix of f^t is the transpose of the matrix of f :

Proof. $\forall u, \lambda : [\lambda]^t [f^t]^t [u] = [\lambda]^t [f] [u] \Leftrightarrow [f^t] = [f]$ ∎

6.2.7   Bilinear forms

Definition 263 *A **multilinear form** is a multilinear map defined on vector spaces on a field K and valued in K.*

So a **bilinear form** g on a vector space E on a field K is a bilinear map on E valued on K:

$g : E \times E \to K$ is such that :

$\forall u, v, w \in E, k, k' \in K : g(ku, k'v) = kk'g(u, v), g(u + w, v) = g(u, v) + g(u, w), g(u, v + w) = g(u, v) + g(u, w)$

Notice that K can be any field.

Warning ! A multilinear form must be valued in the same field as E. A "multilinear form on a complex vector space and valued in $\mathbb{R}$" cannot be defined without a real structure on E.

Symmetric, antisymmetric forms

Definition 264 *A bilinear form g on a vector space E is*
symmetric *if :* $\forall u,v \in E : g(u,v) = g(v,u)$
antisymmetric *if :* $\forall u,v \in E : g(u,v) = -g(v,u)$

Any bilinear symmetric form defines the **quadratic form** : $Q : E \rightarrow K ::$
$Q(u) = g(u,u)$

Conversely $g(u,v) = \frac{1}{2}(Q(u+v) - Q(u) - Q(v))$ (called the **polarization formula**) defines the bilinear symmetric form g form Q.

Non degenerate bilinear forms

Definition 265 *A bilinear symmetric form* $g \in L^2(E^2; K)$ *is* **non degenerate**
if : $\forall v : g(u,v) = 0 \Rightarrow u = 0$

Warning ! one can have g(u,v)=0 with u,v non null.

Theorem 266 *A non degenerate symmetric bilinear form on a finite dimensional vector space E on a field K defines isomorphisms between E and its dual E*:*
$\forall \varpi \in E^*, \exists u \in E : \forall v \in E : \varpi(v) = g(u,v)$
$\forall u \in E, \exists \varpi \in E^* : \forall v \in E : \varpi(v) = g(u,v)$

This is the usual way to "map" vectors to forms and vice versa.

$L^2(E;K) \equiv L(E;L(E;K)) = L(E;E^*)$ .So to each bilinear form g are associated two maps :
$\phi_R : E \rightarrow E^* :: \phi_R(u)(v) = g(u,v)$
$\phi_L : E \rightarrow E^* :: \phi_L(u)(v) = g(v,u)$

which are identical if g is symmetric and opposite from each other if g is skew-symmetric.

If g is non degenerate then ϕ_R, ϕ_L are injective but they are surjective iff E is finite dimensional.

If E is finite dimensional g is non degenerate iff $\phi_R, \phi_L \in L(E;E^*)$ are isomorphisms. As E and its dual have the same dimension iff E is finite dimensional it can happen only if E is finite dimensional. The matrix expression is :
$[\phi_R(u)] = [\phi_L(u)] = [u]^t[g]$

Conversely if $\phi \in L(E;E^*)$ the bilinear forms are : $g_R(u,v) = \phi(u)(v) ; g_L(u,v) = \phi(v)(u)$

Remark : it is usual to say that g is non degenerate if $\phi_R, \phi_L \in L(E;E^*)$ are isomorphisms. The two definitions are equivalent if E is finite dimensional, but we will need non degeneracy for infinite dimensional vector spaces.

Matrix representation of a bilinear form

If E is finite dimensional g is represented in a basis $(e_i)_{i=1}^n$ by a square matrix nxn $[g_{ij}] = g(e_i, e_j)$ with : $g(u,v) = [u]^t [g] [v]$

The matrix $[g]$ is symmetric if g is symmetric, antisymmetric if g is antisymmetric, and its determinant is non zero iff g is non degenerate.

In a change of basis : the new matrice is $[G] = [P]^t [g] [P]$ where $[P]$ is the matrix with the components of the new basis :

$g(u,v) = [u]^t [g] [v], [u] = [P] [U], v = [P] [v] \Rightarrow g(u,v) = [U]^t [P]^t [g] [P] [V] \rightarrow [G] = [P]^t [g] [P]$

Positive bilinear forms

Definition 267 *A bilinear symmetric form g on a real vector space E is* **positive** *if:* $\forall u \in E : g(u,u) \geq 0$

A bilinear symmetric form g on a real vector space E is **definite positive** *if it is positive and* $\forall u \in E : g(u,u) = 0 \Rightarrow u = 0$

definite positive $\Rightarrow$ non degenerate . The converse is not true
Notice that E must be a real vector space.

Theorem 268 *(Schwartz I p.175) If the bilinear symmetric form g on a real vector space E is positive then* $\forall u, v \in E$

i) **Schwarz inegality** *:* $|g(u,v)| \leq \sqrt{g(u,u)g(v,v)}$

if g is positive definite $|g(u,v)| = \sqrt{g(u,u)g(v,v)} \Rightarrow \exists k \in \mathbb{R} : v = ku$

ii) **Triangular inegality** *:* $\sqrt{g(u+v,u+v)} \leq \sqrt{g(u,u)} + \sqrt{g(v,v)}$

$\sqrt{g(u+v,u+v)} = \sqrt{g(u,u)} + \sqrt{g(v,v)} \Leftrightarrow g(u,v) = 0$ *(Pythagore's theorem)*

6.2.8 Sesquilinear forms

Definition 269 *A* **sesquilinear** *form on a complex vector space E is a map* $g : E \times E \rightarrow \mathbb{C}$ *linear in the second variable and antilinear in the first variable:*

$g(\lambda u, v) = \overline{\lambda} g(u,v)$
$g(u + u', v) = g(u,v) + g(u',v)$

So the only difference with a bilinear form is the way it behaves by multiplication by a complex scalar in the first variable.

Remarks :

i) this is the usual convention in physics. One finds also sesquilinear = linear in the first variable, antilinear in the second variable

ii) if E is a real vector space then a bilinear form is the same as a sesquilinear form

The definitions for bilinear forms extend to sesquilinear forms. In most of the results transpose must be replaced by conjugate-transpose.

Hermitian forms

Definition 270 *A **hermitian form** is a sesquilinear form such that :* $\forall u, v \in E : g(v, u) = \overline{g(u, v)}$

Hermitian forms play the same role in complex vector spaces as the symmetric bilinear forms in real vector spaces. If E is a real vector space a bilinear symmetric form is a hermitian form.

The quadratic form associated to an hermitian form is : $Q : E \rightarrow \mathbb{R} ::$ $Q(u, u) = g(u, u) = \overline{g(u, u)}$

Definition 271 *A **skew hermitian form** (also called an anti-symmetric sesquilinear form) is a sesquilinear form such that :*
$\forall u, v \in E : g(v, u) = -\overline{g(u, v)}$

Notice that, on a complex vector space, there are also bilinear form (they must be C-linear), and symmetric bilinear form

Non degenerate hermitian form
To each sesquilinear form g are associated two *antilinear* maps :
$\phi_R : E \rightarrow E^* :: \phi_R(u)(v) = \overline{g(u, v)}$
$\phi_L : E \rightarrow E^* :: \phi_L(u)(v) = \overline{g(v, u)}$
which are identical if g is hermitian and opposite from each other if g is skew-hermitian.

Definition 272 *A hermitian form is **non degenerate** if* $\forall v \in E : g(u, v) = 0 \Rightarrow u = 0$

Warning ! one can have g(u,v)=0 with u,v non null.

Theorem 273 *A non degenerate hermitian form on a finite dimensional vector space defines the anti-isomorphism between E and E* :*
$\forall \varpi \in E^*, \exists u \in E : \forall v \in E : \varpi(v) = g(u, v)$
$\forall u \in E, \exists \varpi \in E^* : \forall v \in E : \varpi(v) = g(u, v)$

Matrix representation of a sequilinear form
If E is finite dimensional a sequilinear form g is represented in a basis $(e_i)_{i=1}^n$ by a square matrix nxn $[g_{ij}] = g(e_i, e_j)$ with : $g(u, v) = \overline{[u]}^t [g][v] = [u]^* [g][v]$

The matrix $[g]$ is hermitan $\left([g] = \overline{[g]}^t = [g]^*\right)$ if g is hermitian, antihermitian $\left([g] = -\overline{[g]}^t = -[g]^*\right)$ if g is skewhermitian, and its determinant is non zero iff g is no degenerate.

In a change of basis : the new matrice is $[G] = [P]^* [g][P]$ where $[P]$ is the matrix with the components of the new basis :
$g(u, v) = [u]^* [g][v], [u] = [P][U], v = [P][v] \Rightarrow g(u, v) = [U]^* [P]^* [g][P][V] \rightarrow [G] = [P]^* [g][P]$

Positive hermitian forms

As $g(u, u) \in \mathbb{R}$ for a hermitian form one can define positive (resp. definite positive) hermitian forms.

Definition 274 *A hermitian form g on a complex vector space E is :*
 positive if: $\forall u \in E : g(u, u) \geq 0$
 definite positive if $\forall u \in E : g(u, u) \geq 0, g(u, u) = 0 \Rightarrow u = 0$

And the Schwarz and triangular inegalities stand for positive hermitian forms :

Theorem 275 *(Schwartz I p.178) If g is a hermitian, positive form on a complex vector space E, then* $\forall u, v \in E$
 Schwarz inegality : $|g(u, v)| \leq \sqrt{g(u, u)g(v, v)}$
 Triangular inegality : $\sqrt{g(u + v, u + v)} \leq \sqrt{g(u, u)} + \sqrt{g(v, v)}$
 and if g is positive definite, in both cases the equality implies $\exists k \in \mathbb{C} : v = ku$

6.2.9 Adjoint of a map

Definition 276 *On a vector space E, endowed with a bilinear symmetric form g if E is real, a hermitian sesquilinear form g if E is complex, the **adjoint** of an endomorphism f with respect to g is the map* $f^* \in L(E; E)$ *such that* $\forall u, v \in E : g(f(u), v) = g(u, f^*(v))$

Warning ! the transpose of a linear map can be defined without a bilinear map, the adjoint is always defined with respect to a form.

Theorem 277 *On a vector space E, endowed with a bilinear symmetric form g if E is real, a hermitian sesquilinear form g if E is complex, which is non degenerate :*
 i) the adjoint of an endormorphism, if it exists, is unique and $(f^*)^* = f$
 ii) If E is finite dimensional any endomorphism has a unique adjoint

The matrix of f* is : $[f^*] = [g]^{-1} [f]^t [g]$ with $[g]$ the matrix of g
Proof. $([f][u])^* [g][v] = [u]^* [g][f^*][v] \Leftrightarrow [f]^* [g] = [g][f^*] \Leftrightarrow [f^*] = [g]^{-1} [f]^* [g]$
∎

And usually $[f^*] \neq [f]^*$

Self-adjoint, orthogonal maps

Definition 278 *An endomorphism f on a vector space E, endowed with a bilinear symmetric form g if E is real, a hermitian sesquilinear form g if E is complex, is:*
 ***self-adjoint** if it is equal to its adjoint :* $f^* = f \Leftrightarrow g(f(u), v) = g(u, f(v))$
 ***orthogonal** (real case),* ***unitary** (complex case) if it preserves the bilinear symmetric form g* $: g(f(u), f(v)) = g(u, v)$

If E is finite dimensional the matrix $[f]$ of a self adjoint map f is such that :
$$[f]^* [g] = [g] [f]$$

Theorem 279 *If the form g is non degenerate then for any unitary endomorphism : $f \circ f^* = f^* \circ f = Id$*

Proof. $\forall u, v : g\left(f\left(u\right), f\left(v\right)\right) = g\left(u, v\right) = g\left(u, f^* f\left(v\right)\right) \Rightarrow g\left(u, \left(Id - f^* f\right) v\right) = 0 \Rightarrow f^* f = Id$ ∎

Definition 280 *The **orthogonal group** denoted O(E,g) of a vector space E endowed with a non degenerate bilinear symmetric form g is the set of orthogonal invertible maps. The **special orthogonal group** denoted SO(E,g) is its subgroup comprised of elements with detf=1;*

*The **unitary group** denoted U(E,g) on a complex vector space E endowed with a hermitian sesquilinear form g is the set of unitary invertible maps denoted U(E,g). The **special unitary group** denoted SU(E,g) is its subgroup comprised of elements with detf=1;*

6.3 Scalar product on a vector space

Many interesting properties of vector spaces occur when there is some non degenerate bilinear form defined on them. Indeed the elementary geometry is defined in an euclidean space, and almost all the properties used in analysis require a metric. So these vector spaces deserve some attention.

There are 4 mains results : existence of orthonormal basis, partition of the vector space, orthogonal complement and isomorphism with the dual.

6.3.1 Definitions

Definition 281 *A **scalar product** on a vector space E on a field K is either a non degenerate, bilinear symmetric form g, or a non degenerate hermitian sesquilinear form g. This is an **inner product** if g is definite positive.*

If g is definite positive then g defines a metric and a norm over E and E is a normed vector space (see Topology). Moreover if E is complete (which happens if E is finite dimensional), it is a Hilbert space. If K=ℝ then E is an **euclidean space**.

If the vector space is finite dimensional the matrix $[g]$ is symmetric or hermitian and its eigen values are all distinct, real and non zero. Their signs defines the **signature** of g. g is definite positive iff all the eigen values are >0.

If K=ℝ then the p in the signature of g is the maximum dimension of the vector subspaces where g is definite positive

With E 4 real dimensional and g the Lorentz metric of signature + + + - E is the Minkowski space of Relativity Theory (remark : quite often in physics

the chosen signature is - - - +, all the following results still stand with the appropriate adjustments).

Definition 282 *An **isometry** is a linear map $f \in L(E; F)$ between two vector spaces $(E, g), (F, h)$ endowed with scalar products, which preserves the scalar product : $\forall u, v \in E, g(f(u), f(v)) = h(u, v)$*

6.3.2 Orthonormal basis

Definition 283 *Two vectors u,v of a vector space endowed with a scalar product are **orthogonal** if g(u,v)=0.*

A vector u and a subset A of a vector space (E,g) endowed with a scalar product are orthogonal if $\forall v \in A, g(u, v) = 0$.

Two subsets A and B of a vector space (E,g) endowed with a scalar product are orthogonal if $\forall u \in A, v \in B, g(u, v) = 0$

Definition 284 *A basis $(e_i)_{i \in I}$ of a vector space (E,g) endowed with a scalar product, such that $\forall i, j \in I : g(e_i, e_j) = \pm \delta_{ij}$ is **orthonormal**.*

Notice that we do not require $g(e_i, e_j) = 1$

Theorem 285 *A finite dimensional vector space (E,g) endowed with a scalar product has orthonormal bases. If E is euclidian $g(e_i, e_j) = \delta_{ij}$. If $K = \mathbb{C}$ it is always possible to choose the basis such that $g(e_i, e_j) = \delta_{ij}$.*

Proof. the matrix $[g]$ is diagonalizable : there are matrix P either orthogonal or unitary such that $[g] = [P]^{-1} [\Lambda] [P]$ with $[P]^{-1} = [P]^* = \overline{[P]}^t$ and $[\Lambda] = Diag(\lambda_1, ... \lambda_n)$ the diagonal matrix with the eigen values of P which are all real.

In a change the basis with new components given by $[P]$, the form is expressed in the matrix $[\Lambda]$

If $K = \mathbb{R}$ take as new basis $[P][D]$ with $[D] = Diag\left(sgn(\lambda_i) \sqrt{|\lambda_i|}\right)$.

If $K = \mathbb{C}$ take as new basis $[P][D]$ with $[D] = Diag(\mu_i)$, with $\mu_i = \sqrt{|\lambda_i|}$ if $\lambda_i > 0, \mu_i = i\sqrt{|\lambda_i|}$ if $\lambda_i < 0$ ∎

In an orthonormal basis g takes the following form (expressed in the components of this basis):

If $K = \mathbb{R}$:$g(u, v) = \sum_{i=1}^n \epsilon_i u_i v_i$ with $\epsilon_i = \pm 1$
If $K = \mathbb{C}$: $g(u, v) = \sum_{i=1}^n \overline{u}_i v_i$
(remember that $u_i, v_i \in K$)

Notation 286 $\eta_{ij} = \pm 1$ *denotes usually the product $g(e_i, e_j)$ for an orthonormal basis and*

$[\eta]$ *is the diagonal matrix $[\eta_{ij}]$*

As a consequence (take orthonormal basis in each vector space):

- all complex vector spaces with hermitian non degenerate form and the same dimension are isometric.

- all real vector spaces with symmetric bilinear form of identical signature and the same dimension are isometric.

6.3.3   Time like and space like vectors

On a real vector space the bilinear form g, it is not definite positive, gives a partition of the vector space.in three subsets which can be or not connected.

1. The quantity g(u,u) is always real, it can be >0, <0,or 0. The sign does not depend on the basis. So one distinguishes the vectors according to the sign of g(u,u) :

- **time-like vectors** : g(u,u)<0
- **space-like vectors** : g(u,u)>0
- **null vectors** : g(u,u)=0

Remark : with the Lorentz metric the definition varies with the basic convention used to define g. The definitions above hold with the signature + + + - . In Physics usually g has the signature - - - + and then time-like vectors are such that g(u,u)>0.

The sign does no change if one takes $u \to ku, k > 0$ so these sets of vectors are half-cones. The cone of null vectors is commonly called the light-cone (as light rays are null vectors).

2. This theorem is new.

Theorem 287 *If g has the signature (+p,-q) a vector space (E,g) endowed with a scalar product is partitioned in 3 subsets :*

E_+ :space-like vectors, open, arc-connected if p>1, with 2 connected components if p=1

E_- : time-like vectors, open, arc-connected if q>1, with 2 connected components if q=1

E_0 : null vectors, closed, arc-connected

Openness an connectedness are topological concepts, but we place this theorem here as it fits the story.

Proof. It is clear that the 3 subsets are disjoint and that their union is E. g being a continuous map E_+ is the inverse image of an open set, and E_0 is the inverse image of a closed set.

For arc-connectedness we will exhibit a continuous path internal to each subset. Choose an orthonormal basis ε_i (with p+ and q- even in the complex case). Define the projections over the first p and the last q vectors of the basis :

$u = \sum_{i=1}^n u^i \varepsilon_i \to P_h(u) = \sum_{i=1}^p u^i \varepsilon_i; P_v(u) = \sum_{i=p+1}^n u^i \varepsilon_i$

and the real valued functions : $f_h(u) = g(P_h(u), P_h(u)); f_v(u) = g(P_v(u), P_v(u))$

so : $g(u, u) = f_h(u) - f_v(u)$

Let be $u_a, u_b \in E_+ : f_h(u_a) - f_v(u_a) > 0, f_h(u_b) - f_v(u_b) > 0 \Rightarrow f_h(u_a), f_h(u_b) > 0$

Define the path $x(t) \in E$ with 3 steps:

a) $t = 0 \to t = 1 : x(0) = u_a \to x(1) = \left(u_a^h, 0\right)$

$x(t) : i \le p : x^i(t) = u_a^i$;if p>1 $: i > p : x^i(t) = (1-t)u_a^i$

$g(x(t), x(t)) = f_h(u_a) - (1-t)^2 f_v(u_a) > f_h(u_a) - f_v(u_a) = g(u_a, u_a) > 0 \Rightarrow$
$x(t) \in E_+$

b) $t = 1 \to t = 2 : x(1) = \left(u_a^h, 0\right) \to x(1) = \left(u_b^h, 0\right)$

$x(t) : i \le p : x^i(t) = (t-1)u_b^i + (2-t)u_a^i = u_a^i$;if p>1$:i > p : x^i(t) = 0$

$g(x(t), x(t)) = f_h((t-1)u_b + (2-t)u_a) > 0 \Rightarrow x(t) \in E_+$

c) $t = 2 \to t = 3 : x(2) = \left(u_b^h, 0\right) \to x(3) = u_b$

$x(t) : i \le p : x^i(t) = u_b^i$;if p>1$:i > p : x^i(t) = (t-2)u_b^i$

$g(x(t), x(t)) = f_h(u_b) - (t-2)^2 f_v(u_b) > f_h(u_b) - f_v(u_b) = g(u_b, u_b) > 0 \Rightarrow$
$x(t) \in E_+$

So if $u_a, u_b \in E_+, x(t) \subset E_+$ whenever p>1.

For E_- we have a similar demonstration.

If q=1 one can see that the two regions t<0 and t>0 cannot be joined : the component along ε_n must be zero for some t and then g(x(t),x(t))=0

If $u_a, u_b \in E_0 \Leftrightarrow f_h(u_a) = f_v(u_a), f_h(u_b) = f_v(u_b)$

The path comprises of 2 steps going through 0 :

a) $t = 0 \to t = 1 : x(t) = (1-t)u_a \Rightarrow g(x(t)) = (1-t)^2 g(u_a, u_a) = 0$

b) $t = 1 \to t = 2 : x(t) = (t-1)u_b \Rightarrow g(x(t)) = (1-t)^2 g(u_b, u_b) = 0$

This path does always exist. ∎

3. The partition of E_- in two disconnected components is crucial, because it gives the distinction between "past oriented" and "future oriented" time-like vectors (one cannot go from one region to the other without being in trouble). This theorem shows that the Lorentz metric is special, in that it is the only one for which this distinction is possible.

One can go a little further. One can show that there is always a vector subspace F of dimension $\min(p, q)$ such that all its. vectors are null vectors. In the Minkowski space the only null vector subspaces are 1-dimensional.

6.3.4   Induced scalar product

Let be F a vector subspace, and define the form $h : F \times F \to K :: \forall u, v \in F :$ $h(u, v) = g(u, v)$. that is the restriction of g to F. h has the same linearity or anti-linearity as g. If F is defined by the nxr matrix A $(u \in F \Leftrightarrow [u] = [A][x])$, then h has the matrix $[H] = [A]^t [g] [A]$.

If g is definite positive, so is h and (F,h) is endowed with an inner product induced by g on F

If not, h can be degenerate,because there are vector subspaces of null-vectors, and its signature is usually different

Definition 288 *A vector subspace, denoted $F^\perp$, of a vector space E endowed with a scalar product is an* **orthogonal complement** *of a vector subspace F of E if $F^\perp$ is orthogonal to F and $E = F \oplus F^\perp$.*

If E is finite dimensional there are always orthogonal vector spaces F' and $\dim F + \dim F' = \dim E$ (Knapp p.50) but we have not necessarily $E = F \oplus F^{\perp}$ (see below) and they are not necessarily unique.

Theorem 289 *In a vector space endowed with an inner product the orthogonal complement always exist and is unique.*

This theorem is important : if F is a vector subspace there is always a vector space B such that $E = F \oplus B$ but B is not unique. This decomposition is useful for many purposes, and it is an hindrance when B cannot be defined more precisely. This is just what g does : $A^{\perp}$ is the orthogonal projection of A. But the theorem is not true if g is not definite positive. The problem of finding the orthogonal complement is linked to the following : starting from a given basis $(e_i)_{i=1}^{n}$ how can we compute an orthonormal basis $(\varepsilon_i)_{i=1}^{n}$? This is the so-called "**Graham-Schmitt's procedure**":

Find a vector of the basis which is not a null-vector. If all the vectors of the basis are null vectors then g=0 on the vector space.

So let be : $\varepsilon_1 = \frac{1}{g(e_1,e_1)} e_1$

Then by recursion : $\varepsilon_i = e_i - \sum_{j=1}^{i-1} \frac{g(e_i,\varepsilon_j)}{g(\varepsilon_j,\varepsilon_j)} \varepsilon_j$

All the ε_i are linearly independant. They are orthogonal :

$g\left(\varepsilon_i,\varepsilon_k\right) = g\left(e_i,\varepsilon_k\right) - \sum_{j=1}^{i-1} \frac{g(e_i,\varepsilon_j)}{g(\varepsilon_j,\varepsilon_j)} g\left(\varepsilon_j,\varepsilon_k\right) = g\left(e_i,\varepsilon_k\right) - \frac{g(e_i,\varepsilon_k)}{g(\varepsilon_k,\varepsilon_k)} g\left(\varepsilon_k,\varepsilon_k\right) = 0$

The only trouble that can occur is if for some i : $g\left(e_i,e_i\right) = g\left(e_i,e_i\right) - \sum_{j=1}^{i-1} \frac{g(e_i,\varepsilon_j)^2}{g(\varepsilon_j,\varepsilon_j)} = 0$. But from the Schwarz inegality :

$g\left(e_i,\varepsilon_j\right)^2 \leq g\left(\varepsilon_j,\varepsilon_j\right) g\left(e_i,e_i\right)$

and, if g is positive definite, equality can occur only if ε_i is a linear combination of the ε_j.

So if g is positive definite the procedure always works.

To find the orthogonal complement of a vector subspace F start with a basis of E such that the first r vectors are a basis of F. Then if there is an orthonormal basis deduced from (e_i) the last n-r vectors are an orthonormal basis of the unique orthogonal complement of F. If g is not positive definite there is not such guaranty.

6.4   Symplectic vector spaces

If the symmetric biliner form of the scalar product is replaced by an antisymmetric form we get a symplectic structure. In many ways the results are similar, and even stronger : all symplectic vector spaces of same dimension are indistiguishable. Symplectic spaces are commonly used in lagrangian mechanics.

6.4.1 Definitions

Definition 290 *A **symplectic vector space** (E,h) is a real vector space E endowed with a non degenerate antisymmetric 2-form h called the **symplectic form***

$$\forall u, v \in E : h(u, v) = -h(v, u) \in \mathbb{R}$$
$$\forall u \in E : \forall v \in E : h(u, v) = 0 \Rightarrow u = 0$$

Definition 291 *2 vectors u,v of a symplectic vector space (E,h) are **orthogonal** if h(u,v)=0.*

Theorem 292 *The set of vectors orthogonal to all vectors of a vector subspace F of a symplectic vector space is a vector subspace denoted $F^\perp$*

Definition 293 *A vector subspace is :*
     ***isotropic** if $F^\perp \subset F$*
     ***co-isotropic** if $F \subset F^\perp$*
     ***self-orthogonal** if $F^\perp = F$*

    The 1-dimensional vector subspaces are isotropic
    An isotropic vector subspace is included in a self-orthogonal vector subspace

Theorem 294 *The symplectic form of symplectic vector space (E,h) induces a map $j : E^* \to E :: \lambda(u) = h(j(\lambda), u)$ which is an isomorphism iff E is finite dimensional.*

6.4.2 Canonical basis

The main feature of symplectic vector spaces if that they admit basis in which any symplectic form is represented by the same matrix. So all symplectic vector spaces of the same dimension are isomorphic.

Theorem 295 *(Hofer p.3) A symplectic (E,h) finite dimensional vector space must have an even dimension n=2m. There are always canonical bases $(\varepsilon_i)_{i=1}^{n}$ such that $h(\varepsilon_i, \varepsilon_j) = 0, \forall |i - j| < m, h(\varepsilon_i, \varepsilon_j) = \delta_{ij}, \forall |i - j| > m$. All finite dimensional symplectic vector space of the same dimension are isomorphic.*

    h reads in any basis : $h(u, v) = [u]^t [h][v]$, with $[h] = [h_{ij}]$ skew-symmetric and det(h)$\neq$ 0.
    In a canonical basis:
$[h] = J_m = \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix}$ so $J_m^2 = -I_{2m}$
$h(u, v) = [u]^t [J_m][v] = \sum_{i=1}^{m} (u_i v_{i+m} - u_{i+m} v_i)$
    The vector subspaces E_1 spanned by $(\varepsilon_i)_{i=1}^{m}$, E_2 spanned by $(\varepsilon_i)_{i=m+1}^{2m}$ are self-orthogonal and $E = E_1 \oplus E$

6.4.3   Symplectic maps

Definition 296 *A symplectic map (or **symplectomorphism***) *between two symplectic vector spaces* $(E_1,h_1),(E_2,h_2)$, *is a linear map* $f \in L(E_1;E_2)$ *such that* $\forall u,v \in E_1 : h_2(f(u),f(v)) = h_1(u,v)$

f is injective so $\dim E_1 \leq \dim E_2$

Theorem 297 *(Hofer p.6) There is always a bijective symplectomorphism between two symplectic vector spaces* $(E_1,h_1),(E_2,h_2)$ *of the same dimension*

So all symplectic vector spaces of the same dimension are indistiguishable.

Definition 298 *A symplectic map (or **symplectomorphism***) *of a symplectic vector space (E,h) is an endomorphism of E which preserves the symplectic form* $h : f \in L(E;E) : h(f(u),f(v)) = h(u,v)$

Theorem 299 *The symplectomorphisms over a symplectic vector space (E,h) constitute the **symplectic group** Sp(E,h).*

In a canonical basis a symplectomorphism is represented by a symplectic matrix A which is such that :
$A^t J_m A = J_m$
because : $h(f(u),f(v)) = (A[u])^t J_m [A[v]] = [u]^t A^t J_m A [v] = [u]^t J_m [v]$
so $\det A = 1$

Definition 300 *The **symplectic group** Sp(2m) is the linear group of 2mx2m real matrices A such that :* $A^t J_m A = J_m$

$A \in Sp(2m) \Leftrightarrow A^{-1}, A^t \in Sp(2m)$

6.4.4   Liouville form

Definition 301 *The **Liouville form** on a 2m dimensional symplectic vector space (E,h) is the 2m form :* $\varpi = \frac{1}{m!} h \wedge h \wedge .. \wedge h$ *(m times). Symplectomorphisms preserve the Liouville form.*

In a canonical basis :
$\varpi = \varepsilon^1 \wedge \varepsilon^{m+1} \wedge .. \wedge \varepsilon^m \wedge \varepsilon^{2m}$
Proof. Put : $h = \sum_{i=1}^m \varepsilon^i \wedge \varepsilon^{i+m} = \sum_{i=1}^m h_i$
$h_i \wedge h_j = 0$ if i=j so $(\wedge h)^m = \sum_{\sigma \in \mathfrak{S}_m} h_{\sigma(1)} \wedge h_{\sigma(2)} .. \wedge h_{\sigma(m)}$
remind that : $h_{\sigma(1)} \wedge h_{\sigma(2)} = (-1)^{2 \times 2} h_{\sigma(2)} \wedge h_{\sigma(1)} = h_{\sigma(2)} \wedge h_{\sigma(1)}$
$(\wedge h)^m = m! \sum_{\sigma \in \mathfrak{S}_m} h_1 \wedge h_2 .. \wedge h_m$ ∎

6.4.5 Complex structure

Theorem 302 *A finite dimensional symplectic vector space (E,h) admits a complex structure*

Take a canonical basis and define : $J : E \rightarrow E :: J\left(\sum_{i=1}^{m} u_i \varepsilon_i + v_i \varphi_i\right) = \sum_{i=1}^{m} \left(-v_i \varepsilon_i + u_i \varphi_i\right)$ So $J^2 = -Id$ (see below)

It sums up to take as complex basis :$(\varepsilon_j, i\varepsilon_{j+m})_{j=1}^{m}$ with complex components. Thus E becomes a m-dimensional complex vector space.

6.5 Complex vector spaces

Complex vector spaces are vector spaces over the field $\mathbb{C}$. They share all the properties listed above, but have some specificities linked to :

- passing from a vector space over $\mathbb{R}$ to a vector space over $\mathbb{C}$ and vice versa
- the definition of the conjugate of a vector

6.5.1 From complex to real vector space

In a complex vector space E the restriction of the multiplication by a scalar to real scalars gives a real vector space, but as a set one must distinguish the vectors u and iu : we need some rule telling which are "real" vectors and "imaginary" vectors *in the same set of vectors.*

There is always a solution but it is not unique and depends on a specific map.

Real structure

Definition 303 *A **real structure** on a complex vector space E is a map :* $\sigma : E \rightarrow E$ *which is antilinear and such that* $\sigma^2 = Id_E$:

$$z \in \mathbb{R}, u \in E, \sigma\left(zu\right) = \overline{z}\sigma\left(u\right) \Rightarrow \sigma^{-1} = \sigma$$

Theorem 304 *There is always a real structure σ on a complex vector space E. Then E is the direct sum of two real vector spaces :* $E = E_{\mathbb{R}} \oplus iE_{\mathbb{R}}$ *where* $E_{\mathbb{R}}$ *, called the **real kernel** of σ, is the subset of vectors invariant by σ*

i) There is always a real structure

Proof. Take any (complex) basis $(e_j)_{j \in I}$ of E and define the map : $\sigma\left(e_j\right) = e_j, \sigma\left(ie_j\right) = -ie_j$

$\forall u \in E : u = \sum_{j \in I} z_j e_j \rightarrow \sigma\left(u\right) = \sum_{j \in I} \overline{z}_j e_j$

$\sigma^2\left(u\right) = \sum_{j \in I} z_j e_j = \sigma\left(u\right)$

It is antilinear :

$\sigma\left((a+ib)u\right) = \sigma\left(\sum_{j \in I}(a+ib)z_j e_j\right) = (a-ib)\sum_{j \in I} \sigma\left(z_j e_j\right) = (a-ib)\sigma\left(u\right)$

∎

This structure is not unique and depends on the choice of a basis.

ii) There is a subset $E_{\mathbb{R}}$ of E which is a real vector subspace of E

Proof. Define $E_{\mathbb{R}}$ as the subset of vectors of E invariant by $\sigma : E_{\mathbb{R}} = \{u \in E : \sigma(u) = u\}$.

It is not empty : with the real structure above any vector with real components in the basis $(e_j)_{j \in I}$ belongs to $E_{\mathbb{R}}$

It is a *real vector subspace* of E. Indeed the multiplication by a real scalar gives : $ku = \sigma(ku) \in E_{\mathbb{R}}$. ∎

iii) E=$E_{\mathbb{R}} \oplus iE_{\mathbb{R}}$

Proof. Define the maps :

$\operatorname{Re} : E \to E_{\mathbb{R}} :: \operatorname{Re} u = \frac{1}{2}(u + \sigma(u))$

$\operatorname{Im} : E \to E_{\mathbb{R}} :: \operatorname{Im} u = \frac{1}{2i}(u - \sigma(u))$

Any vector can be uniquely written with a real and imaginary part : $u \in E$: $u = \operatorname{Re} u + i \operatorname{Im} u$ which both belongs to the real kernel of E. Thus : $E = E_{\mathbb{R}} \oplus iE_{\mathbb{R}}$ ∎

E can be seen as a real vector space with two fold the dimension of E : $E_{\sigma} = E_{\mathbb{R}} \times iE_{\mathbb{R}}$

Conjugate

Warning ! The definition of the conjugate of a vector makes sense only iff E is a complex vector space endowed with a real structure.

Theorem 305 *The conjugate of a vector u on a complex vector space E endowed with a real structure σ is $\sigma(u) = \overline{u}$*

Proof. $: E \to E : \overline{u} = \operatorname{Re} u - i \operatorname{Im} u = \frac{1}{2}(u + \sigma(u)) - i\frac{1}{2i}(u - \sigma(u)) = \sigma(u)$ ∎

Remark : some authors (Wald) define the vector space conjugate $\overline{E}$ to a complex vector space E as the algebraic dual of the vector space of antilinear linear forms over E. One of the objective is to exhibit "mixed tensors" on the tensorial product $E \otimes E^* \otimes \overline{E} \otimes \overline{E}^*$. The algebraic dual E* of a vector space being larger than E, such a construct can be handled safely only if E is finite dimensional. The method presented here is valid whatever the dimension of E. And as one can see conjugation is an involution on E, and $\overline{E} = E$.

Real form

Definition 306 *A real vector space F is a **real form** of a complex vector space E if F is a real vector subspace of E and there is a real structure σ on E for which F is invariant by σ.*

Then E can be written as : $E = F \oplus iF$

As any complex vector space has real structures, there are always real forms, which are not unique.

6.5.2 From real to complex vector space

There are two different ways for endowing a real vector space with a complex vector space.

Complexification

The simplest, and the most usual, way is to enlarge the real vector space itself (as a set). This is always possible and called **complexification.**

Theorem 307 *For any real vector space E there is a structure $E_{\mathbb{C}}$ of complex vector space on ExE, called the **complexification** of E, such that $E_{\mathbb{C}} = E \oplus iE$*

Proof. ExE is a real vector space with the usual operations :

$\forall u, v, u', v' \in E, k \in \mathbb{R} : (u, v) + (u', v') = (u + u', v + v') ; k(u, v) = (ku, kv)$

We add the operation : $i(u, v) = (-v, u)$. Then : $z = a + ib \in \mathbb{C} : z(u, v) = (au - vb, av + bu) \in E \times E$

$i(i(u, v)) = i(-v, u) = -(u, v)$

ExE becomes a vector space $E_{\mathbb{C}}$ over $\mathbb{C}$. This is obvious if we denote : $(u, v) = u + iv$

The direct sum of two vector spaces can be identified with a product of these spaces, so $E_{\mathbb{C}}$ is defined as :

$E_{\mathbb{C}} = E \oplus iE \Leftrightarrow \forall u \in E_{\mathbb{C}}, \exists v, w$ unique $\in E : u = v + iw$ or $u = \operatorname{Re} u + i \operatorname{Im} u$ with $\operatorname{Re} u, \operatorname{Im} u \in E$ ∎

So E and iE are real vector subspaces of $E_{\mathbb{C}}$.

Remark : the complexified is often defined as $E_{\mathbb{C}} = E \otimes_R \mathbb{C}$ the tensoriel product being understood as acting over $\mathbb{R}$. The two definitions are equivalent, but the second is less enlighting...

Definition 308 *The **conjugate** of a vector of $E_{\mathbb{C}}$ is defined by the antilinear map : : $E_{\mathbb{C}} \to E_{\mathbb{C}} :: u \to \operatorname{Re} u - i \operatorname{Im} u$*

Theorem 309 *Any basis $(e_j)_{j \in I}$ of a real vector space E is a basis of the complexified $E_{\mathbb{C}}$ with complex components. $E_{\mathbb{C}}$ has same complex dimension as E.*

As a set $E_{\mathbb{C}}$ is "larger" than E : indeed it is defined through ExE, the vectors $e_i \in E$, and $ie_j \in E_{\mathbb{C}}$ but $ie_j \notin E$. To define a vector in $E_{\mathbb{C}}$ we need two vectors in E. However $E_{\mathbb{C}}$ has the same *complex* dimension as the real vector space E :a complex component needs two real scalars.

Theorem 310 *Any linear map $f \in L(E; F)$ between real vector spaces has a unique prolongation $f_{\mathbb{C}} \in L(E_{\mathbb{C}}; F_{\mathbb{C}})$*

Proof. i) If $f_{\mathbb{C}} \in L(E_{\mathbb{C}}; F_{\mathbb{C}})$ is a C-linear map : $f_{\mathbb{C}}(u + iv) = f_{\mathbb{C}}(u) + if_{\mathbb{C}}(v)$ and if it is the prolongation of f : $f_{\mathbb{C}}(u) = f(u), f_{\mathbb{C}}(v) = f(v)$

ii) $f_{\mathbb{C}}(u + iv) = f(u) + if(v)$ is C-linear and the obvious prolongation of f. ∎

If $f \in L(E; E)$ has $[f]$ for matrix in the basis $(e_i)_{i \in I}$ then its extension $f_c \in L(E_{\mathbb{C}}; E_{\mathbb{C}})$ has the *same* matrix in the basis $(e_i)_{i \in I}$ .. This is exactly what is done to compute the complex eigen values of a real matrix.

Notice that $L(E_{\mathbb{C}}; E_{\mathbb{C}}) \neq (L(E; E))_{\mathbb{C}}$ which is the set : $\{F = f + ig, f, g \in L(E; E)\}$ of maps from E to $E_{\mathbb{C}}$

Similarly $(E_{\mathbb{C}})^* = \{F; F(u + iv) = f(u) + if(v), f \in E^*\}$

and $(E^*)_{\mathbb{C}} = \{F = f + ig, f, g \in E^*\}$

Complex structure

The second way leads to define a complex vector space structure $E_{\mathbb{C}}$ on the *same* set E :

i) the sets are the same : if u is a vector of E it is a vector of $E_{\mathbb{C}}$ and vice versa

ii) the operations (sum and product by a scalar) defined in $E_{\mathbb{C}}$ are closed over $\mathbb{R}$ and $\mathbb{C}$

So the goal is to find a way to give a meaning to the operation : $\mathbb{C} \times E \to E$ and it would be enough if there is an operation with $i \times E \to E$

This is not always possible and needs the definition of a special map.

Definition 311 *A complex structure on a real vector space is a linear map* $J \in L(E; E)$ *such that* $J^2 = -Id_E$

Theorem 312 *A real vector space can be endowed with the structure of a complex vector space iff there a complex structure.*

Proof. a) the condition is necessary :

If E has the structure of a complex vector space then the map : $J : E \to E :: J(u) = iu$ is well defined and $J^2 = -Id$

b) the condition is sufficient :

What we need is to define the multiplication by i such that it is a complex linear operation :

Define on E : $iu = J(u)$. Then $i \times i \times u = -u = J(J(u)) = J^2(u) = -u$ ∎

Theorem 313 *A real vector space has a complex structure iff it has a dimension which is infinite or finite even.*

Proof. a) Let us assume that E has a complex structure, then it can be made a complex vector space and $E = E_{\mathbb{R}} \oplus iE_{\mathbb{R}}$. The two real vector spaces $E_{\mathbb{R}}, iE_{\mathbb{R}}$ are real isomorphic and have same dimension, so $\dim E = 2 \dim E_{\mathbb{R}}$ is either infinite or even ∎

b) The condition is sufficient :

Proof. Pick any basis $(e_{i\in I})_{i\in I}$ of E. If E is finite dimensional or countable we can order I according to the ordinal number, and define the map :

$J(e_{2k}) = e_{2k+1}$

$J(e_{2k+1}) = -e_{2k}$

It meets the condition :

$J^2(e_{2k}) = J(e_{2k+1}) = -e_{2k}$

$J^2(e_{2k+1}) = -J(e_{2k}) = -e_{2k+1}$

So any vector of E_J can be written as :

$u = \sum_{k\in I} u_k e_k = \sum u_{2k} e_{2k} + \sum u_{2k+1} e_{2k+1} = \sum u_{2k} e_{2k} - \sum u_{2k+1} J(e_{2k}) = \sum (u_{2k} - i u_{2k+1}) e_{2k} = \sum (-i u_{2k} + u_{2k+1}) e_{2k+1}$

A basis of the complex structure is then either e_{2k} or e_{2k+1} ■

Remark : this theorem can be extended to the case (of scarce usage !) of uncountable dimensional vector spaces, but this would involve some hypothesis about the set theory which are not always assumed.

The complex dimension of the complex vector space is half the real dimension of E if E is finite dimensional, equal to the dimension of E if E has a countable infinite dimension.

Contrary to the complexification it is not always possible to extend a real linear map $f \in L(E;E)$ to a complex linear map. It must be complex linear : $f(iu) = if(u) \Leftrightarrow f \circ J(u) = J \circ f(u)$ so it must commute with J : $J \circ f = f \circ J$. If so then $f \in L(E_{\mathbb{C}}; E_{\mathbb{C}})$ but it is not represented by the same matrix in the complex basis.

6.5.3 Real linear and complex linear maps

Real linear maps

1. Let E,F be two complex linear maps. A map $f : E \to F$ is **real linear** if :

$\forall u, v \in E, \forall k \in \mathbb{R} : f(u+v) = f(u) + f(v); f(ku) = kf(u)$

A real linear map (or R-linear map) is then a complex-linear maps (that is a linear map according to our definition) iff :

$\forall u \in E : f(iu) = if(u)$

Notice that these properties do not depend on the choice of a real structure on E or F.

2. If E is a real vector space, F a complex vector space, a real linear map : $f : E \to F$ can be uniquely extended to a linear map : $f_{\mathbb{C}} : E_{\mathbb{C}} \to F$ where $E_{\mathbb{C}}$ is the complexification of E. Define : $f_{\mathbb{C}}(u+iv) = f(u) + if(v)$

Cauchy identities

A complex linear map f between complex vector spaces endowed with real structures, must meet some specific identities, which are called (in the homolorphic map context) the Cauchy identities.

Theorem 314 *A linear map $f : E \to F$ between two complex vector spaces endowed with real structures can be written : $f(u) = P_x(\operatorname{Re} u) + P_y(\operatorname{Im} u) + i(Q_x(\operatorname{Re} u) + Q_y(\operatorname{Im} u))$ where P_x, P_y, Q_x, Q_y are real linear maps between the real kernels $E_{\mathbb{R}}, F_{\mathbb{R}}$ which satifsfy the identities : $P_y = -Q_x; Q_y = P_x$*

Proof. Let σ, σ' be the real structures on E,F

Using the sums : $E = E_\mathbb{R} \oplus iE_\mathbb{R}, F = F_\mathbb{R} \oplus iF_\mathbb{R}$ one can write for any vector u of E :

$\operatorname{Re} u = \frac{1}{2}(u + \sigma(u))$

$\operatorname{Im} u = \frac{1}{2i}(u - \sigma(u))$

$f(\operatorname{Re} u + i\operatorname{Im} u) = f(\operatorname{Re} u) + if(\operatorname{Im} u) = \operatorname{Re} f(\operatorname{Re} u) + i\operatorname{Im} f(\operatorname{Re} u) + i\operatorname{Re} f(\operatorname{Im} u) - \operatorname{Im} f(\operatorname{Im} u)$

$P_x(\operatorname{Re} u) = \operatorname{Re} f(\operatorname{Re} u) = \frac{1}{2}(f(\operatorname{Re} u) + \sigma'(f(\operatorname{Re} u)))$

$Q_x(\operatorname{Re} u) = \operatorname{Im} f(\operatorname{Re} u) = \frac{1}{2i}(f(\operatorname{Re} u) - \sigma'(f(\operatorname{Re} u)))$

$P_y(\operatorname{Im} u) = -\operatorname{Im} f(\operatorname{Im} u) = \frac{i}{2}(f(\operatorname{Im} u) - \sigma' f(\operatorname{Im} u))$

$Q_y(\operatorname{Im} u) = \operatorname{Re} f(\operatorname{Im} u) = \frac{1}{2}(f(\operatorname{Im} u) + \sigma' f(\operatorname{Im} u))$

So : $f(\operatorname{Re} u + i\operatorname{Im} u) = P_x(\operatorname{Re} u) + P_y(\operatorname{Im} u) + i(Q_x(\operatorname{Re} u) + Q_y(\operatorname{Im} u))$

As f is complex linear :

$f(i(\operatorname{Re} u + i\operatorname{Im} u)) = f(-\operatorname{Im} u + i\operatorname{Re} u) = if(\operatorname{Re} u + i\operatorname{Im} u)$

which gives the identities :

$f(-\operatorname{Im} u + i\operatorname{Re} u) = P_x(-\operatorname{Im} u) + P_y(\operatorname{Re} u) + i(Q_x(-\operatorname{Im} u) + Q_y(\operatorname{Re} u)$

$if(\operatorname{Re} u + i\operatorname{Im} u) = iP_x(\operatorname{Re} u) + iP_y(\operatorname{Im} u) - Q_x(\operatorname{Re} u) - Q_y(\operatorname{Im} u)$

$P_x(-\operatorname{Im} u) + P_y(\operatorname{Re} u) = -Q_x(\operatorname{Re} u) - Q_y(\operatorname{Im} u)$

$Q_x(-\operatorname{Im} u) + Q_y(\operatorname{Re} u) = P_x(\operatorname{Re} u) + P_y(\operatorname{Im} u)$

$P_y(\operatorname{Re} u) = -Q_x(\operatorname{Re} u)$

$Q_y(\operatorname{Re} u) = P_x(\operatorname{Re} u)$

$P_x(-\operatorname{Im} u) = -Q_y(\operatorname{Im} u)$

$Q_x(-\operatorname{Im} u) = P_y(\operatorname{Im} u)$ ∎

f can then be written : $f(\operatorname{Re} u + i\operatorname{Im} u) = (P_x - iP_y)(\operatorname{Re} u) + (P_y + iP_x)(\operatorname{Im} u)$

Conjugate of a map

Definition 315 *The **conjugate** of a linear map $f : E \to F$ between two complex vector spaces endowed with real structures σ, σ' is the map : $\overline{f} = \sigma' \circ f \circ \sigma$*

so $\overline{f}(u) = \overline{f(\overline{u})}$. Indeed the two conjugations are necessary to ensure that $\overline{f}$ is C-linear.

With the previous notations : $\overline{P}_x = P_x, \overline{P}_y = -P_y$

Real maps

Definition 316 *A linear map $f : E \to F$ between two complex vector spaces endowed with real structures is **real** if it maps a real vector of E to a real vector of F.*

$\operatorname{Im} u = 0 \Rightarrow f(\operatorname{Re} u) = (P_x - iP_y)(\operatorname{Re} u) = P_x(\operatorname{Re} u) \Rightarrow P_y = Q_x = 0$

Then $f = \overline{f}$

But conversely a map which is equal to its conjugate is not necessarily real.

Definition 317 *A multilinear form $f \in L^r(E; \mathbb{C})$ on a complex vector space E, endowed with a real structure σ is said to be **real valued** if its value is real whenever it acts on real vectors.*

A real vector is such that $\sigma(u) = u$ so $f(\sigma u_1, ..., \sigma u_r) \in \mathbb{R}$

Theorem 318 *An antilinear map f on a complex vector space E, endowed with a real structure σ can be uniquely decomposed into two real linear forms.*

Proof. Define the real linear forms :
$g(u) = \frac{1}{2}\left(f(u) + \overline{f(\sigma(u))}\right)$
$h(u) = \frac{1}{2i}\left(f(u) - \overline{f(\sigma(u))}\right)$
$f(u) = g(u) + ih(u)$ ∎
Similarly :

Theorem 319 *Any sesquilinear form γ on a complex vector space E endowed with a real structure σ can be uniquely defined by a C-bilinear form on E. A hermitian sesquilinear form γ is defined by a C-bilinear form g on E such that : $g(\sigma u, \sigma v) = \overline{g(v, u)}$*

Proof. If g is a C-bilinear form on E then : $\gamma(u, v) = g(\sigma u, v)$ defines a sesquilinear form

If g is a C-bilinear form on E such that : $\forall u, v \in E : g(\sigma u, v) = \overline{g(\sigma v, u)}$ then $\gamma(u, v) = g(\sigma u, v)$ defines a hermitian sesquilinear form. In a basis with $\sigma(ie_\alpha) = -ie_\alpha$ g must have components : $g_{\alpha\beta} = \overline{g_{\beta\alpha}}$

$g(\sigma u, v) = \overline{g(\sigma v, u)} \Leftrightarrow g(\sigma u, \sigma v) = \overline{g(\sigma^2 v, u)} = \overline{g(v, u)} \Leftrightarrow \overline{g(\sigma u, \sigma v)} = \overline{g}(u, v) = g(v, u)$

And conversely : $\gamma(\sigma u, v) = g(u, v)$ defines a C-bilinear form on E ∎

This definition is independant of any basis, and always valid. But the expression in components needs attention.

The usual antilinear map σ is expressed in a basis $(e_\alpha)_{\alpha \in A}$ by : $\sigma(e_\alpha) = e_\alpha, \sigma(ie_\alpha) = -ie_\alpha$. In matrix form it reads : $\sigma(u) = \overline{[u]}$ and $g(\sigma u, \sigma v) = \overline{g(v, u)} \Leftrightarrow \overline{[u]}^t [g] \overline{[v]} = \overline{[v]^t [g] [u]} \Leftrightarrow [u]^t \overline{[g]} [v] = [v]^t [g] [u] = [u]^t [g]^t [v]$

So the condition on g in the basis $(e_\alpha)_{\alpha \in A}$ reads : $[g] = [g]^*$

In a change of basis : $e_\alpha \to f_\alpha = \sum_\beta C_\alpha^\beta e_\alpha$ g has the new matrix : $[G] = [C]^t [g] [C]$ and vectors : $u = \sum_\alpha u^\alpha e_\alpha = \sum_\alpha U^\alpha f_\alpha$ with $[U] = [C]^{-1} [u]$.

$\sigma(u) = \sigma(\sum_\alpha U^\alpha f_\alpha) = \sum_\alpha \overline{U}^\alpha \sigma(f_\alpha) = \sum_\alpha \overline{U}^\alpha \sum_\beta \overline{C}_\alpha^\beta e_\beta = \sum_\alpha W^\alpha f_\alpha = \sum_\alpha W^\alpha \sum_\beta C_\alpha^\beta e_\beta \Leftrightarrow \overline{[C]}\, \overline{[U]} = [C] [W]$

So the components of $\sigma(u)$ in the new basis are : $[W] = [C]^{-1} \overline{[C][U]}$
One can check that $g(\sigma u, \sigma v) = \overline{g(v, u)}$

$g(\sigma u, \sigma v) = \left([C]^{-1} \overline{[C][U]}\right)^t [G] [C]^{-1} \overline{[C][U]} = [U]^* [C]^* [C^{-1}]^t [G] [C]^{-1} \overline{[C][V]} = [U]^* [C]^* [g] \overline{[C][V]} = [u]^* [g] \overline{[v]} = \overline{g(v, u)}$
but usually $[G] = [G]^*$ does not hold any longer.

Theorem 320 *A non degenerate scalar product g on a real vector space E can be extended to a hermitian, sesquilinear form γ on the complexified $E_{\mathbb{C}}$.*

Proof. On the complexified $E_{\mathbb{C}} = E \oplus iE$ we define the hermitian, sesquilinear form γ, prolongation of g by :

For any u,v $\in$ E :

$\gamma(u, v) = g(u, v)$

$\gamma(iu, v) = -i\gamma(u, v) = -ig(u, v)$

$\gamma(u, iv) = i\gamma(u, v) = ig(u, v)$

$\gamma(iu, iv) = g(u, v)$

$\gamma(u + iv, u' + iv') = g(u, u') + g(v, v') + i(g(u, v') - g(v, u')) = \overline{\gamma(u' + iv', u + iv)}$

If $(e_i)_{i \in I}$ is an orthonormal basis of E : $g(e_i, e_j) = \eta_{ij} = \pm 1$ then $(e_p)_{p \in I}$ is a basis of $E_{\mathbb{C}}$ and it is orthonormal :

$\gamma(e_p, e_q) = \eta_{pq}$

So the matrix of γ in this basis has a non null determinant and γ is not degenerate. It has the same signature as g, but it is always possible to choose a basis such that $[\gamma] = I_n$. ■

6.6 Affine Spaces

Affine spaces are the usual structures of elementary geometry. However their precise definition requires attention.

6.6.1 Definitions

Definition 321 *An* **affine space** $\left(E, \overrightarrow{E}\right)$ *is a set E with an underlying vector space $\overrightarrow{E}$ over a field K and a map $: \rightarrow : E \times E \rightarrow \overrightarrow{E}$ such that :*

i) $\forall A, B, C \in E : \overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CA} = \overrightarrow{0}$

ii) $\forall A \in E$ fixed the map $\tau_A : \overrightarrow{E} \rightarrow E :: \overrightarrow{AB} = \overrightarrow{u}$ is bijective

Definition 322 *The dimension of an affine space $\left(E, \overrightarrow{E}\right)$ is the dimension of $\overrightarrow{E}$.*

i) $\Rightarrow \forall A, B \in E : \overrightarrow{AB} = -\overrightarrow{BA}$ and $\overrightarrow{AA} = \overrightarrow{0}$

ii) $\Rightarrow \forall \overrightarrow{u} \in \overrightarrow{E}$ there is a unique $B \in E : \overrightarrow{AB} = \overrightarrow{u}$

On an affine space one can define the sum of points : $E \times E \rightarrow E :: \overrightarrow{OA} + \overrightarrow{OB} = \overrightarrow{OC}$ .The result does not depend on the choice of O.

We will usually denote : $\overrightarrow{AB} = \overrightarrow{u} \Leftrightarrow B = A + \overrightarrow{u} \Leftrightarrow B - A = \overrightarrow{u}$

An affine space is fully defined with a point O, and a vector space $\overrightarrow{E}$:

Define : $E = \left\{ A = (O, \overrightarrow{u}), \overrightarrow{u} \in \overrightarrow{E} \right\}, \overrightarrow{(O, \overrightarrow{u})(O, \overrightarrow{v})} = \overrightarrow{v} - \overrightarrow{u}$

So a vector space can be endowed with the structure of an affine space by taking O=$\overrightarrow{0}$.

Frame

Definition 323 *A **frame** in an affine space $\left(E, \overrightarrow{E}\right)$ is a pair $\left(O, (\overrightarrow{e}_i)_{i \in I}\right)$ of a point $O \in E$ and a basis $(\overrightarrow{e}_i)_{i \in I}$ of $\overrightarrow{E}$. The **coordinates** of a point M of E are the components of the vector $\overrightarrow{OM}$ with respect to $(\overrightarrow{e}_i)_{i \in I}$*

If I is infinite only a finite set of coordinates is non zero.

An affine space $\left(E, \overrightarrow{E}\right)$ is real if $\overrightarrow{E}$ is real (the coordinates are real), complex if $\overrightarrow{E}$ is complex (the coordinates are complex).

Affine subspace

Definition 324 *An **affine subspace** F of E is a pair $\left(A, \overrightarrow{F}\right)$ of a point A of E and a vector subspace $\overrightarrow{F} \subset \overrightarrow{E}$ with the condition :*
$$\forall M \in F : \overrightarrow{AM} \in \overrightarrow{F}$$

Thus $A \in F$

The dimension of F is the dimension of $\overrightarrow{F}$

Definition 325 *A **line** is a 1-dimensional affine subspace.*

Definition 326 *A **hyperplane** passing through A is the affine subspace complementary of a line passing through A.*

If E is finite dimensional an hyperplane is an affine subspace of dimension n-1.

If $K = \mathbb{R}, \mathbb{C}$ the **segment** AB between two points A≠B is the set :
$$AB = \left\{ M \in E : \exists t \in [0, 1], t\overrightarrow{AM} + (1 - t)\overrightarrow{BM} = 0 \right\}$$

Theorem 327 *The intersection of a family finite or infinite of affine subspaces is an affine subspace. Conversely given any subset F of an affine space the affine subspace generated by F is the intersection of all the affine subspaces which contains F.*

Definition 328 *Two affine subspaces are said to be **parallel** if they share the same underlying vector subspace $\overrightarrow{F}$:*

$$\left(A, \overrightarrow{F}\right) // \left(B, \overrightarrow{G}\right) \Leftrightarrow \overrightarrow{F} = \overrightarrow{G}$$

Product of affine spaces

1. If $\overrightarrow{E}, \overrightarrow{F}$ are vector spaces over the same field, $\overrightarrow{E} \times \overrightarrow{F}$ can be identified with $\overrightarrow{E} \oplus \overrightarrow{F}$. Take any point O and define the affine space : $\left(O, \overrightarrow{E} \oplus \overrightarrow{F}\right)$. It can be identified with the set product of the affine spaces : $\left(O, \overrightarrow{E}\right) \times \left(O, \overrightarrow{F}\right)$.

2. A real affine space $\left(E, \overrightarrow{E}\right)$ becomes a complex affine space $\left(E, \overrightarrow{E}_{\mathbb{C}}\right)$ with the complexification $\overrightarrow{E}_{\mathbb{C}} = \overrightarrow{E} \oplus i\overrightarrow{E}$.

$\left(E, \overrightarrow{E}_{\mathbb{C}}\right)$ can be dentified with the product of real affine space $\left(O, \overrightarrow{E}\right) \times \left(O, i\overrightarrow{E}\right)$.

3. Conversely a complex affine space $\left(E, \overrightarrow{E}\right)$ endowed with a real structure can be identified with the product of two real affine space $\left(O, \overrightarrow{E}_{\mathbb{R}}\right) \times \left(O, i\overrightarrow{E}_{\mathbb{R}}\right)$. The "complex plane" is just the affine space $\mathbb{C} \simeq \mathbb{R} \times i\mathbb{R}$

6.6.2 Affine transformations

Definition 329 *The **translation** by the vector $\overrightarrow{u} \in \overrightarrow{E}$ on of an affine space $\left(E, \overrightarrow{E}\right)$ is the map $\tau : E \to E :: \tau(A) = B :: \overrightarrow{AB} = \overrightarrow{u}$.*

Definition 330 *An **affine map** $f : E \to E$ on an affine space $\left(E, \overrightarrow{E}\right)$ is such that there is a map $\overrightarrow{f} \in L\left(\overrightarrow{E}; \overrightarrow{E}\right)$ and :*

$$\forall M, P \in E : M' = f(M), P' = f(P) : \overrightarrow{M'P'} = \overrightarrow{f}\left(\overrightarrow{MP}\right)$$

If is fully defined by a triple $\left(O, \overrightarrow{a}, \overrightarrow{f}\right), O \in E, \overrightarrow{a} \in \overrightarrow{F}, \overrightarrow{f} \in L\left(\overrightarrow{E}; \overrightarrow{E}\right)$: then

$$\overrightarrow{Of(M)} = \overrightarrow{a} + \overrightarrow{f}\left(\overrightarrow{OM}\right) \text{ so } A = f(O) \text{ with } \overrightarrow{OA} = \overrightarrow{a}$$

With another point O', the vector $\overrightarrow{a'} = \overrightarrow{O'f(O')}$ defines the same map:

Proof. $\overrightarrow{O'f'(M)} = \overrightarrow{a}' + \overrightarrow{f}\left(\overrightarrow{O'M}\right) = \overrightarrow{O'O} + \overrightarrow{Of(O')} + \overrightarrow{f}\left(\overrightarrow{O'O}\right) + \overrightarrow{f}\left(\overrightarrow{OM}\right) = \overrightarrow{O'O} + \overrightarrow{Of'(M)}$

$\overrightarrow{Of(O')} = \overrightarrow{a} + \overrightarrow{f}\left(\overrightarrow{OO'}\right)$

$\overrightarrow{Of'(M)} = \overrightarrow{a} + \overrightarrow{f}\left(\overrightarrow{OO'}\right) + \overrightarrow{f}\left(\overrightarrow{O'O}\right) + \overrightarrow{f}\left(\overrightarrow{OM}\right) = \overrightarrow{a} + \overrightarrow{f}\left(\overrightarrow{OM}\right) = \overrightarrow{Of(M)}$ ∎

It can be generalized to an affine map between affine spaces : $f : E \to F$:

take $(O, O', \overrightarrow{a'}, \overrightarrow{f}) \in E \times F \times \overrightarrow{F} \times L\left(\overrightarrow{E}; \overrightarrow{F}\right)$: then

$$\overrightarrow{O'f(M)} = \overrightarrow{a'} + \overrightarrow{f}\left(\overrightarrow{OM}\right) \Rightarrow \overrightarrow{O'f(O)} = \overrightarrow{a'}$$

If E is finite dimensional $\overrightarrow{f}$ is defined by a matrix $[F]$ in a basis and the coordinates of the image f(M) are given by the affine relation : $[y] = A + [F][x]$ with $\overrightarrow{OA} = \sum_{i \in I} a_i \overrightarrow{e}_i, \overrightarrow{OM} = \sum_{i \in I} x_i \overrightarrow{e}_i, \overrightarrow{Of(M)} = \sum_{i \in I} y_i \overrightarrow{e}_i$

Theorem 331 *(Berge p.144) A hyperplane in an affine space E over K is defined by f(x)=0 where $f : E \rightarrow K$ is an affine, non constant, map.*

f is not unique.

Theorem 332 *Affine maps are the morphisms of the affine spaces over the same field K, which is a category*

Theorem 333 *The set of affine transformations over an affine space $\left(E, \overrightarrow{E}\right)$ is a group with the composition law :*
$$\left(O, \overrightarrow{a}_1, \overrightarrow{f}_1\right) \circ (O, \overrightarrow{a}_2, \overrightarrow{f}_2) = (O, \overrightarrow{a}_1 + \overrightarrow{f}_1(\overrightarrow{a}_2), \overrightarrow{f}_1 \circ \overrightarrow{f}_2);$$
$$(O, \overrightarrow{a}, \overrightarrow{f})^{-1} = (O, -\overrightarrow{f}^{-1}(\overrightarrow{a}), \overrightarrow{f}^{-1})$$

6.6.3   Convexity

Barycenter

Definition 334 *A set of points $(M_i)_{i \in I}$ of an affine space E is said to be **independant** if all the vectors $\left(\overrightarrow{M_i M_j}\right)_{i,j \in I}$ are linearly independant.*

If E has the dimension n at most n+1 points can be independant.

Definition 335 *A weighted family in an affine space $\left(E, \overrightarrow{E}\right)$ over a field K is a family $(M_i, w_i)_{i \in I}$ where $M_i \in E$ and $w_i \in K$*
*The **barycenter** of a weighted family is the point G such that for each finite subfamily J of I : $\sum_{i \in J} m_i \overrightarrow{GM}_i = 0$.*

One writes : $G = \sum_{i \in J} m_i M_i$
In any coordinates : $(x_G)_i = \sum_{j \in I} m_j (x_{Mj})_i$

Convex subsets
Convexity is a purely geometric property. However in many cases it provides a "proxy" for topological concepts.

Definition 336 *A subset A of an affine space $\left(E, \overrightarrow{E}\right)$ is **convex** iff the barycenter of any weighted family $(M_i, 1)_{i \in I}$ where $M_i \in A$ belongs to A*

Theorem 337 *A subset A of an affine space $\left(E, \overrightarrow{E}\right)$ over $\mathbb{R}$ or $\mathbb{C}$ is convex iff*
$$\forall t \in [0,1], \forall M, P \in A, Q : t\overrightarrow{QM} + (1-t)\overrightarrow{QP} = 0, Q \in A$$

that is if any point in the segment joining M and P is in A.
Thus in $\mathbb{R}$ convex sets are closed intervals $[a, b]$

Theorem 338 *In an affine space* $\left(E, \overrightarrow{E}\right)$:

the empty set $\varnothing$ and E are convex.
the intersection of any collection of convex sets is convex.
the union of a non-decreasing sequence of convex subsets is a convex set.
if A_1, A_2 are convex then $A_1 + A_2$ is convex

Definition 339 *The **convex hull** of a subset A of an affine space $\left(E, \overrightarrow{E}\right)$ is the intersection of all the convex sets which contains A. It is the smallest convex set which contains A.*

Definition 340 *A convex subset C of a real affine space $\left(E, \overrightarrow{E}\right)$ is*

*i) A **cone** if for every x in C and $0 \leq \lambda \leq 1, \lambda x$ is in C.*
*ii) **Balanced** if for all x in C, $|\lambda| \leq 1, \lambda x$ is in C*
*iii) **Absorbent** or absorbing if the union of tC over all t > 0 is all of E, or equivalently for every x in E, tx is in C for some t > 0. The set C can be scaled out to absorb every point in the space.*
*iv) **Absolutely convex** if it is both balanced and convex. A set C is absolutely convex iff :*
$$\forall \lambda, \mu \in K, |\lambda| + |\mu| \leq 1, \forall M \in C, \forall O : \lambda \overrightarrow{OM} + \mu \overrightarrow{OM} \in C$$

There is a separation theorem which does not require any topological structure (but uses the Zorn lemna).

Theorem 341 *Kakutani (Berge p.162): If X,Y are two disjunct convex subset of an affine space E, there are two convex subsets X',Y' such that : $X \subset X', Y \subset Y', X' \cap Y' = \varnothing, X' \cup Y' = E$*

Definition 342 *A point a is an **extreme point** of a convex subset C of a real affine space if it does not lie in any open segment of C*

Meaning : $\forall M, P \in C, \forall t \in]0, 1[: tM + (1-t)P \neq a$

Convex function

Definition 343 *A real valued function $f : A \to \mathbb{R}$ defined on a convex set A of a real affine space $\left(E, \overrightarrow{E}\right)$ is **convex** if, for any two points M, P in A :$\forall t \in [0, 1] : f(Q) \leq tf(M) + (1-t)f(P)$ with Q such that : $t\overrightarrow{QM} + (1-t)\overrightarrow{QP} = 0$ It is stricly convex if $\forall t \in]0, 1[: f(Q) < tf(M) + (1-t)f(P)$*

Definition 344 *A function f is said to be (strictly) **concave** if -f is (strictly) convex.*

Theorem 345 *If g is an affine map : $g : A \to A$ and f is convex, then $f \circ g$ is convex*

6.6.4 Homology

Homology is a branch of abstract algebra. We will limit here to the definitions and results which are related to simplices, which can be seen as solids bounded by flat faces and straight edges. Simplices appear often in practical optimization problems : whenever one has to find the extremum of a linear function under linear constraints (what is called a linear program) the solution is on the simplex delimited by the constraints.

Definitions and results can be found in Nakahara p.110, Gamelin p.171

Simplex
(plural simplices)
1. Definition:

Definition 346 *A **k-simplex** denoted $\langle A_0, ...A_k \rangle$ where $(A_i)_{i=0}^k$ are k+1 independent points of a n dimensional real affine space $\left(E, \overrightarrow{E} \right)$, is the convex subset*
: $\langle A_0, ...A_k \rangle = \{ P \in E : P = \sum_{i=0}^k t_i A_i ; 0 \leq t_i \leq 1, \sum_{i=0}^k t_i = 1 \}$
*A **vertex** (plural vertices) is a 0-simplex (a point)*
*An **edge** is a 1-simplex (the segment joining 2 points)*
*A **polygon** is a 2-simplex in a 3 dimensional affine space*
*A **polyhedron** is a 3-simplex in a 3 dimensional affine space (the solid delimited by 4 points)*
*A **p-face** is a p-simplex issued from a k-simplex.*

So a k-simplex is a convex subset of a k dimensional affine subspace delimited by straight lines.

A regular simplex is a simplex which is symmetric for some group of affine transformations.

The standard simplex is the n-1-simplex in $\mathbb{R}^n$ delimited by the points of coordinates $A_i = (0, ..0, 1, 0, ...0)$

Remark : the definitions vary greatly, but these above are the most common and easily understood. The term simplex is sometimes replaced by polytope.

2. Orientation of a k-simplex:

Let be a path connecting any two vertices A_i, A_j of a simplex. This path can be oriented in two ways (one goes from A_i to A_j or from A_j to A_i). So for any path connecting all the vertices, there are only two possible consistent orientations given by the parity of the permutation $(A_{i_0}, A_{i_1}, ..., A_{i_k})$ of $(A_0, A_1, ...A_k)$. So a k-simplex can be oriented.

3. Simplicial complex:

Let be $(A_i)_{i \in I}$ a family of points in E. For any finite subfamily J one can define the simplex delimited by the points $(A_i)_{i \in J}$ denoted $\langle A_i \rangle_{i \in J} = C_J$. .The set $C = \cup_J C_J$ is a **simplicial complex** if : $\forall J, J' : C_J \cap C_{J'} \subset C$ or is empty

The dimension m of the simplicial complex is the maximum of the dimension of its simplices.

The **Euler characteristic** of a n dimensional simplicial complex is : $\chi(C) = \sum_{r=0}^{n} (-1)^r I_r$ where I_r is the number of r-simplices in C (non oriented). It is a generalization of the Euler Number in 3 dimensions :

Number of vertices - Number of edges + Number of 2-faces = Euler Number

r-chains

It is intuitive that, given a simplicial complex, one can build many different simplices by adding or removing vertices. This is formalized in the concept of chain and homology group, which are the basic foundations of algebraic topology (the study of "shapes" of objects in any dimension).

1. Definition:

Let C a simplicial complex, whose elements are simplices, and $C_r(C)$ its subset comprised of all r-simplices. $C_r(C)$ is a finite set with I_r different non oriented elements.

A **r-chain** is a formal finite linear combination of r-simplices belonging to the same simplicial complex. The set of all r-chains of the simplicial complex C is denoted $G_r(C)$:

$G_r(C) = \left\{ \sum_{i=1}^{I_r} z_i S_i, S_i \in C_r(C), z_i \in \mathbb{Z} \right\}, i = $ index running over all the elements of $C_r(C)$

Notice that the coefficients $z_i \in \mathbb{Z}$.

2. Group structure:

$G_r(C)$ is an abelian group with the following operations :

$\sum_{i=1}^{I_r} z_i S_i + \sum_{i=1}^{I_r} z_i' S_i = \sum_{i=1}^{I_r} (z_i + z_i') S_i$

$0 = \sum_{i=1}^{I_r} 0 S_i$

$-S_i = $ the same r-simplex with the opposite orientation

The group $G(C) = \oplus_r G_r(C)$

3. Border:

Any r-simplex of the complex can be defined from r+1 independant points. If one point of the simplex is removed we get a r-1-simplex which still belongs to the complex. The **border** of the simplex $\langle A_0, A_1, ... A_r \rangle$ is the r-1-chain :

$\partial \langle A_0, A_1, ... A_r \rangle = \sum_{k=0}^{r} (-1)^k \left\langle A_0, A_1, ., \hat{A}_k, ... A_r \right\rangle$ where the point A_k has been removed

Conventionnaly : $\partial \langle A_0 \rangle = 0$

The operator ∂ is a morphism $\partial \in \hom(G_r(C), G_{r-1}(C))$ and there is the exact sequence :

$0 \rightarrow G_n(C) \xrightarrow{\partial} G_{n-1}(C) \xrightarrow{\partial} .... G_0(C) \xrightarrow{\partial} 0$

3. Cycle:

A simplex such that $\partial S = 0$ is a **r-cycle**. The set $Z_r(C) = \ker(\partial)$ is the r-cycle subgroup of $G_r(C)$ and $Z_0(C) = G_0(C)$

Conversely if there is $A \in G_{r+1}(C)$ such that $B = \partial A \in G_r(C)$ then B is called a **r-border**. The set of r-borders is a subgroup $B_r(C)$ of $G_r(C)$ and $B_n(C) = 0$

$B_r(C) \subset Z_r(C) \subset G_r(C)$

4. Homology group:

The **r-homology group** of C is the quotient set : $H_r(C) = Z_r(C)/B_r(C)$

The rth **Betti number** is $b_r(C) = \dim H_r(C)$

Euler-Poincaré theorem : $\chi(C) = \sum_{r=0}^{n}(-1)^r b_r(C)$

The situation is very similar to the exact $(\varpi = d\pi)$ and closed $(d\varpi = 0)$ forms on a manifold, and there are strong relations between the groups of homology and cohomology.

7 TENSORS

Tensors are mathematical objects defined over a space vector. As they are ubiquituous in mathematics, they deserve a full section. Many of the concepts presented here stand in vector bundles, due to the functorial nature of tensors constructs, so it is good to have a good grasp at these concepts in the simpler framework of vector space in order to get along with the more difficult cases of differential geometry.

7.1 Tensorial product of vector spaces

All definitions and theorems of this section can be found in Knapp Annex A.

7.1.1 Definition

Universal property

Definition 347 *The **tensorial product** $E \otimes F$ of two vector spaces on the same field K is defined by the following universal property : there is a map $\imath :$ $E \times F \to E \otimes F$ such that for any vector space S and bilinear map $f : E \times F \to S$, there is a unique linear map : $F : E \otimes F \to S$ such that $f = F \circ \imath$*

This definition can be seen as abstract, but it is in fact the most natural introduction of tensors. Let f be a bilinear map so :

$f(u,v) = f\left(\sum_i u_i e_i, \sum_j v_j f_j\right) = \sum_{i,j} u_i v_j f(e_i, f_j) = \sum_{ijk} F_{ijk} u_i v_j \varepsilon_k$

it is intuitive to extend the map by linearity to something like : $\sum_{ijk} F_{ijk} U_{ij} \varepsilon_k$ meaning that $U = u \otimes v$

This can be expressed in category parlance (Lane p.58). Let be $\mathfrak{V}$ the category of vector spaces, Set the category of sets, H the functor $\mathfrak{V} \mapsto Set$ which assigns to each vector space S the set of all bilinear maps to S $:L^2(V \times V'; S)$. The pair $(E \otimes F, \imath)$ is a universal morphism from $\mathfrak{V} \times \mathfrak{V}$ to H.

Definition is not proof of existence. So to prove that the tensorial product does exist the construct is the following :

1. Take the product $E \times F$ with the obvious structure of vector space.

2. Take the equivalence relation : $(x,0) \sim (0,y) \sim 0$ and 0 as identity element for addition

3. Define $E \otimes F = E \times F/\sim$

Example The set $K_p[x_1, ...x_n]$ of polynomials of degree p in n variables is a vector space over K.

$P_p \in K_p[x]$ reads : $P_p(x) = \sum_{r=0}^n a_r x^r = \sum_{r=0}^n a_r e_r$ with as basis the monomials : $e_r = x^r, r = 0..n$

Consider the bilinear map :

$f : K_p[x] \times K_q[x] \to K_{p+q}[x,y] :: f(P_p(x), P_q(y)) = P_p(x) \times P_q(y) = \sum_{r=0}^{p} \sum_{s=0}^{q} a_r b_s x^r y^s$

So there is a linear map : $F : K_p[x] \otimes K_q[x] \to K_{p+q}[x,y] :: f = F \circ \imath$

$\imath(e_r, e_s) = e_r \otimes e_s$

$\imath(P_p(x), P_q(y)) = \sum_{r=0}^{p} \sum_{s=0}^{q} a_r b_s e_r \otimes e_s$

$\sum_{r=0}^{p} \sum_{s=0}^{q} a_r b_s x^r y^s = \sum_{r=0}^{p} \sum_{s=0}^{q} a_r b_s e_r \otimes e_s$

So $e_r \otimes e_s = x^r y^s$

And one can write : $K_p[x] \otimes K_q[y] = K_{p+q}[x,y]$

7.1.2 Properties

Theorem 348 *The tensorial product $E \otimes F$ of two vector space on a field K is a vector space is a vector space on K whose vectors are called **tensors**.*

Definition 349 *The bilinear map : $\imath : E \times F \to E \otimes F :: \imath(u,v) = u \otimes v$ is the **tensor product of vectors***

with the properties :
$\forall T, U, T', U' \in E \otimes F, \forall a, b \in K$
$aT + bU \in E \otimes F$
$(aT + T') \otimes U = aT \otimes U + T' \otimes U$
$T \otimes (aU + U') = aT \otimes U + T \otimes U'$
$0 \otimes T = T \otimes 0 = 0 \in E \otimes F$
But if E=F *it is not commutative* : $u, v \in E, u \otimes v = v \otimes u \Leftrightarrow \exists k \in K : v = ku$

Theorem 350 *If $(e_i)_{i \in I}, (f_j)_{j \in J}$ are basis of E and F, $(e_i \otimes f_j)_{I \times J}$ is a basis of $E \otimes F$ called a tensorial basis.*

So tensors are linear combinations of $e_i \otimes f_j$. If E and F are finite dimensional with dimensions n,p then $E \otimes F$ is finite dimensional with dimensions nxp.

If $u = \sum_{i \in I} U_i e_i, v = \sum_{j \in J} V_j f_j : u \otimes v = \sum_{(i,j) \in I \times J} U_i V_j e_i \otimes f_j$

The components of the tensorial product are the sum of all combination of the components of the vectors

If $T \in E \otimes F : T = \sum_{(i,j) \in I \times J} T_{ij} e_i \otimes f_j$

A tensor which can be put in the form : $t \in E \otimes F : t = u \otimes v, u \in E, v \in F$ is said to be **decomposable**.

Warning ! all tensors are not decomposable : they are sum of decomposable tensors

Theorem 351 *The vector spaces $E \otimes F \simeq F \otimes E, E \otimes K \simeq E$ are canonically isomorphic and can be identified whenever $E \neq F$*

7.1.3 Tensor product of more than two vector spaces

Definition 352 *The **tensorial product** $E_1 \otimes E_2 ... \otimes E_r$ of the vector spaces $(E_i)_{i=1}^r$ on the same field K is defined by the following universal property : there is a multilinear map : $\imath : E_1 \times E_2 ... \times E_r \to E_1 \otimes E_2 ... \otimes E_r$ such that for any vector space S and multilinear map $f : E_1 \times E_2 ... \times E_r \to S$ there is a unique linear map : $F : E_1 \otimes E_2 ... \otimes E_r \to S$ such that $f = F \circ \imath$*

The **order** of a tensor is the number r of vectors spaces.

In components with : $u_k = \sum_{j \in I_k} U_{kj} e_{kj}$

$u_1 \otimes u_2 ... \otimes u_r = \sum_{(j_1, j_2, ... j_r) \in I_1 \times I_2 ... \times I_r} U_{1j_1} U_{2j_2} ... U_{rj_r} e_{1j_1} \otimes e_{2j_2} ... \otimes e_{rj_r}$

The multilinear map $:\imath : E_1 \times E_2 ... \times E_r \to E_1 \otimes E_2 ... \otimes E_r$ is the tensor product of vectors

As each tensor product $E_{i_1} \otimes E_{i_2} ... \otimes E_{i_k}$ is itself a vector space the **tensorial product of tensors** can be defined.

Theorem 353 *The tensorial product of tensors is associative, and distributes over direct sums, even infinite sums :*

$E \otimes (\oplus_I F_i) = \oplus_I (E \otimes F_i)$

In components :

$T = \sum_{(i_1, i_2, ... i_r) \in I_1 \times I_2 ... \times I_r} T_{i_1 i_2 .. i_r} e_{1i_1} \otimes e_{2i_2} ... \otimes e_{ri_r}$

$S = \sum_{(i_1, i_2, ... i_r) \in J_1 \times J_2 ... \times J_s} S_{j_1 j_2 .. i_r} f_{1i_1} \otimes f_{2j_2} ... \otimes f_{sj_s}$

$T \otimes S = \sum_{(i_1, i_2, ... i_r) \in I_1 \times I_2 ... \times I_r} \sum_{(i_1, i_2, ... i_r) \in J_1 \times J_2 ... \times J_s} T_{i_1 i_2 .. i_r} S_{j_1 j_2 .. i_r} e_{1i_1} \otimes e_{2i_2} ... \otimes e_{ri_r} \otimes f_{1i_1} \otimes f_{2j_2} ... \otimes f_{sj_s}$

The sets L(E;E'), L(F;F') of linear maps are vector spaces, so one can define the tensorial product $L(E; E') \otimes L(F; F')$:

$L(E; E') \otimes L(F; F') \in L(E \otimes F; E' \otimes F')$

and it has the property : $\forall f \in L(E; E'), g \in L(F; F'), \forall u \in E, v \in F :$ $(f \otimes g)(u \otimes v) = f(u) \otimes g(v)$

There is more on this topic in the following.

7.1.4 Tensorial algebra

Definition 354 *The **tensorial algebra**, denoted $T(E)$, of the vector space E on the field K is the direct sum $T(E) = \oplus_{n=0}^\infty (\otimes^n E)$ of the tensorial products $\otimes^n E = E \otimes E ... \otimes E$ where for n=0 $\otimes^0 E = K$*

Theorem 355 *The tensorial algebra of the vector space E on the field K is an algebra on the field K with the tensor product as internal operation and the unity element is 1$\in$ K.*

The elements of $\otimes^n E$ are homogeous tensors of order n. Their components in a basis $(e_i)_{i \in I}$ are such that :

$T = \sum_{(i_1 ... i_n)} t^{i_1 ... i_n} e_{i_1} \otimes ... \otimes e_{i_n}$ with the sum over all finite n-sets of indices $(i_1 ... i_n), i_k \in I$

Theorem 356 *The tensorial algebra $T(E)$, of the vector space E on the field K has the universal property : for any algebra A on the field K and linear map $l : E \to A$ there is a unique algebra morphism $L : T(E) \to A$ such that : $l = L \circ \jmath$ where $\jmath : E \to \otimes^1 E$*

Definition 357 *A derivation D over the algebra $T(E)$ is a map $D : T(E) \to T(E)$ such that :*
$$\forall u, v \in T(E) : D\left(u \otimes v\right) = D(u) \otimes v + u \otimes D(v)$$

Theorem 358 *The tensorial algebra $T(E)$ of the vector space E has the universal property that for any linear map $d : E \to T(E)$ there is a unique derivation $D : T(E) \to T(E)$ such that : $d = D \circ \jmath$ where $\jmath : E \to \otimes^1 E$*

7.1.5 Covariant and contravariant tensors

Definition 359 *Let be E a vector space and E* its algebraic dual*

*The tensors of the tensorial product of p copies of E are p **contravariant** tensors*

The tensors of the tensorial product of q copies of E are q **covariant** tensors*

The tensors of the tensorial product of p copies of E and q copies of E are mixed, p contravariant,q covariant tensors (or a **type (p,q)** tensor)*

The tensorial product is not commutative if E=F, so in a mixed product (p,q) the order between contravariant on one hand, covariant on the other hand, matters, but not the order between contravariant and covariant. So :
$$\overset{p}{\underset{q}{\otimes}}E = (\otimes E)^p \otimes (\otimes E^*)^q = (\otimes E^*)^q \otimes (\otimes E)^p$$

Notation 360 *$\otimes_q^p E$ is the vector space of type (p,q) tensors over E :*

Components of contravariant tensors are denoted as upper index: $a^{ij...m}$

Components of covariant tensors are denoted as lower index: $a_{ij...m}$.

Components of mixed tensors are denoted with upper and lower indices: $a^{ij...m}_{qr...t}$

The order of the upper indices (resp.lower indices) matters

Basis vectors e_i of E are denoted with lower index, and

Basis vectors e^i of E* are denoted with upper index.

Notice that a covariant tensor is a multilinear map acting on vectors the usual way :

If $T = \sum t_{ij} e^i \otimes e^j$ then T(u,v)=$\sum_{ij} t_{ij} u^i v^j \in K$

Similarly a contravariant tensor can be seen as a linear map acting on 1-forms :

If $T = \sum t^{ij} e_i \otimes e_j$ then T(λ,μ)=$\sum_{ij} t^{ij} \lambda_i \mu_j \in K$

And a mixed tensor is a map acting on vectors and giving vectors (see below)

Isomorphism L(E;E)$\simeq E \otimes E^*$

Theorem 361 *If the vector space E is finite dimensional, there is an isomorphism between L(E;E) and $E \otimes E^*$*

Proof. Define the bilinear map :

$\lambda : E \times E^* \to L(E;E) :: \lambda(u,\varpi)(v) = \varpi(u)v$

$\lambda \in L^2(E, E^*; L(E;E))$

From the universal property of tensor product :

$\imath : E \times E^* \to E \otimes E^*$

$\exists$ unique $\Lambda \in L(E \otimes E^*; L(E;E)) : \lambda = \Lambda \circ \imath$

$t \in E \otimes E^* \to f = \Lambda(t) \in L(E;E)$

Conversely :

$\forall f \in L(E;E), \exists f^* \in L(E^*; E^*) : f^*(\varpi) = \varpi \circ f$

$\exists f \otimes f^* \in L(E \otimes E^*; E \otimes E^*) :: (f \otimes f^*)(u \otimes \varpi) = f(u) \otimes f^*(\varpi) = f(u) \otimes (\varpi \circ f) \in E \otimes E^*$

Pick up any basis of E : $(e_i)_{i \in I}$ and its dual basis $(e^i)_{i \in I}$

Define : $T = \sum_{i,j} (f \otimes f^*)(e_i \otimes e^j) \in E \otimes E^*$

In components : $f(u) = \sum_{ij} g_i^j u^i e_j \to T(f) = \sum_{ij} g_i^j e^i \otimes e_j$ ■

Warning ! E must be finite dimensional

This isomorphism justifies the notation of matrix elements with upper indexes (rows, for the contravariant part) and lower indexes (columns, for the covariant part) : the matrix A$=\left[a_j^i\right]$ is the matrix of the linear map : f$\in L(E;E) ::$ $f(u) = \sum_{i,j} \left(a_j^i u^j\right) e_i$ which is identified with the mixed tensor in $E \otimes E^*$ acting on a vector of E.

Definition 362 *The **Kronecker tensor** is $\delta = \sum_{i=1}^n e^i \otimes e_i = \sum_{ij} \delta_j^i e^i \otimes e_j \in E \otimes E^*$*

It has the same components in any basis, and is isomorphic to the identity map $E \to E$

The trace operator

Theorem 363 *If E is a vector space on the field K there is a unique linear map called the trace $Tr : E^* \otimes E \to K$ such that :$Tr(\varpi \otimes u) = \varpi(u)$*

Proof. This is the consequence of the universal property :

For : $f : E^* \times E \to K :: f(\varpi, u) = \varpi(u)$

we have : $f = Tr \circ \imath \Leftrightarrow f(\varpi, u) = F(\varpi \otimes u) = \varpi(u)$ ■

So to any (1,1) tensor S is associated one scalar Tr(S) called the **trace** of the tensor, whose value does not depend on a basis. In components it reads :

$S = \sum_{i,j \in I} S_i^j e^i \otimes e_j \to Tr(S) = \sum_{i \in I} S_i^i$

If E is finite dimensional there is an isomorphism between L(E;E) and $E \otimes E^*$, and $E^* \otimes E \equiv E \otimes E^*$. So to any linear map $f \in L(E;E)$ is associated a scalar. In a basis it is $Tr(f) = \sum_{i \in I} f_{ii}$. This is the geometric (basis independant) definition of the **Trace** operator of an endomorphism.

Remark : this is an algebraic definition of the trace operator. This definition uses the algebraic dual E* which is replaced in analysis by the topological dual. So there is another definition for the Hilbert spaces, they are equivalent in finite dimension.

Theorem 364 *If E is a finite dimensional vector space and $f, g \in L(E; E)$ then $Tr(f \circ g) = Tr(g \circ f)$*

Proof. Check with a basis :

$f = \sum_{i \in I} f_i^j e^i \otimes e_j, g = \sum_{i \in I} g_i^j e^i \otimes e_j$

$f \circ g = \sum_{i,j,k \in I} f_k^j g_i^k e^i \otimes e_j$

$Tr(f \circ g) = \sum_{i,k \in I} f_k^i g_i^k = \sum_{i,k \in I} g_k^i f_i^k$ ■

Contraction of tensors

Over mixed tensors there is an additional operation, called **contraction**.

Let $T \in \otimes_q^p E$. One can take the trace of T over one covariant and one contravariant component of T (or similarly one contravariant component and one covariant component of T). The resulting tensor $\in \otimes_{q-1}^{p-1} E$. The result depends of the choice of the components which are to be contracted (but not of the basis).

Example :

Let $T = \sum_{ijk} a_{jk}^i e_i \otimes e^j \otimes e^k \in \overset{1}{\underset{2}{\otimes}} E$, the contracted tensor is $\sum_i \sum_k a_{ik}^i e_i \otimes e^k \in \overset{1}{\underset{1}{\otimes}} E$

$\sum_i \sum_k a_{ik}^i e_i \otimes e^k \neq \sum_i \sum_k a_{ki}^i e_i \otimes e^k \in \overset{1}{\underset{1}{\otimes}} E$

Einstein summation convention :

In the product of components of mixed tensors, whenever a index has the same value in a upper and in a lower position it is assumed that the formula is the sum of these components. This convention is widely used and most convenient for the contraction of tensors.

Examples :

$a_{jk}^i b_i^l = \sum_i a_{jk}^i b_i^l$

$a^i b_i = \sum_i a^i b_i$

So with this convention $a_{ik}^i = \sum_i a_{ik}^i$ is the contracted tensor

Change of basis

Let E a finite dimensional n vector space. So the dual E* is well defined and is n dimensional.

A basis $(e_i)_{i=1}^n$ of E and the dual basis $(e^i)_{i=1}^n$ of E*

In a change of basis : $f_i = \sum_{j=1}^n P_i^j e_j$ the components of tensors change according to the following rules :

$[P] = [Q]^{-1}$

- the contravariant components are multiplied by Q (as for vectors)
- the covariant components are multiplied by P (as for forms)

$T = \sum_{i_1...i_p} \sum_{j_1....j_q} t^{i_1...i_p}_{j_1...j_q} e_{i_1} \otimes e_{i_2}... \otimes e_{i_p} \otimes e^{j_1} \otimes ... \otimes e^{j_q}$
$\rightarrow$
$T = \sum_{i_1...i_p} \sum_{j_1....j_q} \widetilde{t}^{i_1...i_p}_{j_1...j_q} f_{i_1} \otimes f_{i_2}... \otimes f_{i_p} \otimes f^{j_1} \otimes ... \otimes f^{j_q}$

with :
$\widetilde{t}^{i_1...i_p}_{j_1...j_q} = \sum_{k_1...k_p} \sum_{l_1....l_q} t^{k_1...k_p}_{l_1...l_q} Q^{i_1}_{k_1}...Q^{i_p}_{k_p} P^{l_1}_{j_1}...P^{l_q}_{j_q}$

Bilinear forms

Let E a finite dimensional n vector space. So the dual E* is well defined and is n dimensional. Let $(e_i)_{i=1}^n$ be a basis of E with its the dual basis $(e^i)_{i=1}^n$ of E*.

1. Bilinear forms: $g : E \times E \rightarrow K$ can be seen as tensors : $G : E^* \otimes E^*$:
$g(u,v) = \sum_{ij} g_{ij} u^i u^j \rightarrow G = \sum_{ij} g_{ij} e^i \otimes e^j$

Indeed in a change of basis the components of the 2 covariant tensor $G = \sum_{ij} g_{ij} e^i \otimes e^j$ change as :

$G = \sum_{ij} \widetilde{g}_{ij} f^i \otimes f^j$ with $\widetilde{g}_{ij} = \sum_{kl} g_{kl} P_i^k P_j^l$ so $[\widetilde{g}] = [P]^t [g] [P]$ is transformed according to the rules for bilinear forms.

Similarly let be $[g]^{-1} = [g^{ij}]$ and $H = \sum_{ij} g^{ij} e_i \otimes e_j$. h is a 2 contravariant tensor $h \in \otimes^2 E$

2. Let E be a a n-dimensional vector space over $\mathbb{R}$ endowed with a bilinear symmetric form g, non degenerate (but not necessarily definite positive). Its matrix is $[g] = [g_{ij}]$ and $[g]^{-1} = [g^{ij}]$

By contraction with the 2 covariant tensor $G = \sum_{ij} g_{ij} e^i \otimes e^j$ one "lowers" a contravariant tensor :

$T = \sum_{i_1...i_p} \sum_{j_1....j_q} t^{i_1...i_p}_{j_1...j_q} e_{i_1} \otimes e_{i_2}... \otimes e_{i_p} \otimes e^{j_1} \otimes ... \otimes e^{j_q}$
$\rightarrow \widetilde{T} = \sum_{i_2...i_p} \sum_{j_1....j_{q+1}} \sum_{i_1} g_{j_{q+1}i_1} t^{i_1...i_p}_{j_1...j_q} e_{i_2}... \otimes e_{i_p} \otimes e^{j_1} \otimes ... \otimes e^{j_{q+1}}$

so $T \in \otimes^p_q \rightarrow \widetilde{T} \in \otimes^{p-1}_{q+1}$

This operation can be done on any (or all) contravariant components (it depends of the choice of the component) and the result does not depend of the basis.

Similarly by contraction with the 2 covariant tensor $H = \sum_{ij} g^{ij} e_i \otimes e_j$ one "lifts" a covariant tensor :

$T = \sum_{i_1...i_p} \sum_{j_1....j_q} t^{i_1...i_p}_{j_1...j_q} e_{i_1} \otimes e_{i_2}... \otimes e_{i_p} \otimes e^{j_1} \otimes ... \otimes e^{j_q}$
$\rightarrow \widetilde{T} = \sum_{i_1...i_{p+1}} \sum_{j_2....j_q} \sum_{j_1} g^{i_{p+1}j_1} t^{i_1...i_p}_{j_1...j_q} e_{i_1}... \otimes e_{i_{p+1}} \otimes e^{j_2} \otimes ... \otimes e^{j_q}$

so $T \in \otimes^p_q \rightarrow \widetilde{T} \in \otimes^{p+1}_{q-1}$

These operations are just the generalization of the isomorphism $E \simeq E^*$ using a bilinear form.

Derivation

The tensor product of any mixed tensor defines the algebra of tensors over a vector space E :

Notation 365 $\otimes E = \oplus_{r,s=0}^{\infty} (\otimes_s^r E)$ *is the algebra of all tensors over* E

Theorem 366 *The tensorial algebra* $\otimes E$ *of the vector space* E *on the field* K *is an algebra on the field* K *with the tensor product as internal operation and the unity element is* $1 \in K$.

Definition 367 *A* **derivation on the tensorial algebra** $\otimes E$ *is a linear map* $D : \otimes E \to \otimes E$ *such that :*
i) it preserves the tensor type : $\forall r, s, T \in \otimes_s^r E : DT \in \otimes_s^r E$
ii) it follows the Leibnitz rule for tensor product :
$\forall S, T \in \otimes E : D(S \otimes T) = D(S) \otimes T + S \otimes D(T)$
iii) it commutes with the trace operator.

So it will commute with the contraction of tensors.
A derivation on the tensorial algebra is a derivation as defined previously (see Algebras) with the i),iii) additional conditions.

Theorem 368 *The set of all derivations on* $\otimes E$ *is a vector space and a Lie algebra with the bracket :* $[D, D'] = D \circ D' - D' \circ D$.

Theorem 369 *(Kobayashi p.25) If* E *is a finite dimensional vector space, the Lie algebra of derivations on* $\otimes E$ *is isomorphic to the Lie algebra of endomorphisms on* E. *This isomorphism is given by assigning to each derivation its value on* E.

So given an endomorphism $f \in L(E; E)$ there is a unique derivation D on $\otimes E$ such that :
$\forall u \in E, \varpi \in E^* : Du = f(u), D(\varpi) = -f^*(\varpi)$ where f* is the dual of f and we have $\forall k \in K : D(k) = 0$

7.2 Algebras of symmetric and antisymmetric tensors

There are two ways to look at the set of symmetric (resp.antisymmetric) tensors :

- the geometric way : this is a vector subspace of tensors, and using their specificities one can define some additional operations, which fully come from the tensorial product. But the objects stay tensors.
- the algebraic way : as a symmetric tensor can be defined by a restricted set of components one can take the quotient of the vector subspace by the equivalence relations. One gets another set, with a structure of algebra, whose

objects are no longer tensors but classes of equivalence of tensors, upon which specific operationscan be defined.

The way which is taken depends upon the authors, and of course of their main topic of interest, but it is rarely explicited, and that brings much confusion on the subject. I will expose below the two ways, but in all the rest of this book I will clearly take the geometric way because it is by far the most convenient in geometry, with which we will have to deal.

We will use contravariant tensors, but everything is valid with covariant tensors as well (but not mixed tensors).

Notation 370 *For any finite set I of indices:*

$(i_1, i_2, ...i_n)$ *is any subset of n indexes chosen in I, two subsets deduced by permutation are considered distinct*

$\sum_{(i_1,i_2,...i_n)}$ *is the sum over all permutations of n indices in I*

$\{i_1, i_2, ...i_n\}$ *is any strictly ordered permutation of n indices in I: $i_1 < i_2 < .. < i_n$*

$\sum_{\{i_1,i_2,...i_n\}}$ *is the sum over all ordered permutations of n indices chosen in I*

$[i_1, i_2, ...i_n]$ *is any set of n indexes in I such that: $i_1 \leq i_2 \leq .. \leq i_n$*

$\sum_{[i_1,i_2,...i_n]}$ *is the sum over all distinct such sets of indices chosen in I*

We remind the notations:

$\mathfrak{S}(\mathfrak{n})$ is the symmetric group of permutation of n indexes

$\sigma(i_1, i_2, ...i_n) = (\sigma(i_1), \sigma(i_2), ...\sigma(i_n))$ is the image of the set $(i_1, i_2, ...i_n)$ by $\sigma \in \mathfrak{S}(\mathfrak{n})$

$\epsilon(\sigma)$ where $\sigma \in \mathfrak{S}(\mathfrak{n})$ is the signature of σ

Permutation is a set operation, without respect for the possible equality of some of the elements of the set. So $\{a, b, c\}$ and $\{b, a, c\}$ are two distinct permutations of the set even if it happens that a=b.

7.2.1 Algebra of symmetric tensors

Symmetric tensors

1. Symmetrizer :

Definition 371 *On a vector space E the **symmetrisation operator** or symmetrizer is the map :*

$s_r : E^r \to \overset{r}{\otimes}E :: s_r(u_1, .., u_r) = \sum_{\sigma \in \mathfrak{S}(r)} u_{\sigma(1)} \otimes .. \otimes u_{\sigma(r)}$

It is a multilinear symmetric map : $s_r \in L^r(E^r; \otimes^r E)$

$s_r((u_{\sigma(1)}, u_{\sigma(2)}, ..., u_{\sigma(r)})) = \sum_{\sigma' \in \mathfrak{s}_r} u_{\sigma'\sigma(1)} \otimes .. \otimes u_{\sigma'\sigma(r)} = \sum_{\theta \in \mathfrak{s}_r} u_{\theta(1)} \otimes .. \otimes u_{\theta(r)} = s_r(u_1, u_2, ..., u_r)$

So there is a unique linear map : $S_r : \overset{r}{\otimes}E \to \overset{r}{\otimes}E$: such that : $s_r = S_r \circ \imath$ with $\imath : E^r \to \overset{r}{\otimes}E$

$s_r(e_1, .., e_r) = S_r \circ \imath(e_1, ..., e_r) = S_r(e_1 \otimes .. \otimes e_r) = \sum_{\sigma \in \mathfrak{s}_r} e_{\sigma(1)} \otimes .. \otimes e_{\sigma(r)}$

108

For any tensor $T \in \overset{r}{\otimes} E$:

$S_r(T) = \sum_{(i_1...i_r)} t^{i_1...i_r} S_r(e_{i_1} \otimes ... \otimes e_{i_r}) = \sum_{(i_1...i_r)} t^{i_1...i_r} \sum_{\sigma \in \mathfrak{s}_r} e_{\sigma(1)} \otimes .. \otimes e_{\sigma(r)}$

2. Symmetric tensors:

Definition 372 *A symmetric r contravariant tensor is a tensor T such that* $S_r(T) = r!T$

In a basis a symmetric r contravariant tensor reads : $T = \sum_{(i_1...i_r)} t^{i_1...i_r} e_{i_1} \otimes ... \otimes e_{i_r}$, where $t^{i_1...i_r} = t^{\sigma(i_1...i_r)}$ with σ is any permutation of the set of r-indices.
Example :
$T = t^{111} e_1 \otimes e_1 \otimes e_1 + t^{112} e_1 \otimes e_1 \otimes e_2 + t^{121} e_1 \otimes e_2 \otimes e_1 + t^{122} e_1 \otimes e_2 \otimes e_2$
$+t^{211} e_2 \otimes e_1 \otimes e_1 + t^{212} e_2 \otimes e_1 \otimes e_2 + t^{221} e_2 \otimes e_2 \otimes e_1 + t^{222} e_2 \otimes e_2 \otimes e_2$
$S_3(T) = 6t^{111} e_1 \otimes e_1 \otimes e_1 + 6t^{222} e_2 \otimes e_2 \otimes e_2$
$+2\left(t^{112} + t^{121} + t^{211}\right)(e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_1 + e_2 \otimes e_1 \otimes e_1)$
$+2\left(t^{122} + t^{212} + t^{221}\right)(e_1 \otimes e_2 \otimes e_2 + e_2 \otimes e_1 \otimes e_2 + e_2 \otimes e_2 \otimes e_1)$
If the tensor is symmetric : $t^{112} = t^{121} = t^{211}, t^{122} = t^{212} = t^{221}$ and
$S_3(T) = 6\{t^{111} e_1 \otimes e_1 \otimes e_1 + t^{112}(e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_1 + e_2 \otimes e_1 \otimes e_1)$
$+t^{122}(e_1 \otimes e_2 \otimes e_2 + e_2 \otimes e_1 \otimes e_2 + e_2 \otimes e_2 \otimes e_1) + t^{222} e_2 \otimes e_2 \otimes e_2\}$

3. Space of symmetric tensors:

Notation 373 $\odot^r E$ *is the set of symmetric r-contravariant tensors on E*
$\odot_r E^*$ *is the set of symmetric r-covariant tensors on E*

Theorem 374 *The set of symmetric r-contravariant tensors $\odot^r E$ is a vector subspace of $\otimes^r E$.*

A symmetric tensor is uniquely defined by a set of components $t^{i_1...i_r}$ for all ordered indices $[i_1...i_r]$ with the rule :
$t^{\sigma(i_1...i_r)} = t^{i_1...i_r}$
If $(e_i)_{i \in I}$ is a basis of E, with I an ordered set, the set of ordered products
$e_{i_1} \otimes e_{i_2} \otimes ... \otimes e_{i_r}, i_1 \leq i_2.. \leq i_r \equiv (\otimes e_{i_1})^{j_1} \otimes (\otimes e_{i_2})^{j_2} \otimes ... (\otimes e_{i_k})^{j_k}, i_1 < i_2.. < i_k, \sum_{l=1}^{k} j_l = r$ is a basis of $\odot^r E$
If E is n-dimensional dim $\odot^r E = C_{n-1+r}^{n-1}$

4. Universal property:
For any r-linear symmetric map $f \in L^r(E; E')$:
$\forall u_i \in E, i = 1...r, \sigma \in \mathfrak{S}_r : f(u_1, u_2, ..., u_r) = f(u_{\sigma(1)}, u_{\sigma(2)}, ..., u_{\sigma(r)})$
There is a unique linear map : $F \in L\left(\overset{r}{\otimes} E; E'\right)$ such that : $f = F \circ \imath$
$F \circ s_r(u_1, u_2, ..., u_r) = \sum_{\sigma \in \mathfrak{s}_r} F\left(u_{\sigma(1)} \otimes .. \otimes u_{\sigma(r)}\right) = \sum_{\sigma \in \mathfrak{s}_r} F \circ i\left(u_{\sigma(1)}, .., u_{\sigma(r)}\right) = \sum_{\sigma \in \mathfrak{s}_r} f\left(u_{\sigma(1)}, .., u_{\sigma(r)}\right)$
$= \sum_{\sigma \in \mathfrak{s}_r} f(u_1, .., u_r) = r! f(u_1, .., u_r)$
So :

Theorem 375 *For any multilinear symmetric map $f \in L^r(E; E')$ there is a unique linear map $F \in L\left(\overset{r}{\otimes}E; E'\right)$ such that : $F \circ s_r = r!f$*

The symmetrizer is a multilinear symmetric map : $s_r : E^r \to \odot^r E :: s_r \in L^r(E^r; \odot^r E) : F = S_r$

By restriction of F on $\odot^r E$ the property still holds : for any multilinear symmetric map $f \in L^r(E; E')$ there is a unique linear map $F \in L(\odot^r E; E')$ such that : $F \circ s_r = r!f$

Symmetric tensorial product

1. Symmetric tensorial product of r vectors

The tensorial product of two symmetric tensors is not necessarily symmetric so, in order to have an internal operation for $S^r(E)$ one defines :

Definition 376 *The symmetric tensorial product of r vectors of E, denoted by $\odot$, is the map :*
$\odot : E^r \to \odot^r E :: u_1 \odot u_2.. \odot u_r = \sum_{\sigma \in \mathfrak{s}_r} u_{\sigma(1)} \otimes .. \otimes u_{\sigma(r)} = s_r(u_1, .., u_r) = S_r \circ \imath(u_1, .., u_r)$

notice that there is no r!

2. Properties of the symmetric tensorial product of r vectors

Theorem 377 *The symmetric tensorial product of r vectors is a multilinear, distributive over addition, symmetric map : $\odot : E^r \to \odot^r E$*

$u_{\sigma(1)} \odot u_{\sigma(2)}.. \odot u_{\sigma(r)} = u_1 \odot u_2.. \odot u_r$
$(\lambda u + \mu v) \odot w = \lambda u \odot w + \mu v \odot w$
Examples:
$u \odot v = u \otimes v + v \otimes u$
$u_1 \odot u_2 \odot u_3 =$
$u_1 \otimes u_2 \otimes u_3 + u_1 \otimes u_3 \otimes u_2 + u_2 \otimes u_1 \otimes u_3 + u_2 \otimes u_3 \otimes u_1 + u_3 \otimes u_1 \otimes u_2 + u_3 \otimes u_2 \otimes u_1$

3. Basis of $\odot^r E$:

Notation 378 *$\odot^r E$ is the subset of $\otimes^r E$ comprised of symmetric tensors*

Theorem 379 *If $(e_i)_{i \in I}$ is a basis of E, with I an ordered set, the set of ordered products $e_{i_1} \odot e_{i_2} \odot ... \odot e_{i_r}, i_1 \le i_2.. \le i_r$ is a basis of $\odot^r E$*

Any r symmetric contravariant tensor can be written equivalentely :
i) $T = \sum_{[i_1...i_r]} t^{i_1...i_r} e_{i_1} \odot e_{i_2} \odot ... \odot e_{i_r}$ with ordered indices
ii) $T = \frac{1}{r!} \sum_{(i_1...i_r)} t^{i_1...i_r} e_{i_1} \otimes ... \otimes e_{i_r}$ with non ordered indices

4. Symmetric tensorial product of symmetric tensors :
The symmetric tensorial product is generalized for symmetric tensors :

a) define for vectors :

$(u_1 \odot ... \odot u_p) \odot (u_{p+1} \odot ... \odot u_{p+q}) = \sum_{\sigma \in \mathfrak{s}_{p+q}} u_{\sigma(1)} \otimes u_{\sigma(2)} ... \otimes u_{\sigma(+q)} = u_1 \odot ... \odot u_p \odot u_{p+1} \odot ... \odot u_{p+q}$

This product is commutative

b) so for any symmetric tensor :

$T = \sum_{[i_1...i_p]} t^{i_1...i_p} e_{i_1} \odot e_{i_2} \odot ...e_{i_p}, U = \sum_{[i_1...j_q]} u^{j_1...j_q} e_{j_1} \odot e_{j_2} \odot ...e_{j_q}$

$T \odot U = \sum_{[i_1...i_p]} \sum_{[i_1...j_q]} t^{i_1...i_p} u^{j_1...j_q} e_{i_1} \odot e_{i_2} \odot ...e_{i_p} \odot e_{j_1} \odot e_{j_2} \odot ...e_{j_q}$

$T \odot U = \sum_{[k_1,..k_{p+q}]_{p+q}} \sum_{[i_1...i_p],[i_1...j_q] \subset [k_1,..k_{p+q}]} t^{i_1...i_p} u^{j_1...j_q} e_{k_1} \odot e_{k_2} ... \odot e_{k_{p+q}}$

Theorem 380 *The symmetric tensorial product of symmetric tensors is a bilinear, distributive over addition, associative, commutative map :* $\odot : \odot^p E \times \odot^q E \to \odot^{p+q} E$

5. Algebra of symmetric tensors:

Theorem 381 *If E is a vector space over the field K, the set* $\odot E = \oplus_{r=0}^{\infty} \odot^r E \subset T(E)$, *with* $\odot^0 E = K, \odot^1 E = E$ *is, with symmetric tensorial product, a graded unital algebra over K, called the* **symmetric algebra S(E)**

Notice that $\odot E \subset T(E)$

Algebraic definition   (Knapp p.645)

The **symmetric algebra S(E)** is the quotient set :

$S(E) = T(E)/$ (two-sided ideal generated by the tensors of the kind $u \otimes v - v \otimes u$ with $u, v \in E$)

The tensor product translates in a symmetric tensor product $\odot$ which makes S(E) an algebra.

With this definition difficulties arise because the elements of S(E) are not tensors (but classes of equivalence) so in practical calulations it is rather confusing.

7.2.2   The set of antisymmetric tensors

Antisymmetric tensors

1. Antisymmetrizer:

Definition 382 *On a vector space E the* **antisymmetrisation operator** *or antisymmetrizer is the map :*

$a_r : E^r \to \overset{r}{\otimes} E :: a_r(u_1, .., u_r) = \sum_{\sigma \in \mathfrak{S}(r)} \epsilon(\sigma) u_{\sigma(1)} \otimes .. \otimes u_{\sigma(r)}$

The antisymmetrizer is an antisymmetric multilinear map : $a_r \in L^r(E^r; A^r(E))$

$a_r\left((u_{\sigma(1)}, u_{\sigma(2)}, ..., u_{\sigma(r)})\right) = \sum_{\sigma' \in \mathfrak{S}(r)} \epsilon(\sigma') u_{\sigma'\sigma(1)} \otimes .. \otimes u_{\sigma'\sigma(r)} = \sum_{\sigma\sigma' \in \mathfrak{S}(r)} \epsilon(\sigma) \epsilon(\sigma\sigma') u_{\sigma'\sigma(1)} \otimes .. \otimes u_{\sigma'\sigma(r)}$

$= \epsilon(\sigma) \sum_{\theta \in \mathfrak{S}(r)} \epsilon(\theta) u_{\theta(1)} \otimes .. \otimes u_{\theta(r)} = \epsilon(\sigma) a_r(u_1, u_2, ..., u_r)$

It is a multilinear map so there is a unique linear map : $\mathfrak{A}_r : \overset{r}{\otimes} E \to \overset{r}{\otimes} E$:
such that : $a_r = \mathfrak{A}_r \circ \imath$ with $\imath : E^r \to \overset{r}{\otimes} E$

$a_r(e_1,..,e_r) = \mathfrak{A}_r \circ \imath (e_1,..,e_r) = \mathfrak{A}_r(e_1 \otimes .. \otimes e_r) = \sum_{\sigma \in \mathfrak{s}_r} \epsilon(\sigma) e_{\sigma(1)} \otimes .. \otimes e_{\sigma(r)}$

For any tensor $T \in \overset{r}{\otimes} E$:

$\mathfrak{A}_r(T) = \sum_{(i_1...i_r)} t^{i_1...i_r} \mathfrak{A}_r(e_{i_1} \otimes ... \otimes e_{i_r}) = \sum_{(i_1...i_r)} t^{i_1...i_r} \sum_{\sigma \in \mathfrak{S}(r)} \epsilon(\sigma) e_{\sigma(1)} \otimes .. \otimes e_{\sigma(r)}$

2. Antisymmetric tensor :

Definition 383 *An antisymmetric r contravariant tensor is a tensor T such that $\mathfrak{A}_r(T) = r!T$*

In a basis a r contravariant antisymmetric tensor $T = \sum_{(i_1...i_r)} t^{i_1...i_r} e_{i_1} \otimes ... \otimes e_{i_r}$ is such that :

$t^{i_1...i_r} = \epsilon(\sigma) t^{\sigma(i_1...i_r)} \Leftrightarrow i_1 < i_2.. < i_k : t^{\sigma(i_1...i_r)} = \epsilon(\sigma(i_1,...i_r)) t^{i_1...i_r}$

where σ is any permutation of the set of r-indices.

It implies that $t^{i_1...i_r} = 0$ whenever two of the indices have the same value.

Thus one can write :

$T = \sum_{\{i_1...i_r\}} t^{i_1...i_r} \left(\sum_{\sigma \in \mathfrak{S}(r)} \epsilon(\sigma) e_{\sigma(i_1)} \otimes ... \otimes e_{\sigma(i_r)} \right)$

An antisymmetric tensor is uniquely defined by a set of components $t^{i_1...i_r}$ for all ordered indices $\{i_1...i_r\}$ with the rule :

$t^{\sigma(i_1...i_r)} = \epsilon(\sigma) t^{i_1...i_r}$

3. Vector space of r antisymmetric tensors

Notation 384 *$\Lambda^r E$ is the set of antisymmetric r-contravariant tensors on E*
$\Lambda_r E^*$ is the set of antisymmetric r-covariant tensors on E

Theorem 385 *The set of antisymmetric r-contravariant tensors $\Lambda^r E$ is a vector subspace of $T^r(E)$.*

A basis of the vector subspace $\Lambda^r E$ is : $e_{i_1} \otimes e_{i_2} \otimes ... \otimes e_{i_r}, i_1 < i_2.. < i_n$
If E is n-dimensional $\dim \Lambda^r E = C_n^r$ and :
- there is no antisymmetric tensor of order r>N
- $\dim \Lambda^n E = 1$ so all antisymmetric n-tensors are proportionnal
- $\Lambda^{n-r} E \simeq \Lambda^r E$: they are isomorphic vector spaces

4. Universal property:
A r-linear antisymmetric map $f \in L^r(E;E')$ is such that :
$\forall u_i \in E, i = 1...r, \sigma \in \mathfrak{S}_r : f(u_1,u_2,...,u_r) = \epsilon(\sigma) f(u_{\sigma(1)},u_{\sigma(2)},...,u_{\sigma(r)})$

There is a unique linear map : $F \in L\left(\overset{r}{\otimes} E;E'\right)$ such that : $f = F \circ \imath$

$F \circ a_r(u_1,u_2,...,u_r) = \sum_{\sigma \in \mathfrak{S}(r)} \epsilon(\sigma) F\left(u_{\sigma(1)} \otimes .. \otimes u_{\sigma(r)}\right) = \sum_{\sigma \in \mathfrak{S}(r)} \epsilon(\sigma) F \circ \imath\left(u_{\sigma(1)},..,u_{\sigma(r)}\right) = \sum_{\sigma \in \mathfrak{S}(r)} \epsilon(\sigma) f\left(u_{\sigma(1)},..,u_{\sigma(r)}\right)$

$= \sum_{\sigma \in \mathfrak{S}(r)} f(u_1,..,u_r) = r! f(u_1,..,u_r)$

So :

Theorem 386 *For any multilinear antisymmetric map $f \in L^r(E; E')$ there is a unique linear map $F \in L\left(\overset{r}{\otimes} E; E'\right)$ such that :*

Theorem 387 $F \circ a_r = r!f$

For $f = a_r : F = \mathfrak{A}_r$

By restriction of F on $\Lambda^r E$ the property still holds : for any multilinear antisymmetric map $f \in L^r(E; E')$ there is a unique linear map $F \in L(\Lambda^r E; E')$ such that : $F \circ a_r = r!f$

Exterior product

1. Exterior product of vectors:

The tensor product of 2 antisymmetric tensor is not necessarily antisymmetric so, in order to have an internal operation for $A^r(E)$ one defines :

Definition 388 *The **exterior product** (or **wedge product**) of r vectors is the map :*

$\wedge : E^r \rightarrow \Lambda^r E :: u_1 \wedge u_2 .. \wedge u_r = \sum_{\sigma \in \mathfrak{S}(r)} \epsilon(\sigma) u_{\sigma(1)} \otimes .. \otimes u_{\sigma(r)} = a_r(u_1, .., u_r)$

notice that there is no r!

Theorem 389 *The exterior product of vectors is a multilinear, antisymmetric map , which is distributive over addition*

$u_{\sigma(1)} \wedge u_{\sigma(2)} .. \wedge u_{\sigma(r)} = \epsilon(\sigma) u_1 \wedge u_2 .. \wedge u_r$
$(\lambda u + \mu v) \wedge w = \lambda u \wedge w + \mu v \wedge w$
Moreover :
$u_1 \wedge u_2 .. \wedge u_r = 0 \Leftrightarrow$ the vectors are linearly dependant
$u \wedge v = 0 \Leftrightarrow \exists k \in K : u = kv$
Examples :
$u \wedge v = u \otimes v - v \otimes u$
$u_1 \wedge u_2 \wedge u_3 =$
$u_1 \otimes u_2 \otimes u_3 - u_1 \otimes u_3 \otimes u_2 - u_2 \otimes u_1 \otimes u_3 + u_2 \otimes u_3 \otimes u_1 + u_3 \otimes u_1 \otimes u_2 - u_3 \otimes u_2 \otimes u_1$

2. Basis of $\Lambda^r E$:

Theorem 390 *The set of antisymmetric tensors : $e_{i_1} \wedge e_{i_2} \wedge ... e_{i_r}, i_1 < i_2 .. < i_n$, is a basis of $\Lambda^r E$*

3. Exterior product of antisymmetric tensors:

The exterior product is generalized between antisymmetric tensors:

a) define for vectors :

$(u_1 \wedge ... \wedge u_p) \wedge (u_{p+1} \wedge ... \wedge u_{p+q}) = \sum_{\sigma \in \mathfrak{s}_{p+q}} \epsilon(\sigma) u_{\sigma(1)} \otimes u_{\sigma(2)} ... \otimes u_{\sigma(p+q)} = u_1 \wedge ... \wedge u_p \wedge u_{p+1} \wedge ... \wedge u_{p+q}$

113

Notice that *it is not anticommutative* : $(u_1 \wedge ... \wedge u_p) \wedge (u_{p+1} \wedge ... \wedge u_{p+q}) =$
$(-1)^{pq} (u_{p+1} \wedge ... \wedge u_{p+q}) \wedge (u_1 \wedge ... \wedge u_p)$

b) so for any antisymmetric tensor :

$T = \sum_{\{i_1...i_p\}} t^{i_1...i_p} e_{i_1} \wedge e_{i_2} \wedge ...e_{i_p}, U = \sum_{\{i_1...j_q\}} u^{j_1...j_q} e_{j_1} \wedge e_{j_2} \wedge ...e_{j_q}$

$T \wedge U = \sum_{\{i_1...i_p\}} \sum_{\{i_1...j_q\}} t^{i_1...i_p} u^{j_1...j_q} e_{i_1} \wedge e_{i_2} \wedge ...e_{i_p} \wedge e_{j_1} \wedge e_{j_2} \wedge ...e_{j_q}$

$T \wedge U = \frac{1}{p!q!} \sum_{(i_1...i_p)} \sum_{(i_1...j_q)} t^{i_1...i_p} u^{j_1...j_q} e_{i_1} \wedge e_{i_2} \wedge ...e_{i_p} \wedge e_{j_1} \wedge e_{j_2} \wedge ...e_{j_q}$

$e_{i_1} \wedge e_{i_2} \wedge ...e_{i_p} \wedge e_{j_1} \wedge e_{j_2} \wedge ...e_{j_q} = \epsilon (i_1, ..i_p, j_1, ...j_q) e_{k_1} \wedge e_{k_2} \wedge .. \wedge e_{k_{p+q}}$

where $(k_1, ...k_{p+q})$ is the ordered set of indices : $(i_1, ..i_p, j_1, ...j_q)$

Expressed in the basis $e_{i_1} \wedge e_{i_2} ... \wedge e_{i_{p+q}}$ of $\Lambda^{p+q} E$:

$T \Lambda S =$
$\sum_{\{j_1, ..j_{p+q}\}_{p+q}} \left(\sum_{\{j_1, ..j_p\}, \{j_{p+1}, ...j_{p+q}\} \subset \{i_1, ..i_{p+q}\}} \epsilon (j_1, ..j_p, j_{p+1}, ...j_{p+q}) T^{\{j_1, ..j_p\}} S^{\{j_{p+1}, ...j_{p+q}\}} \right)$
$e_{j_1} \wedge e_{j_2} ... \wedge e_{j_{p+q}}$

or with

$\{A\} = \{j_1, ..j_p\}, \{B\} = \{j_{p+1}, ...j_{p+q}\}, \{C\} = \{j_1, ..j_p, j_{p+1}, ...j_{p+q}\} = \{\{A\} \cup \{B\}\}$
$\{B\} = \{j_{p+1}, ...j_{p+q}\} = \{C / \{A\}\}$

$T \Lambda S = \sum_{\{C\}_{p+q}} \left(\sum_{\{A\}_p} \epsilon (\{A\}, \{C / \{A\}\}) T^{\{A\}} S^{\{C/\{A\}\}} \right) \wedge e_{\{C\}}$

4. Properties of the exterior product of antisymmetric tensors:

Theorem 391 *The wedge product of antisymmetric tensors is a multilinear, distributive over addition, associative map :*
$\wedge : \Lambda^p E \times \Lambda^q E \to \Lambda^{p+q} E$

Moreover:
$T \wedge U = (-1)^{pq} U \wedge T$
$k \in K : T \wedge k = kT$

5. Algebra of antisymmetric tensors:

Theorem 392 *For the vector space E over the field K, the set denoted : $\Lambda E = \oplus_{n=0}^{\dim E} \Lambda^n E$ with $\Lambda^n E = K$ is, with the exterior product, a graded unital algebra (the identity element is $1 \in K$) over K*

$\dim \Lambda E = 2^{\dim E}$

The elements T of ΛE which can be written as : $T = u_1 \wedge u_2 ... \wedge u_r$ are homogeneous.

Theorem 393 *An antisymmetric tensor is homogeneous iff $T \wedge T = 0$*

Warning ! usually $T \wedge T \neq 0$
There are the algebra isomorphisms :
$\hom (\Lambda^r E, F) \simeq L_A^r (E^r; F)$ antisymmetric multilinear maps
$\Lambda^r E^* \simeq (\Lambda^r E)^*$

6. Determinant of an endomorphism

Theorem 394 *On a finite dimensional vector space E on a field K there is a unique map, called* **determinant** *:*

$\det : L(E;E) \to K$ *such that* $\forall u_1, u_2, ...u_n \in E : f(u_1) \wedge f(u_2) ... \wedge f(u_n) = (\det f) u_1 \wedge u_2 ... \wedge u_n$

Proof. $F = a_r \circ f : E^n \to \Lambda^n E :: F(u_1, ..., u_n) = f(u_1) \wedge f(u_2) ... \wedge f(u_n)$
is a multilinear, antisymmetric map. So there is a unique linear map $D : \Lambda^n E \to \Lambda^n E$ such that

$D \circ a_r = n! F$

$F(u_1, ..., u_n) = f(u_1) \wedge f(u_2) ... \wedge f(u_n) = \frac{1}{n!} D(u_1 \wedge ... \wedge u_n)$

As all the n-antisymmetric tensors are proportional, $D(u_1 \wedge ... \wedge u_n) = k(f)(u_1 \wedge ... \wedge u_n)$ with $k : L(E;E) \to K$. ∎

Algebraic definition
(see Knapp p.651).

The algebra $A(E)$ is defined as the quotient set : $A(E) = T(E)/(I)$ where I =two-sided ideal generated by the tensors of the kind $u \otimes v + v \otimes u$ with $u, v \in E$. The set of its homogeneous elements of order r is denoted $A^r(E)$, $A^0(E) = K$

The interior product of T(E), that is the tensor product, goes in $A(E)$ as an interior product denoted $\wedge$ and called wedge product, with which $A(E)$ is an algebra.

If $(e_i)_{i \in I}$ is a basis of E, with I an ordered set, the set of ordered products $e_{i_1} \wedge e_{i_2} \wedge ... e_{i_n}, i_1 < i_2 .. < i_n$, is a basis of $A^n(E)$

So the properties are the same than above, but $A^r(E)$ *is not a subset* of $\overset{r}{\otimes} E$.

$A^r(E)$ (as defined algebraically here) is isomorphic (as vector space) to $\Lambda^r E$ with :

$T = \sum_{\{i_1 ... i_r\}} t^{i_1 ... i_r} e_{i_1} \wedge e_{i_2} \wedge ... e_{i_r} \in A^r(E) \leftrightarrow T' = \sum_{\{i_1 ... i_r\}} \left(\frac{1}{r!} \sum_{\sigma \in \mathfrak{S}(r)} \epsilon(\sigma) t^{\sigma(i_1 ... i_r)} \right) e_{i_1} \otimes ... \otimes e_{i_r} \in \Lambda^r E$

The wedge product is computed differently. Algebraically :

$T, U \in A(E) \to T \wedge U = T \otimes U \pmod{I}$

and more plainly one identifies $A^r(E)$ with its image in $\Lambda^r E$ and writes :

$u_1 \wedge .. \wedge u_r = \frac{1}{r!} \sum_{\sigma \in \mathfrak{s}_n} \epsilon(\sigma(1, ..r)) u_{\sigma(1)} \otimes .. \otimes u_{\sigma(r)}$

With this definition :

$u_1 \wedge u_2 = \frac{1}{2}(u_1 \otimes u_2 - u_2 \otimes u_1)$

$u_1 \wedge u_2 \wedge u_3 =$

$\frac{1}{3!}(u_1 \otimes u_2 \otimes u_3 - u_1 \otimes u_3 \otimes u_2 - u_2 \otimes u_1 \otimes u_3 + u_2 \otimes u_3 \otimes u_1 + u_3 \otimes u_1 \otimes u_2 - u_3 \otimes u_2 \otimes u_1)$

But now to define the wedge product of $u_1 \wedge u_2 \in A^2(E)$ and u_3 is not so easy (there is no clear and indisputable formula).

So, in order to avoid all these factorials, in this paper we will only consider antisymmetric tensors (and not bother with the quotient space). But it is common to meet the wedge product defined with factorials.

7.2.3   Exterior algebra

All the previous material can be easily extended to the dual E* of a vector space, but the exterior algebra ΛE^* is by far more widely used than ΛE and has some specific properties which must be known.

r-forms
1. Definition:

Definition 395 *The **exterior algebra** (also called Grassman algebra) of a vector space E is the algebra $\Lambda E^* = \Lambda\left(E^*\right) = \left(\Lambda E\right)^*$.*

So $\Lambda E^* = \oplus_{r=0}^{\dim E}\Lambda_r E^*$ and $\Lambda_0 E^* = K, \Lambda_1 E^* = E^*$ (all indices down)

The tensors of $\Lambda_r E^*$ are called **r-forms** : they are antisymmetric multilinear functions $E^r \to K$

2. Components:

In the following E is a n-dimensional vector space with basis $\left(e_i\right)_{i=1}^n$, and the dual basis $\left(e^i\right)_{i=1}^n$ of E*:$e^i\left(e_j\right) = \delta_j^i$

So $\varpi \in \Lambda_r E^*$ can be written equivalently :

i) $\varpi = \sum_{\{i_1...i_r\}} \varpi_{i_1...i_r} e^{i_1} \wedge e^{i_2} \wedge ... \wedge e^{i_r}$ with ordered indices

ii) $\varpi = \frac{1}{r!}\sum_{(i_1...i_r)} \varpi_{i_1...i_r} e^{i_1} \wedge e^{i_2} \wedge ... \wedge e^{i_r}$ with non ordered indices

iii) $\varpi = \sum_{\{i_1...i_r\}} \varpi_{i_1...i_r} e^{i_1} \otimes e^{i_2} \otimes ... \otimes e^{i_r}$ with ordered indices

iv) $\varpi = \frac{1}{r!}\sum_{(i_1...i_r)} \varpi_{i_1...i_r} e^{i_1} \otimes e^{i_2} \otimes ... \otimes e^{i_r}$ with non ordered indices

3. Change of basis:

In a change of basis : $f_i = \sum_{j=1}^n P_i^j e_j$ the components of tensors change according to the following rules :

$\varpi = \sum_{(i_1...i_r)} \varpi_{i_1...i_r} e^{i_1} \otimes e^{i_2} \otimes ... \otimes e^{i_r} \to \varpi = \sum_{(i_1...i_r)} \widetilde{\varpi}_{i_1...i_r} f^{i_1} \otimes f^{i_2} \otimes$
$... \otimes f^{i_r} = \sum_{\{i_1...i_r\}} \widetilde{\varpi}_{i_1...i_r} f^{i_1} \wedge f^{i_2} \wedge ... \wedge f^{i_r}$

with $\widetilde{\varpi}_{i_1...i_r} = \sum_{(j_1....j_r)} \varpi_{j_1...j_r} P_{i_1}^{j_1}...P_{i_r}^{j_r} = \sum_{\{j_1....j_r\}} \epsilon\left(\sigma\right)\varpi_{j_1...j_r} P_{i_1}^{\sigma(j_1)}...P_{i_r}^{\sigma(j_r)} = \sum_{\{j_1....j_r\}} \varpi_{j_1...j_r} \det\left[P\right]_{i_1...i_r}^{j_1...j_r}$

where $\det\left[P\right]_{i_1...i_r}^{j_1...j_r}$ is the determinant of the matrix with r column $(i_1,..i_r)$ comprised each of the components $(j_1...j_r)$ of the new basis vectors

Interior product
1. Value of a r forms over r vectors:

The value of a r-form over r vectors of E is :

$\varpi = \sum_{(i_1...i_r)} \varpi_{i_1...i_r} e^{i_1} \otimes e^{i_2} \otimes ... \otimes e^{i_r}$

$\varpi\left(u_1,...,u_r\right) = \sum_{(i_1...i_r)} \varpi_{i_1...i_r} e^{i_1}\left(u_1\right) e^{i_2}\left(u_2\right)...e^{i_r}\left(u_r\right)$

$\varpi\left(u_1,...,u_r\right) = \sum_{(i_1...i_r)} \varpi_{i_1...i_r} u_1^{i_1} u_2^{i_2}...u_r^{i_r}$

The value of the exterior product of a p-form and a q-form $\varpi \wedge \pi$ for p+q vectors is given by the formula (Kolar p.62):

$\varpi \Lambda \pi \left(u_1, ..., u_{p+q} \right) = \frac{1}{p!q!} \sum_{\sigma \in \mathfrak{S}_{p+q}} \epsilon \left(\sigma \right) \varpi \left(u_{\sigma(1)}, ... u_{\sigma(p)} \right) \pi \left(u_{\sigma(p+1)}, ... u_{\sigma(p+q)} \right)$

If $r = \dim E : \varpi_{i_1 ... i_n} = \epsilon \left(i_1, ... i_n \right) \varpi_{12...n}$

$\varpi \left(u_1, ..., u_n \right) = \varpi_{12...n} \sum_{\sigma \in \mathfrak{S}_n} \epsilon \left(\sigma \right) u_1^{\sigma(1)} u_2^{\sigma(2)} ... u_n^{\sigma(n)} = \varpi_{12...n} \det \left[u_1, u_2, ... u_n \right]$

This is the determinant of the matrix with columns the components of the vectors u

2. Interior product:

Definition 396 *The **interior product** of a r form $\varpi \in \Lambda_r E^*$ and a vector $u \in E$, denoted $i_u \varpi$, is the r-1-form :*

$i_u \varpi = \sum_{\{i_1 ... i_r\}} \sum_{k=1}^r (-1)^{k-1} u^{i_k} \varpi_{\{i_1 ... i_r\}} e^{\{i_1} \Lambda ... \Lambda \widehat{e^{i_k}} ... \Lambda e^{i_p\}}$ *where ^ means that the vector shall be omitted*

with $\left(e^i \right)_{i \in I}$ a basis of E*.

3. Properties:

For u fixed the map : $i_u : \Lambda_r E^* \rightarrow \Lambda_{r-1} E^*$ is linear :$i_u \in L \left(\Lambda E; \Lambda E \right)$

$i_u \circ i_v = -i_v \circ i_u$

$i_u \circ i_u = 0$

$i_u \left(\lambda \wedge \mu \right) = \left(i_u \lambda \right) \wedge \mu + (-1)^{\deg \lambda} \lambda \wedge \mu$

Orientation of a vector space

For any n dimensional vector space E a basis can be chosen and its vectors labelled $e_1, ... e_n$. One says that there are two possible orientations : direct and indirect according to the value of the signature of any permutation of these vectors. A vector space is orientable if it is possible to compare the orientation of two different bases.

A change of basis is defined by an endomorphism $f \in GL \left(E; E \right)$. Its determinant is such that :

$\forall u_1, u_2, ... u_n \in E : f \left(u_1 \right) \wedge f \left(u_2 \right) ... \wedge f \left(u_n \right) = \left(\det f \right) u_1 \wedge u_2 ... \wedge u_n$

So if E is a real vector space det(f) is a non null real scalar, and two bases have the same orientation if det(f) > 0.

If E is a complex vector space, it has a real structure such that : $E = E_{\mathbb{R}} \oplus i E_{\mathbb{R}}$. So take any basis $\left(e_i \right)_{i=1}^n$ of $E_{\mathbb{R}}$ and say that the basis : $\left(e_1, i e_1, e_2, i e_2, ... e_n, i e_n \right)$ is direct. It does not depend on the choice of $\left(e_i \right)_{i=1}^n$ and is called the canonical orientation of E.

To sum up :

Theorem 397 *All finite dimensional vector spaces over $\mathbb{R}$ or $\mathbb{C}$ are orientable.*

Volume

Definition 398 *A volume form on a n dimensional vector space $\left(E, g \right)$ with scalar product is a n-form ϖ such that its value on any direct orthonormal basis is 1.*

Theorem 399 *In any direct basis* $\left(e^i\right)_{i=1}^n$ *a volume form is* $\varpi = \sqrt{|\det g|}e_1 \wedge e_2.. \wedge e_n$

In any orthonormal basis $(\varepsilon_i)_{i=1}^n$ $\varpi = \varepsilon_1 \wedge \varepsilon_2.. \wedge \varepsilon_n$

Proof. (E,g) is endowed with a bilinear symmetric form g, non degenerate (but not necessarily definite positive).

In $\left(e^i\right)_{i=1}^n$ g has for matrix is $[g] = [g_{ij}]$. $g_{ij} = g\left(e_i, e_j\right)$

Let $\varepsilon_i = \sum_j P_i^j e_j$ then g has for matrix in $\varepsilon_i : [\eta] = [P]^* [g] [P]$ with $\eta_{ij} = \pm\delta_{ij}$

The value of $\varpi\left(\varepsilon_1, \varepsilon_2, ...\varepsilon_n\right) = \varpi_{12...n} \det\left[\varepsilon_1, \varepsilon_2, ...\varepsilon_n\right] = \varpi_{12...n} \det\left[P\right]$

But : $\det\left[\eta\right] = \det\left([P]^* [g] [P]\right) = |\det\left[P\right]|^2 \det\left[g\right] = \pm 1$ depending on the signature of g

If E is a real vector space, then $\det\left[P\right] > 0$ as the two bases are direct. So : $\det\left[P\right] = 1/\sqrt{|\det g|}$ and :

$\varpi = \sqrt{|\det g|}e_1 \wedge e_2.. \wedge e_n = \varepsilon_1 \wedge \varepsilon_2.. \wedge \varepsilon_n$

If E is a complex vector space $[g] = [g]^*$ and $\det[g]^* = \overline{\det[g]} = \det[g]$ so $\det[g]$ is real. it is always possible to choose an orthonormal basis such that : $\eta_{ij} = \delta_{ij}$ so we can still take $\varpi = \sqrt{|\det g|}e_1 \wedge e_2.. \wedge e_n = \varepsilon_1 \wedge \varepsilon_2.. \wedge \varepsilon_n$ ∎

Definition 400 *The **volume** spanned by n vectors* $(u_1, ..., u_n)$ *of a real n dimensional vector space* (E, g) *with scalar product endowed with the volume form* ϖ *is* $\varpi\left(u_1, ..., u_n\right)$

It is null if the vectors are linearly dependant.

Maps of the special orthogonal group SO(E,g) preserve both g and the orientation, so they preserve the volume.

7.3 Tensorial product of maps

7.3.1 Tensorial product of maps

1. Maps on contravariant or covariant tensors:

The following theorems are the consequences of the universal property of the tensorial product, implemented to the vector spaces of linear maps.

Theorem 401 *For any vector spaces* E_1, E_2, F_1, F_2 *on the same field,* $\forall f_1 \in L\left(E_1; F_1\right), f_2 \in L(E_2; F_2)$ *there is a unique map denoted* $f_1 \otimes f_2 \in L\left(E_1 \otimes F_1; E_2 \otimes F_2\right)$ *such that :* $\forall u \in E_1, v \in F_1 : (f_1 \otimes f_2)(u \otimes v) = f_1(u) \otimes f_2(v)$

Theorem 402 *For any vector spaces E,F on the same field,* $\forall r \in \mathbb{N}$. $\forall f \in L\left(E; F\right)$

i) there is a unique map denoted $\otimes^r f \in L\left(\otimes^r E; \otimes^r F\right)$ *such that :*
$\forall u_k \in E, k = 1...r : \left(\otimes^r f\right)\left(u_1 \otimes u_2... \otimes u_r\right) = f\left(u_1\right) \otimes f\left(u_2\right)... \otimes f\left(u_r\right)$

ii) there is a unique map denoted $\otimes_s f^t \in L\left(\otimes^s F^*; \otimes^s E^*\right)$ *such that :*
$\forall \lambda_k \in F^*, k = 1...s : \left(\otimes_s f^t\right)\left(\lambda_1 \otimes \lambda_2... \otimes \lambda_r\right) = f^t\left(\lambda_1\right) \otimes f^t\left(\lambda_2\right)... \otimes f^t\left(\lambda_r\right) = \left(\lambda_1 \circ f\right) \otimes \left(\lambda_2 \circ f\right)... \otimes \left(\lambda_s \circ f\right)$

2. Maps on mixed tensors

If f is inversible : $f^{-1} \in L(F;E)$ and $\left(f^{-1}\right)^t \in L(E^*;F^*)$. So to extend a map from $L(E;F)$ to $L(\otimes_s^r E; \otimes_s^r F)$ an inversible map $f \in GL(E;F)$ is required.

Take as above :

$E_1 = \otimes^r E, E_2 = \otimes^s E^*, F_1 = \otimes^r F, F_2 = \otimes_s F^* = \otimes^s F^*,$

$f \in L(E_1;F_1), \otimes^r f \in L(\otimes^r E; \otimes^r F)$

$f^{-1} \in L(F_2;E_2), \otimes^s \left(f^{-1}\right)^t = L(\otimes^s E^*; \otimes^s F^*)$

There is a unique map : $(\otimes^r f) \otimes \left(\otimes^s \left(f^{-1}\right)^t\right) \in L(\otimes_s^r E; \otimes_s^r F)$ such that :

$\forall u_k \in E, \lambda_l \in E^*, k = ..r, l = ..s :$

$(\otimes^r f) \otimes \left(\otimes^s \left(f^{-1}\right)^t\right) ((u_1 \otimes u_2 ... \otimes u_r) \otimes (\lambda_1 \otimes \lambda_2 ... \otimes \lambda_s)) = f(u_1) \otimes f(u_2) ... \otimes f(u_r) \otimes f(\lambda_1) \otimes f(\lambda_2) ... \otimes f(\lambda_r)$

This can be done for any r,s and from a map $f \in L(E;F)$ build a family of linear maps $\otimes_s^r f = (\otimes^r f) \otimes \left(\otimes^s \left(f^{-1}\right)^t\right) \in L(\otimes_s^r E; \otimes_s^r F)$ such that the maps commute with the trace operator and preserve the tensorial product :

$S \in \otimes_s^r E, T \in \otimes_{s'}^{r'} E : F_{s+s'}^{r+r'}(S \otimes T) = F_s^r(S) \otimes F_{s'}^{r'}(T)$

3. These results are summarized in the following theorem :

Theorem 403 *(Kobayashi p.24) For any vector spaces E,F on the same field, there is an isomorphism between the isomorphisms in L(E;F) and the isomorphisms of algebras $L(\otimes E; \otimes F)$ which preserves the tensor type and commute with contraction. So there is a unique extension of an isomorphism f$\in L(E,F)$ to a linear bijective map $F \in L(\otimes E; \otimes F)$ such that $F(S \otimes T) = F(S) \otimes F(T)$, F preserves the type and commutes with the contraction. And this extension can be defined independantly of the choice of bases.*

Let E be a vector space and G a subgroup of GL(E;E). Then any fixed f in G is an isomorphism of L(E;E) and can be extended to a unique linear bijective map $F \in L(\otimes E; \otimes E)$ such that $F(S \otimes T) = F(S) \otimes F(T)$, F preserves the type and commutes with the contraction. For $F(T,1) : \otimes E \to \otimes E$ we have a linear map.

7.3.2   Tensorial product of bilinear forms

Derivatives of maps are multilinear symmetric linear maps. It can be handy to extend a scalar product from the vector spaces to the spaces of these multilinear maps. We see here how it can be done.

1. Bilinear form on $\otimes^r E$:

Theorem 404 *A bilinear symmetric form on a finite n dimensional vector space E over the field K can be extended to a bilinear symmetric form : G_r : $\otimes^r E \times \otimes^r E \to K :: G_r = \otimes^r g$*

Proof. $g \in E^* \otimes E^*$. g reads in in a basis $\left(e^i\right)_{i=1}^n$ of E* : $g = \sum_{i,j=1}^n g_{ij} e^i \otimes e^j$

The r tensorial product of g : $\otimes^r g \in \otimes^{2r} E^*$ reads : $\otimes^r g = \sum_{i_1 \ldots i_2 = 1}^n g_{i_1 i_2} \ldots g_{i_{2r-1} i_{2r}} e^{i_1} \otimes e^{i_2} \ldots \otimes e^{i_{2r}}$

It acts on tensors $U \in \otimes^{2r} E$: $\otimes^r g (U) = \sum_{i_1 \ldots i_2 = 1}^n g_{i_1 i_2} \ldots g_{i_{2r-1} i_{2r}} U^{i_1 \ldots i_{2r}}$

Take two r contravariant tensors S,T$\in \otimes^r E$ then

$\otimes^r g (S \otimes T) = \sum_{i_1 \ldots i_2 = 1}^n g_{i_1 i_2} \ldots g_{i_{2r-1} i_{2r}} S^{i_1 \ldots i_r} T^{i_{r+1} \ldots i_{2r}}$

From the properties of the tensorial product :

$\otimes^r g \left((kS + k'S') \otimes T\right) = k \otimes^r g (S \otimes T) + k' \otimes^r g (S' \otimes T)$

So it can be seen as a bilinear form acting on $\otimes^r E$. Moreover it is symmetric :

$$G_r (S,T) = \otimes^r g (S \otimes T) = G_r (T, S) \quad \blacksquare$$

2. Bilinear form on L^r(E;E):

Theorem 405 *A bilinear symmetric form on a finite n dimensional vector space E over the field K can be extended to a bilinear symmetric form : B_r : $\otimes^r E^* \times \otimes^r E \to K :: B_r = \otimes^r g^* \otimes g$*

Proof. The vector space of r linear maps L^r(E;E) is isomorphic to the tensorial subspace : $\otimes^r E^* \otimes E$

We define a bilinear symmetric form on L^r(E;E) as follows :

$\varphi, \psi \in L^r(E; E)$:$B_r (\varphi, \psi) = B_r (\varphi \otimes \psi)$

with : $B_r = \otimes^r g^* \otimes g = \sum_{i_1 \ldots i_2 = 1}^n g^{i_1 i_2} \ldots g^{i_{2r-1} i_{2r}} g_{j_1 j_2} e_{i_1} \otimes \ldots \otimes e_{i_{2r}} \otimes e^{j_1} \otimes e^{j_2}$

This is a bilinear form, and it is symmetric because g is symmetric. $\quad \blacksquare$

Notice that if E is a complex vector space and g is hermitian we do not have a hermitian scalar product.

7.3.3 Hodge duality

Hodge duality is a special case of the previous construct : if the tensors are anti-symmetric then we get the determinant. However we will extend the study to the case of hermitian maps, because it will be used later.

Remind that a vector space (E,g) on a field K is endowed with a scalar product if g is either a non degenerate, bilinear symmetric form, or a non degenerate hermitian form.

Scalar product of r-forms

Theorem 406 *If (E,g) is a finite dimensional vector space endowed with a scalar product, then the map :*

$G_r : \Lambda_r E^* \times \Lambda_r E^* \to \mathbb{R} :: G_r (\lambda, \mu) = \sum_{\{i_1 \ldots i_r\} \{j_1 \ldots j_r\}} \overline{\lambda}_{i_1 \ldots i_r} \mu_{j_1 \ldots j_r} \det \left[g^{-1}\right]^{\{i_1 \ldots i_r\}, \{j_1 \ldots j_r\}}$

is a non degenerate hermitian form and defines a scalar product which does not depend on the basis.

It is definite positive if g is definite positive

In the matrix $\left[g^{-1}\right]$ one takes the elements $g^{i_k j_l}$ with $i_k \in \{i_1..i_r\}, j_l \in \{j_1..j_r\}$

$G_r\left(\lambda, \mu\right) = \sum_{\{i_1...i_r\}} \overline{\lambda}_{\{i_1...i_r\}} \sum_{j_1...j_r} g^{i_1 j_1}...g^{i_r j_r} \mu_{j_1...j_r} = \sum_{\{i_1...i_r\}} \overline{\lambda}_{\{i_1...i_r\}} \mu^{\{i_1 i_2...i_r\}}$

where the indexes are lifted and lowered with g.

In an orthonormal basis : $G_r\left(\lambda, \mu\right) = \sum_{\{i_1..i_r\}\{j_1..j_r\}} \overline{\lambda}_{i_1..i_r} \mu_{j_1...j_r} \eta^{i_1 j_1}...\eta^{i_r j_r}$

This is the application of the first theorem of the previous subsection, where the formula for the determinant is used.

For $r = 1$ one gets the usual bilinear symmetric form over $E^* : G_1\left(\lambda, \mu\right) = \sum_{ij} \overline{\lambda}_i \mu_j g^{ij}$

Theorem 407 *For a vector u fixed in (E, g), the map : $\lambda\left(u\right) : \Lambda_r E \to \Lambda_{r+1} E ::$ $\lambda\left(u\right)\mu = u \wedge \mu$ has an adjoint with respect to the scalar product of forms : $G_{r+1}\left(\lambda\left(u\right)\mu, \mu'\right) = G_r\left(\mu, \lambda^*\left(u\right)\mu'\right)$ which is $\lambda^*\left(u\right) : \Lambda_r E \to \Lambda_{r-1} E ::$ $\lambda^*\left(u\right)\mu = i_u \mu$*

It suffices to compute the two quantities.

Hodge duality

g can be used to define the isomorphism $E \simeq E^*$. Similarly this scalar product can be used to define the isomorphism $\Lambda_r E \simeq \Lambda_{n-r} E$

Theorem 408 *If (E,g) is a n dimensional vector space endowed with a scalar product with the volume form ϖ_0, then the map :*
$* : \Lambda_r E^* \to \Lambda_{n-r} E$ defined by the condition $\forall \mu \in \Lambda_r E^* : *\lambda_r \wedge \mu = G_r\left(\lambda, \mu\right)\varpi_0$
is an anti-isomorphism

A direct computation gives the value of the Hodge dual $*\lambda$ in the basis $\left(e^i\right)_{i=1}^n$ of E*:

$*\left(\sum_{\{i_1...i_r\}} \lambda_{\{i_1...i_r\}} e^{i_1} \wedge .. \wedge e^{i_r}\right)$
$= \sum_{\{i_1..i_{n-r}\}\{j_1..j_r\}} \epsilon\left(j_1..j_r, i_1, ...i_{n-r}\right) \overline{\lambda}^{j_1...j_r} \sqrt{|\det g|} e^{i_1} \wedge e^{i_2}... \wedge e^{i_{n-r}}$
With $\epsilon = sign \det[g]$ (which is always real)
For r=0:
$*\lambda = \overline{\lambda}\varpi_0$
For r=1 :
$*\left(\sum_i \lambda_i e^i\right) = \sum_{j=1}^n \left(-1\right)^{j+1} g^{ij} \overline{\lambda}_j \sqrt{|\det g|} e^1 \wedge ..\widehat{e^j} \wedge ... \wedge e^n$
For r=n-1:
$*\left(\sum_{i=1}^n \lambda_{1..\widehat{i}...n} e^1 \wedge ..\widehat{e^i} \wedge ... \wedge e^n\right) = \sum_{i=1}^n \left(-1\right)^{i-1} \overline{\lambda}^{1..\widehat{i}...n} \sqrt{|\det g|} dx^i$
For r=n:
$*\left(\lambda e^1 \wedge ..... \wedge e^n\right) = \epsilon \frac{1}{\sqrt{|\det g|}} \overline{\lambda}$
The usual cross product of 2 vectors in an 3 dimensional euclidean vector space can be defined as $u \times v = *\left(a \Lambda b\right)$ where the algebra $\Lambda^r E$ is used
The inverse of the map * is :
$*^{-1}\lambda_r = \epsilon(-1)^{r(n-r)} * \lambda_r \Leftrightarrow * * \lambda_r = \epsilon(-1)^{r(n-r)}\lambda_r$
$G_q(\lambda, *\mu) = G_{n-q}(*\lambda, \mu)$

$G_{n-q}(*\lambda, *\mu) = G_q(\lambda, \mu)$

Contraction is an operation over ΛE^*. It is defined, on a real vector space by:

$\lambda \in \Lambda_r E, \mu \in \Lambda_q E : \lambda \vee \mu = \epsilon \left(-1\right)^{p+(r-q)n} * \left(\lambda \wedge *\mu\right) \in \wedge_{r-q} E^*$

It is distributive over addition and not associative

$* \left(\lambda \vee \mu\right) = \epsilon \left(-1\right)^{(r-q)n} \epsilon (-1)^{(r-q)(n-(r-q))} \left(\lambda \wedge *\mu\right) = (-1)^{q^2+r^2} \left(\lambda \wedge *\mu\right)$

$\lambda \vee \left(\lambda \vee \mu\right) = 0$

$\lambda \in E^*, \mu \in \Lambda_q E :$

$* \left(\lambda \wedge \mu\right) = (-1)^q \lambda \vee *\mu$

$* \left(\lambda \vee \mu\right) = (-1)^{q-1} \lambda \wedge *\mu$

7.3.4 Tensorial Functors

These functors will be used later in several parts.

Theorem 409 *The vector spaces over a field K with their morphisms form a category $\mathfrak{V}$.*

The vector spaces isomorphic to some vector space E form a subcategory $\mathfrak{V}_E$

Theorem 410 *The functor $\mathfrak{D} : \mathfrak{V} \mapsto \mathfrak{V}$ which associates :*
to each vector space E its dual : $\mathfrak{D}(E) = E^*$
to each linear map $f : E \to F$ its dual : $f^* : F^* \to E^*$
is contravariant : $\mathfrak{D}\left(f \circ g\right) = \mathfrak{D}\left(g\right) \circ \mathfrak{D}\left(f\right)$

Theorem 411 *The r-tensorial power of vector spaces is a faithful covariant functor $\mathfrak{T}^r : \mathfrak{V} \mapsto \mathfrak{V}$*

$\mathfrak{T}^r \left(E\right) = \otimes^k E$

$f \in L\left(E; F\right) : \mathfrak{T}^r \left(f\right) = \otimes^r f \in L\left(\otimes^r E; \otimes^r F\right)$

$\mathfrak{T}^r \left(f \circ g\right) = \mathfrak{T}^r \left(f\right) \circ \mathfrak{T}^r \left(g\right) = \left(\otimes^r f\right) \circ \left(\otimes^r g\right)$

Theorem 412 *The s-tensorial power of dual vector spaces is a faithful contravariant functor $\mathfrak{T}_s : \mathfrak{V} \mapsto \mathfrak{V}$*

$\mathfrak{T}_s \left(E\right) = \otimes_s E = \otimes^s E^*$

$f \in L\left(E; F\right) : \mathfrak{T}_s \left(f\right) = \otimes_s f \in L\left(\otimes_s F^*; \otimes_s E^*\right)$

$\mathfrak{T}_s \left(f \circ g\right) = \mathfrak{T}_s \left(g\right) \circ \mathfrak{T}_s \left(f\right) = \left(\otimes_s g^*\right) \circ \left(\otimes_s f^*\right)$

Theorem 413 *The (r,c)-tensorial product of vector spaces is a faithful bifunctor : $\mathfrak{T}_s^r : \mathfrak{V}_E \mapsto \mathfrak{V}_E$*

The following functors are similarly defined:

the covariant functors $\mathfrak{T}_S^r : \mathfrak{V} \mapsto \mathfrak{V} :: \mathfrak{T}_S^r \left(E\right) = \odot^r E$ for symmetric r tensors

the covariant functors $\mathfrak{T}_A^r : \mathfrak{V} \mapsto \mathfrak{V} :: \mathfrak{T}_A^r \left(E\right) = \wedge^r E$ for antisymmetric r contravariant tensors

the contravariant functors $\mathfrak{T}_{As} : \mathfrak{V} \mapsto \mathfrak{V} :: \mathfrak{T}_{As} \left(E\right) = \wedge_s E$ for antisymmetric r covariant tensors

Theorem 414 *Let $\mathfrak{A}$ be the category of algebras over the field K. The functor $\mathfrak{T} : \mathfrak{V} \mapsto \mathfrak{A}$ is defined as :*

$\mathfrak{V}(E) = \otimes E = \sum_{r,s=0}^{\infty} \otimes_s^r E$

$\forall f \in L(E;F) : \mathfrak{T}(f) \in \hom(\otimes E; \otimes F) = L(\otimes E; \otimes F)$

is faithful : there is a unique map $\mathfrak{T}(f) \in L(\otimes E; \otimes F)$ such that :

$\forall u \in E, v \in F : \mathfrak{T}(f)(u \otimes v) = f(u) \otimes f(v)$

7.3.5 Invariant and equivariant tensors

These results are used in the part Fiber bundles.

Let E be a vector space, GL(E) the group of linear inversible endomorphisms, G a subgroup of GL(E).

The action of $g \in G$ on E is : $f(g) : E \to E$ and we have the dual action: $f^*(g) : E^* \to E^* :: f^*(g)\lambda = \lambda \circ f\left(g^{-1}\right)$

This action induces an action $F_s^r(g) : \otimes_s^r E \to \otimes_s^r E$ with $F_s^r(g) = (\otimes^r f(g)) \otimes \left(\otimes^s (f(g))^*\right)$

Invariant tensor

A tensor $T \in \otimes_s^r E$ is said to be invariant by G if : $\forall g \in G : F_s^r(g) T = T$

Definition 415 *The elementary invariant tensors of rank r of a finite dimensional vector space E are the tensors $T \in \otimes_r^r E$ with components : $T_{j_1..j_r}^{i_1..i_r} = \sum_{\sigma \in \mathfrak{S}(r)} C_\sigma \delta_{j_1}^{\sigma(i_1)} \delta_{j_2}^{\sigma(i_2)} ... \delta_{j_r}^{\sigma(i_r)}$*

Theorem 416 *Invariant tensor theorem (Kolar p.214): On a finite dimensional vector space E, any tensor $T \in \otimes_s^r E$ invariant by the action of GL(E) is zero if $r \neq s$. If r=s it is a linear combination of the elementary invariant tensors of rank r*

Theorem 417 *Weyl (Kolar p.265) : The linear space of all linear maps $\otimes^k \mathbb{R}^m \to \mathbb{R}$ invariant by the orthogonal group $O(\mathbb{R},m)$ is spanned by the elementary invariants tensors if k is even, and 0 if k is odd.*

Equivariant map

A map : $f : \otimes E \to \otimes E$ is said to be equivariant by the action of GL(E) if :

$\forall g \in G, T \in \otimes_s^r E : f(F_s^r(g) T) = F_s^r(g) f(T)$

Theorem 418 *(Kolar p.217) : all smooth GL(E) equivariant maps (not necessarily linear) :*

i) $\wedge^r E \to \wedge^r E$ are multiples of the identity

ii) $\otimes^r E \to \odot^r E$ are multiples of the symmetrizer

iii) $\otimes^r E \to \wedge^r E$ are multiples of the antisymmetrizer

iv) $\wedge^r E \to \otimes^r E$ or $\odot^r E \to \otimes^r E$ are multiples of the inclusion

7.3.6 Invariant polynomials

These results are used mainly in the Chern theory (Fiber Bundles part). They are located here because they can be useful for other applications.

Invariant maps

Definition 419 *Let E be a vector space on a field K, G a subgroup of GL(E),
f a map :$f : E^r \times E^{*s} \to K$ with r,s$\in \mathbb{N}$*

f is said to be invariant by G if :

$\forall g \in G, \forall (u_i)_{i=1..r} \in E, \forall (\lambda_j)_{j=1}^s \in E^* : f\left((gu_1, ..gu_r), \left(g^{-1}\lambda_1\right), ... \left(g^{-1}\lambda_s\right)\right) =
f(u_1, ..u_r, \lambda_1, ..\lambda_s)$

Theorem 420 *Tensor evaluation theorem (Kolar p.223) Let E a finite dimensional real space. A smooth map $f : E^r \times E^{*s} \to \mathbb{R}$ (not necessarily linear) is invariant by GL(E) iff $\exists F \in C_\infty (\mathbb{R}^{rs}; \mathbb{R})$ such that :*

$\forall (u_i)_{i=1..r} \in E, \forall (\lambda_j)_{j=1}^s \in E^* : f(u_1, ..u_r, \lambda_1, ..\lambda_s) = F(..\lambda_i(u_j)...)$ *for all*
i,j

As an application, all smooth GL(E) equivariant maps :
$f : E^r \times E^{*s} \to E$ are of the form :
$f(u_1, ..u_r, \lambda_1, ..\lambda_s) = \sum_{\beta=1}^k F_\beta(..\lambda_i(u_j)...)u_\beta$ where $F_\beta(..\lambda_i(u_j)...) \in C_\infty (\mathbb{R}^{rs}; \mathbb{R})$
$f : E^r \times E^{*s} \to E^*$ are of the form :
$f(u_1, ..u_r, \lambda_1, ..\lambda_s) = \sum_{\beta=1}^l F_\beta(..\lambda_i(u_j)...)\lambda_\beta$ where $F_\beta(..\lambda_i(u_j)...) \in C_\infty (\mathbb{R}^{rs}; \mathbb{R})$

Polynomials on a vector space

Definition 421 *A map $f : V \to W$ between two finite dimensional vector spaces on a field K is said to be polynomial if in its coordinate expression in any bases : $f_i(x_{1,...}x_j, ..x_m) = y_i$ are polynomials in the x_j.*

i) Then f reads : $f_i = f_{i0} + f_{i1} + .. + f_{ir}$ where f_{ik} , called a homogeneous component, is, for each component, a monomial of degree k in the components
: $f_{ik} = x_1^{\alpha_1} x_2^{\alpha_2} ..x_m^{\alpha_m}, \alpha_1 + ...\alpha_m = k$
ii) let $f : V \to K$ be a homogeneous polynomial map of degree r. The
polarization of f is defined as P_r such that $r!P_r(u_1, ..., u_r)$ is the coefficient of
$t_1 t_2 .. t_r$ in $f(t_1 u_1 + t_2 u_2 + ... + t_r u_r)$
P_r is a r linear symmetric map : $P_r \in L^r(V; K)$
Conversely if P_r is a r linear symmetric map a homogeneous polynomial map
of degree r is defined with : f(u)=P_r(u,u,..,u)
iii) by the universal property of the tensor product, the r linear symmetric
map P_r induces a unique map : $\widehat{P}_r : \odot^r V \to K$ such that : $P_r(u_1, ..., u_r) =
\widehat{P}_r(u_1 \otimes ... \otimes u_r)$
iv) So if f is a polynomial map of degree r : $f : V \to K$ there is a linear map
: P:$\odot_{k=0}^{k=r} V \to K$ given by the sum of the linear maps $\widehat{P}_r$.

Invariant polynomial
Let V a finite dimensional vector space on a field K, G a group with action
on V : $\rho : G \to L(E; E)$

A map : $f : V \to K$ is said to be invariant by this action if : $\forall g \in G, \forall u \in V : f\left(\rho\left(g\right)u\right) = f\left(u\right)$

Similarly a map $f : V^r \to K$ is invariant if : $\forall g \in G, \forall u \in V : f\left(\rho\left(g\right)u_1, ..\rho\left(g\right)u_r\right) = f\left(u_1, .., u_r\right)$

A polynomial $f : V \to K$ is invariant iff each of its homogeneous components f_k is invariant

An invariant polynomial induces by polarizarion a r linear symmetric invariant map, and conversely a r linear, symmetric, invariant map induces an invariant polynomial.

Theorem 422 *(Kolar p.266) Let f:$\mathbb{R}\left(m\right) \to \mathbb{R}$ a polynomial map from the vector space $\mathbb{R}\left(m\right)$ of square mxm real matrices to $\mathbb{R}$ such that : f(OM)=M for any orthogonal matrix $O \in O\left(\mathbb{R}, m\right)$. Then there is a polynomial map F:$\mathbb{R}\left(m\right) \to \mathbb{R}$ such that : $f(M) = F(M^t M)$*

8 MATRICES

8.1 Operations with matrices

8.1.1 Definitions

Definition 423 *A rxc **matrix** over a field K is a table A of K scalars arranged in r rows and c columns, indexed as : a_{ij} i=1...r, j=1...c (the fist index is for row, the second is for columns).*

We will use also the tensor like indexes : a^i_j, up=row, low=column. When necessary a matrix is denoted within brackets : $A = [a_{ij}]$

When r=c we have the set of **square r-matrices** over K

Notation 424 *$K(r,c)$ is the set of rxc matrices over the field K.*
$K(r)$ is the set of square r-matrices over the field K

8.1.2 Basic operations

Addition and multiplication by a scalar

Theorem 425 *With addition and multiplication by a scalar the set K(r,c) is a vector space over K, with dimension rc.*

$A, B \in K(r,c) : A + B = [a_{ij} + b_{ij}]$
$A \in K(r,c), k \in K : kA = [ka_{ij}]$

Product of matrices

Definition 426 *The **product of matrices is the operation** : $K(c,s) \times K(c,s) \to K(r,s) :: AB = [\sum_{k=1}^{c} a_{ik} b_{kj}]$*

When defined the product distributes over addition and multiplication by a scalar and is associative :
$A(B + C) = AB + AC$
$A(kB) = kAB$
$(AB)C = A(BC)$
The product is not commutative.
The identity element for multiplication is the **identity matrix** : $I_r = [\delta_{ij}]$

Square matrices

Theorem 427 *With these operations the set K(r) of square r-matrices over K is a ring and a unital algebra over K.*

Definition 428 *The **commutator** of 2 matrices is :* $[A, B] = AB - BA$.

Theorem 429 *With the commutator as bracket K(r) is a Lie algebra.*

Notation 430 *K(r) is the group of square invertible (for the product) r-matrices.*

When a matrix has an inverse, denoted A^{-1} , it is unique and a right and left inverse : $AA^{-1} = A^{-1}A = I_r$ and $(AB)^{-1} = B^{-1}A^{-1}$

Diagonal

Definition 431 *The **diagonal** of a squared matrix A is the set of elements :* $\{a_{11}, a_{22}, ..., a_{rr}\}$

A square matrix is diagonal if all its elements $=0$ but for the diagonal.

A diagonal matrix is commonly denoted as $Diag\,(m_1, m_2, ...m_r)$ with $m_i = a_{ii}$

Remark : the diagonal is also called the "main diagonal", with reverse diagonal = the set of elements : $\{a_{r1}, a_{r-12}, ..., a_{1r}\}$

Theorem 432 *The set of diagonal matrices is a commutative subalgebra of K(r).*

A diagonal matrix is invertible if there is no zero on its diagonal.

Triangular matrices

Definition 433 *A **triangular** matrix is a square matrix A such that :* $a_{ij} = 0$ *whenever i>j . Also called upper triangular (the non zero elements are above the diagonal). A lower triangular matrix is such that A^t is upper triangular (the non zero elements are below the diagonal)*

8.1.3   Transpose

Definition 434 *The **transpose** of a matrix $A = [a_{ij}] \in K\,(r, c)$ is the matrix $A^t = [a_{ji}] \in K(c, r)$*

Rows and columns are permuted:

$$A = \begin{bmatrix} a_{11} & ... & a_{1c} \\ ... & & ... \\ a_{r1} & ... & a_{rc} \end{bmatrix} \rightarrow A^t = \begin{bmatrix} a_{11} & ... & a_{r1} \\ ... & & ... \\ a_{1c} & ... & a_{rc} \end{bmatrix}$$

Remark : there is also the old (and rarely used nowodays) notation ^tA

For $A, B \in K\,(r, c)\,, k, k' \in K$:

$(kA + k'B)^t = kA^t + k'B^t$

$(AB)^t = B^t A^t$

$(A_1 A_2...A_n)^t = A_n^t A_{n-1}^t...A_1^t$

$(A^t)^{-1} = (A^{-1})^t$

Definition 435 *A square matrix A is :*
   symmetric *if $A = A^t$*
   skew-symmetric *(or antisymmetric) if $A = -A^t$*
   orthogonal *if : $A^t = A^{-1}$*

Notation 436 *O(r,K) is the set of orthogonal matrix in K(r)*

   So $A \in O(r, K) \Rightarrow A^t = A^{-1}, AA^t = A^t A = I_r$
   Notice that O(r,K) is not an algebra: the sum of two orthogonal matrices is generally not orthogonal.

8.1.4 Adjoint

Definition 437 *The **adjoint** of a matrix $A = [a_{ij}] \in \mathbb{C}(r, c)$ is the matrix $A^* = [\overline{a}_{ji}] \in K(c, r)$*

   Rows and columns are permuted and the elements are conjugated :

$$A = \begin{bmatrix} a_{11} & ... & a_{1c} \\ ... & & ... \\ a_{r1} & ... & a_{rc} \end{bmatrix} \rightarrow A^* = \begin{bmatrix} \overline{a}_{11} & ... & \overline{a}_{r1} \\ ... & & ... \\ \overline{a}_{1c} & ... & \overline{a}_{rc} \end{bmatrix}$$

   Remark : the notation varies according to the authors
   For $A, B \in \mathbb{C}(r, c), k, k' \in K$:
   $(kA + k'B)^* = \overline{k}A^* + \overline{k'}B^*$
   $(AB)^* = B^*A^*$
   $(A_1 A_2 ... A_n)^* = A_n^* A_{n-1}^* ... A_1^t$
   $(A^*)^{-1} = (A^{-1})^*$

Definition 438 *A square matrix A is*
   hermitian *if $A = A^*$*
   skew-hermitian *if $A = -A^*$*
   unitary *if : $A^* = A^{-1}$*
   normal *if $AA^* = A^*A$*

Notation 439 *U(r) is the set of unitary matrices is a group denoted :*

   So $A \in U(r) \Rightarrow A^* = A^{-1}, AA^* = A^*A = I_r$
   U(r) is not an algebra: the sum of two unitary is generally not unitary

Theorem 440 *The real symmetric, real antisymmetric, real orthogonal, complex hermitian, complex antihermitian, unitary matrices are normal.*

   Normal matrices have many nice properties.
   Remark : $\mathbb{R}(c)$ is a subset of $\mathbb{C}(r)$. Matrices in $\mathbb{C}(r)$ with real elements are matrices in $\mathbb{R}(r)$. So hermitian becomes symmetric, skew-hermitian becomes skew-symmetric, unitary becomes orthogonal, normal becomes $AA^t = A^tA$. Any theorem for $\mathbb{C}(r)$ can be implemented for $\mathbb{R}(r)$ with the proper adjustments.

8.1.5   Trace

Definition 441 *The **trace** of a square matrix $A \in K(r)$ is the sum of its diagonal elements*

$Tr(A) = \sum_{i=1}^{r} a_{ii}$
It is the trace of the linear map whose matrix is A
$Tr : K(r) \to K$ is a linear map $Tr \in K(r)^*$
$Tr(A) = Tr(A^t)$
$Tr(A) = \overline{Tr(A^*)}$
$Tr(AB) = Tr(BA) \Rightarrow Tr(ABC) = Tr(BCA) = Tr(CAB)$
$Tr(A^{-1}) = (Tr(A))^{-1}$
$Tr(PAP^{-1}) = Tr(A)$
Tr(A)= sum of the eigenvalues of A
$Tr(A^k)$ = sum of its $(\text{eigenvalues})^k$
If A is symmetric and B skew-symmetric then Tr(AB)=0
$Tr([A,B]) = 0$ where $[A,B] = AB - BA$

Definition 442 *The Frobenius norm (also called the Hilbert-Schmidt norm) is the map : $K(r,c) \to \mathbb{R} :: Tr(AA^*) = Tr(A^*A)$*

Whenever $A \in \mathbb{C}(r,c) : AA^* \in \mathbb{C}(r,r), A^*A \in \mathbb{C}(c,c)$ are square matrix, so $Tr(AA^*)$ and $Tr(A^*A)$ are well defined
$Tr(AA^*) = \sum_{i=1}^{r} \left(\sum_{j=1}^{c} a_{ij}\overline{a}_{ij} \right) = \sum_{j=1}^{c} \left(\sum_{i=1}^{r} \overline{a}_{ij}a_{ij} \right) = \sum_{i=1}^{r} \sum_{j=1}^{c} |a_{ij}|^2$

8.1.6   Permutation matrices

Definition 443 *A **permutation matrix** is a square matrix $P \in K(r)$ which has on each row and column all elements =0 but one =1*

$\forall i,j : P_{ij} = 0$ but for one unique couple $(I,J) : P_{IJ} = 1$
It implies that $\forall i,j : \sum_c P_{ic} = \sum_r P_{rj} = 1$

Theorem 444 *The set P(K,r) of permutation matrices is a subgroup of the orthogonal matrices O(K,r).*

The right multiplication of a matrix A by a permutation matrix is a permutation of the rows of A

The left multiplication of a matrix A by a permutation matrix is a permutation of the columns of A

So given a permutation $\sigma \in \mathfrak{S}(r)$ of $(1,2,...r)$ the matrix : $S(\sigma) = P : [P_{ij}] = \delta_{\sigma(j)j}$ is a permutation matrix (remark : one can also take $P_{ij} = \delta_{i\sigma(i)}$ but it is less convenient) and this map : $S : \mathfrak{S}(r) \to P(K,r)$ is a group isomorphism : $P_{S(\sigma \circ \sigma')} = P_{S(\sigma)}P_{S(\sigma')}$

The identity matrix is the only diagonal permutation matrix.

As any permutation of a set can be decomposed in the product of transpositions, any permutation matrix can be decomposed in the product of elementary permutation matrices which transposes two columns (or two rows).

8.1.7 Determinant

Definition 445 *The **determinant** of a square matrix $A \in K(r)$ is the quantity :*

$$\det A = \sum_{\sigma \in \mathfrak{S}(r)} \varepsilon(\sigma) \, a_{1\sigma(1)} a_{2\sigma(2)}...a_{n\sigma(n)} = \sum_{\sigma \in \mathfrak{S}(r)} \epsilon(\sigma) \, a_{\sigma(1)1} a_{\sigma(2)2}...a_{\sigma(n)n}$$

$\det A^t = \det A$

$\det A^* = \overline{\det A}$ so the determinant of a Hermitian matrix is real

$\det(kA) = k^r \det A$ (Beware !)

$\det(AB) = \det(A)\det(B) = \det(BA)$

$\exists A^{-1} \Leftrightarrow \det A \neq 0$ and then $\det A^{-1} = (\det A)^{-1}$

The determinant of a permutation matrix is equal to the signature of the corresponding permutation

For $K = \mathbb{C}$ the determinant of a matrix is equal to the product of its eigen values

As the product of a matrix by a permutation matrix is the matrix with permuted rows or columns, the determinant of the matrix with permuted rows or columns is equal to the determinant of the matrix x the signature of the permutation.

The determinant of a triangular matrix is the product of the elements of its diagonal

Theorem 446 *Sylvester's determinant theorem : Let $A \in K(r,c)$, $B \in K(c,r)$, $X \in GL(K,r)$*

then : $\det(X + AB) = \det X \det(I_c + BX^{-1}A)$ *so with X=I:* $\det(I + AB) = \det(I_c + BA)$

Computation of a determinant : Determinant is the unique map : $D : K(r) \rightarrow K$ with the following properties :

a) For any permutation matrix P, $D(P) =$ signature of the corresponding permutation

b) $D(AP) = D(P)D(A) = D(A)D(P)$ where P is a permutation matrix

Moreover D has the following linear property : D(A')=kD(A)+D(A') where A' is (for any i) the matrix

$A = [A_1, A_2, ...A_r] \rightarrow A' = [A_1, A_2, ., A_{i-1}, B, A_{i+1}..A_r]$

where A_i is the i column of A, B is rx1 matrix and k a scalar

So for $A \in K(r)$ and A' the matrix obtained from A by adding to the row i a scalar multiple of another row i' :$\det A = \det A'$.

There is the same result with columns (but one cannot mix rows and columns in the same operation). This is the usual way to compute determinants, by gaussian elimination : by successive applications of the previous rules one strives to get a triangular matrix.

There are many results for the determinants of specific matrices. Many Internet sites offer results and software for the computation.

Definition 447 *The (i,j)* **minor** *of a square matrix $A=[a_{ij}] \in K(r)$ is the determinant of the (r-1,r-1) matrix denoted A_{ij} deduced from A by removing the row i and the column j.*

Theorem 448 $\det A = \sum_{i=1}^{r} (-1)^{i+j} a_{ij} \det A_{ij} = \sum_{j=1}^{r} (-1)^{i+j} a_{ij} \det A_{ij}$

The row i or the column j are arbitrary. It gives a systematic way to compute a determinant by a recursive calculus.

This formula is generalized in the **Laplace's development** :

For any sets of p ordered indices

$I = \{i_1, i_2, ...i_p\} \subset (1, 2, ...r), J = \{j_1, j_2, ...j_p\} \subset (1, 2, ...r)$

Let us denote $[A^c]_J^I$ the matrices deduced from A by removing all rows with indexes in I, and all columns with indexes in J

Let us denote $[A]_J^I$ the matrices deduced from A by keeping only the rows with indexes in I, and the columns with indexes in J

Then :

$\det A = \sum_{(j_1,...,j_p)} (-1)^{i_1+i_2+...+i_p+j_1+...+j_p} \left(\det [A]_{\{j_1,..,j_p\}}^{\{i_1,...i_p\}} \right) \left(\det [A^c]_{(j_1,...,j_p)}^{\{i_1,...i_p\}} \right)$

The cofactor of a square matrix $A=[a_{ij}] \in K(r)$ is the quantity $(-1)^{i+j} \det A_{ij}$ where $\det A_{ij}$ is the minor.

The matrix of cofactors is the matrix : $C(A) = \left[(-1)^{i+j} \det A_{ij} \right]$ and (cf Matrix cook book): $A^{-1} = \frac{1}{\det A} C(A)^t$ So

Theorem 449 *The elements $\left[A^{-1} \right]_{ij}$ of A^{-1} are given by the formula : $\left[A^{-1} \right]_{ij} = \frac{1}{\det A} (-1)^{i+j} \det [A_{ji}]$ where A_{ij} is the (r-1,r-1) matrix denoted A_{ij} deduced from A by removing the row i and the column j.*

Beware of the inverse order of indexes on the right hand side!

8.1.8   Kronecker's product

Also called tensorial product of matrices

For $A \in K(m, n), B(p, q)$, $C = A \otimes B \in K(mp, nq)$ is the matrix $[C_{ij}] = [a_{ij}] B$ built as follows : to each element $[a_{ij}]$ one associates one block equal to $[a_{ij}] B$

The useful relation is : $(A \otimes B) \times (C \otimes D) = AC \otimes BD$

Thus : $(A_1 \otimes ... \otimes A_p) \times (B_1 \otimes ...B_p) = A_1 B_1 \otimes ... \otimes A_p B_p$

If the matrices are square the Kronecker product of two symmetric matrices is still symmetric, the Kronecker product of two hermitian matrices is still hermitian.

8.2   Eigen values

There are two ways to see a matrix : as a vector in the vector space of matrices, and as the representation of a map in K^n.

A matrix in K(r,c) can be seen as tensor in $K^r \otimes (K^c)^*$ so a morphism in K(r,c) is a 4th order tensor. As this is not the most convenient way to

work, usually matrices are seen as representations of maps, either linear maps or bilinear forms.

8.2.1 Canonical isomorphims

1. The set K^n has an obvious n-dimensional vector space structure, with canonical basis $\varepsilon_i = (0, 0, .., 0, 1, 0, ..0)$

Vectors are represented as nx1 column matrices

$(K^n)^*$ has the basis $\varepsilon_i = (0, 0, .., 0, 1, 0, ..0)$ with vectors represented as 1xn row matrices

So the action of a form on a vector is given by : $[x] \in K^n, [\varpi] \in K^{n*}$: $\varpi(x) = [\varpi][x]$

2. To any matrix $A \in K(r, c)$ is associated a **linear map** $L_A \in L(K^c; K^r)$ with the obvious definition :

$[y] = A[x] : (r, 1) = (r, c)(c, 1)$ Beware of the dimensions!

The rank of A is the rank of L_A.

Similarly for the dual map $L_A^* \in L((K^r)^*; (K^c)^*)$: $[\mu] = [\lambda]A : (1, c) = (1, r)(r, c)$

So : $\forall \lambda \in (K^r)^*, x \in K^c : L_A^*(\lambda)(x) = [\lambda][a^*][x] = [\lambda]A[x] = \lambda \circ a(x) \Leftrightarrow [L_A^*] = A$

Warning ! The map : K(r,c)$\rightarrow L(K^c; K^r)$ is basis dependant. With another basis we would have another map. And the linear map L_A is represented by another matrix in another basis.

If r=c, in a change of basis $e_i = \sum_j P_i^j \varepsilon_j$ the new matrix of a is : $B = P^{-1}AP$. Conversely, for $A, B, P \in K(r)$ such that : $B = P^{-1}AP$ the matrices A and B are said to be similar : they represent the same linear map L_A.. Thus they have same determinant, rank, eigen values.

3. Similarly to each square matrix $A \in K(r)$ is associated a **bilinear form** B_A whose matrix is A in the canonical basis.

$A \in K(r) \rightarrow b \in L(K^r, K^r; K) :: B_A(x, y) = [y]^t A[x]$

and if K=$\mathbb{C}$ a **sequilinear form** B_A defined by : $A \in \mathbb{C}(r) \rightarrow B_A \in L(\mathbb{C}^r, \mathbb{C}^r; \mathbb{C}) :: B_A(x, y) = [y]^* A[x]$

B_A is symmetric (resp.skew symmetric, hermitian, skewhermitian) is A is symmetric (resp.skew symmetric, hermitian, skewhermitian)

To the unitary matrix I_r is associated the canonical bilinear form : $B_I(x, y) = \sum_{i=1}^r x_i y_i = [x]^t [y]$. The canonical basis is orthonormal. And the associated isomorphism $K^r \rightarrow K^{r*}$ is just passing from column vectors to rows vectors. With respect to this bilinear form the map a associated to a matrix A is orthogonal if A is orthogonal.

If K=$\mathbb{C}$, to the unitary matrix I_r is associated the canonical hermitian form : $B_I(x, x) = \sum_{i=1}^r \overline{x}_i y_i = [x]^* [y]$.With respect to this hermitian form the map a associated to a matrix A is unitary if A is unitary.

Remark : the property for a matrix to be symmetric (or hermitian) is not linked to the associated linear map, but to the associated bilinear or sesquilinear map. It is easy to check that if a linear map is represented by a symmetric matrix in a basis, this property is not conserved in a change of basis.

4. Warning ! A matrix in $\mathbb{R}(r)$ can be considered as a matrix in $\mathbb{C}(r)$ with real elements. As a matrix A in $\mathbb{R}(r)$ is associated a linear map $L_A \in L(\mathbb{R}^r; \mathbb{R}^r)$. As a matrix in $\mathbb{C}(r)$ is associated $M_A \in L(\mathbb{C}^r; \mathbb{C}^r)$ which is the complexified of the map L_A in the complexified of $\mathbb{R}^r$. L_A and M_A have same value for real vectors, and same matrix. It works only with the classic complexification (see complex vector spaces), and not with complex structure.

5. Definite positive matrix

Definition 450 *A matrix $A \in \mathbb{R}(r)$ is **definite positive** if $\forall [x] \neq 0 : [x]^t A [x] > 0$*

An hermitian matrix A is definite positive if $\forall [x] \neq 0 : [x]^* A [x] > 0$

8.2.2 Eigen values

Definition 451 *The **eigen values** λ of a square matrix $A \in K(r)$ are the eigen values of its associated linear map $L_A \in L(K^r; K^r)$*

So there is the equation : $A[x] = \lambda[x]$ and the vectors $[x] \in K^r$ meeting this relation are the **eigen vectors** of A with respect to λ

Definition 452 *The **characteristic equation** of a matrix $A \in K(r)$ is the polynomial equation of degree r over K in λ :*
$\det(A - \lambda I_r) = 0$ *reads :* $\sum_{i=0}^r \lambda^i P_i = 0$

$A[x] = \lambda[x]$ is a set of r linear equations with respect to x, so the eigen values of A are such the solutions of $\det(A - \lambda I_r) = 0$

The coefficient of degree 0 is just detA : $P_0 = \det A$

If the field K is algebraically closed then this equation has always a solution. So matrices in $\mathbb{R}(r)$ can have no (real) eigen value and matrices in $\mathbb{C}(r)$ have r eigen values (possibly identical). And similarly the associated real linear maps can have no (real) eigen value and complex linear maps have r eigen values (possibly identical)

As any $A \in \mathbb{R}(r)$ can be considered as the same matrix (with real elements) in $\mathbb{C}(r)$ it has always r eigen values (possibly complex and identical) and the corresponding eigen vectors can have complex components in $\mathbb{C}^r$. These eigen values and eigen vectors are associated to the complexified M_A of the real linear map L_A and not to L_A.

The matrix has no zero eigen value iff the associated linear form is injective. The associated bilinear form is non degenerate iff there is no zero eigen value, and definite positive iff all the eigen values are >0 .

If all eigen values are real the (non ordered) sequence of signs of the eigen values is the **signature** of the matrix.

Theorem 453 *Hamilton-Cayley's Theorem: Any square matrix is a solution of its characteristic equation :* $\sum_{i=0}^r A^i P_i = 0$

The following are used very often:

Theorem 454 *Any symmetric matrix $A \in \mathbb{R}(r)$ has real eigen values*
Any hermitian matrix $A \in \mathbb{C}(r)$ has real eigen values

8.2.3   Diagonalization

The eigen spaces E_λ (set of eigen vectors corresponding to the same eigen value λ) are independant. Let be dim $E_\lambda = d_\lambda$ so $\sum_\lambda d_\lambda \leq r$

The matrix A is said to be **diagonalizable** iff $\sum_\lambda d_\lambda = r$. If it is so $K^r = \oplus_\lambda E_\lambda$ and it is possible to find a basis $(e_i)_{i=1}^r$ of K^r such that the linear map a associated with A is expressed in a diagonal matrix D=Diag$(\lambda_1, ... \lambda_r)$ (several λ can be identical).

With a basis of each vector subspace (e_λ), together they constitute a basis for K^r and : $u \in E_\lambda \Leftrightarrow L_A u = \lambda u$

Matrices are not necessarily diagonalizable.

Let be m_λ the order of multiplicity of λ in the characteristic equation. The matrix A is diagonalizable iff $m_\lambda = d_\lambda$. Thus if there are r distincts eigen values the matrix is diagonalizable.

Let be P the matrix whose columns are the components of the eigen vectors (in the canonical basis), P is also the matrix of the new basis : $e_i = \sum_j P_i^j \varepsilon_j$ and the new matrix of L_A is : $D = P^{-1}AP \Leftrightarrow A = PDP^{-1}$. The basis (e_i) is not unique : the vectors e_i are defined up to a scalar, and the vectors can be permuted.

Let be $A, P, Q, D, D' \in K(r), D, D'$ diagonal such that : $A = PDP^{-1} = QD'Q^{-1}$ then there is a permutation matrix π such that : $D' = \pi D \pi^t ; P = Q\pi$

Theorem 455 *Normal matrices admit a complex diagonalization*

Proof. Let K=$\mathbb{C}$. the Schur decomposition theorem states that any matrix A can be written as : $A = U^*TU$ where U is unitary $(UU^* = I)$ and T is a triangular matrix whose diagonal elements are the eigen values of A.

T is a diagonal matrix iff A is normal : AA*=A*A. So A can be written as : $A = U^*DU$ iff it is normal. The diagonal elements are the eigen values of A. ■

Hermitian matrices and real symmetric matrices are normal, they can be written as :

real symmetric : $A = P^t D P$ with P orthogonal : $P^t P = PP^t = I$. The eigen vectors are real and orthogonal for the canonical bilinear form

hermitian : $A = U^*DU$ (also called Takagi's decomposition)

8.3   Matrix calculus

There are many theorems about matrices. Here are the most commonly used.

8.3.1   Decomposition

The decomposition of a matrix A is a way to write A as the product of matrices with interesting properties.

Singular values

Theorem 456 *Any matrix $A \in K(r,c)$ can be written as : $A = VDU$ where V,U are unitary and D is the matrix :*

$$D = \begin{bmatrix} diag(\sqrt{\lambda_1},...\sqrt{\lambda_c}) \\ 0_{(r-c) \times c} \end{bmatrix}_{r \times c}$$

with λ_i the eigen values of A^*A
(as A^*A is hermitian its eigen values are real, and it is easy to check that $\lambda_i \geq 0$)
If K=$\mathbb{R}$ the theorem stands and V,U are orthogonal.
Remark : the theorem is based on the study of the eigen values and vectors of A*A and AA*.

Definition 457 *A scalar $\lambda \in K$ is a **singular value** for $A \in K(r,c)$ if there are vectors $[x] \in K^c, [y] \in K^r$ such that :*
$A[x] = \lambda[y]$ and $A^*[y] = \lambda[x]$

Jordan's decomposition

Theorem 458 *Any matrix $A \in K(r)$ can be uniquely written as : $A = S + N$ where S is diagonalizable, N is nilpotent (there is $k \in \mathbb{N} : N^k = 0$),and $SN = NS$ .Furthermore there is a polynomial such that : $S = \sum_{j=1}^{p} a_j A^j$*

Schur's decomposition

Theorem 459 *Any matrix $A \in K(r)$ can be written as : $A = U^*TU$ where U is unitary ($UU^* = I$) and T is a triangular matrix whose diagonal elements are the eigen values of A.*

T is a diagonal matrix iff A is normal : AA*=A*A. So A can be written as : $A = U^*DU$ iff it is normal (see Diagonalization).

With triangular matrices

Theorem 460 *Lu decomposition : Any square matrix $A \in K(r)$ can be written : $A = LU$ with L lower triangular and U upper triangular*

Theorem 461 *QR decomposition : any matrix $A \in \mathbb{R}(r,c)$ can be written : $A = QR$ with Q orthogonal and R upper triangular*

Theorem 462 *Cholesky decomposition : any symmetric positive definite matrix can be uniquely written $A = T^tT$ where T is triangular with positive diagonal entries*

Spectral decomposition

Let be $\lambda_k, k = 1...p$ the eigen values of A$\in \mathbb{C}(n)$ with multiplicity m_k, A diagonalizable with $A = PDP^{-1}$

B_k the matrix deduced from D by putting 1 for all diagonal terms related to λ_k and 0 for all the others and $E_k = PB_kP^{-1}$

Then $A = \sum_{k=1}^{p} \lambda_k E_k$ and :

$E_j^2 = E_j; E_i E_j = 0, i \neq j$

$\sum_{k=1}^{p} E_k = I$

rank $E_k = m_k$

$(\lambda_k I - A) E_k = 0$

$A^{-1} = \sum \lambda_k^{-1} E_k$

A matrix commutes with A iff it commutes with each E_k

If A is normal then the E_k are hermitian

Other

Theorem 463 *Any non singular real matrix $A \in \mathbb{R}(r)$ can be written A=CP (or A=PC) where C is symmetric definite positive and P orthogonal*

8.3.2 Block calculus

Quite often matrix calculi can be done more easily by considering sub-matrices, called blocks.

The basic identities are :

$$\begin{bmatrix} A_{np} & B_{nq} \\ C_{rp} & D_{rq} \end{bmatrix} \begin{bmatrix} A'_{pn'} & B'_{pp'} \\ C'_{qn'} & D'_{qp'} \end{bmatrix} = \begin{bmatrix} A_{np}A'_{pn'} + B_{nq}C'_{qn'} & A_{np}B'_{pp'} + B_{nq}D'_{qp'} \\ C_{rp}A'_{pn'} + D_{rq}C'_{qn'} & C_{rp}B'_{pp'} + D_{rq}D'_{qp'} \end{bmatrix}$$

so we get nicer results if some of the blocks are 0.

Let be M=$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$; $A(m,m); B(m,n); C(n,m); D(n,n)$

Then :

$\det M = \det(A)\det(D - CA^{-1}B) = \det(D)\det(A - BD^{-1}C)$

If A=I,D=I:det(M)=det$(I_{mm} - BC) = det(I_{nn} - CB)$

$P[n,n] = D - CA^{-1}B$ and $Q[m,m] = A - BD^{-1}C$ are respectively the Schur Complements of A and D in M.

$$M^{-1} = \begin{bmatrix} Q^{-1} & -Q^{-1}BD^{-1} \\ -D^{-1}CQ^{-1} & D^{-1}\left(I + CQ^{-1}BD^{-1}\right) \end{bmatrix}$$

8.3.3 Complex and real matrices

Any matrix $A \in \mathbb{C}(r,c)$ can be written as : $A = \operatorname{Re} A + i \operatorname{Im} A$ where $\operatorname{Re} A, \operatorname{Im} A \in \mathbb{R}(r,c)$

For square matrices $M \in \mathbb{C}(n)$ it can be useful to introduce :

$$Z(M) = \begin{bmatrix} \operatorname{Re} M & -\operatorname{Im} M \\ \operatorname{Im} M & \operatorname{Re} M \end{bmatrix} \in \mathbb{R}(2n)$$

It is the real representation of GL(n,$\mathbb{C}$) in GL(2n;$\mathbb{R}$)
and :
Z(MN)=Z(M)Z(N)
Z(M*)=Z(M)*
TrZ(M)=2ReTrM
detZ(M)=|detM|^2

8.3.4 Pauli's matrices

They are (with some differences according to authors and usages) the matrices in $\mathbb{C}(2)$:

$$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}; \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix};$$

the multiplication tables are:
$\sigma_i \sigma_j + \sigma_j \sigma_i = 2\delta_{ij}.\sigma_0 : i,j = 1,2,3$ that is :
$\sigma_1 \sigma_2 = i\sigma_3$
$\sigma_2 \sigma_3 = i\sigma_1$
$\sigma_3 \sigma_1 = i\sigma_2$

$$\sigma_0 \sigma_i \sigma_j = \sigma_i \sigma_j = \begin{bmatrix} \sigma_i \sigma_j & 0 & 1 & 2 & 3 \\ 0 & \sigma_0 & \sigma_1 & \sigma_2 & \sigma_3 \\ 1 & \sigma_1 & \sigma_0 & i\sigma_3 & -i\sigma_2 \\ 2 & \sigma_2 & -i\sigma_3 & \sigma_0 & i\sigma_1 \\ 3 & \sigma_3 & i\sigma_2 & -i\sigma_1 & \sigma_0 \end{bmatrix};$$

$$\sigma_1 \sigma_i \sigma_j = \begin{bmatrix} \sigma_i \sigma_j & 0 & 1 & 2 & 3 \\ 0 & \sigma_1 & \sigma_0 & i\sigma_3 & -i\sigma_2 \\ 1 & \sigma_0 & \sigma_1 & \sigma_2 & \sigma_3 \\ 2 & i\sigma_3 & -\sigma_2 & \sigma_1 & i\sigma_0 \\ 3 & -i\sigma_2 & -\sigma_3 & -i\sigma_0 & \sigma_1 \end{bmatrix}$$

$$\sigma_2 \sigma_i \sigma_j = \begin{bmatrix} \sigma_i \sigma_j & 0 & 1 & 2 & 3 \\ 0 & \sigma_2 & -i\sigma_3 & \sigma_0 & i\sigma_1 \\ 1 & -i\sigma_3 & \sigma_2 & -\sigma_1 & -i\sigma_0 \\ 2 & \sigma_0 & -i\sigma_1 & \sigma_2 & \sigma_3 \\ 3 & i\sigma_1 & i\sigma_0 & -\sigma_3 & \sigma_2 \end{bmatrix};$$

$$\sigma_3 \sigma_i \sigma_j = \begin{bmatrix} \sigma_i \sigma_j & 0 & 1 & 2 & 3 \\ 0 & \sigma_3 & i\sigma_2 & -i\sigma_1 & \sigma_0 \\ 1 & i\sigma_2 & \sigma_3 & i\sigma_0 & -\sigma_1 \\ 2 & -i\sigma_1 & -i\sigma_0 & \sigma_3 & -\sigma_2 \\ 3 & \sigma_0 & \sigma_1 & \sigma_2 & \sigma_3 \end{bmatrix}$$

8.3.5 Matrix functions

We have to introduce some bits of analysis but it seems logical to put these results in this section.

$\mathbb{C}(r)$ is a finite dimensional vector space, thus a normed vector space and a Banach vector space (and a C* algebra).

All the noms are equivalent. The two most common are :

i) the Frobenius norm (also called the Hilbert-Schmidt norm): $\|A\|_{HS} = Tr(A^*A) = \sum_{ij} |a_{ij}|^2$

ii) the usual norm on $L(\mathbb{C}^n; \mathbb{C}^n) : \|A\|_2 = \inf_{\|u\|=1} \|Au\|$

$\|A\|_2 \le \|A\|_{HS} \le n \|A\|_2$

Exponential

Theorem 464 *The series :* $\exp A = \sum_0^\infty \frac{A^n}{n!}$ *converges always*

$\exp 0 = I$

$(\exp A)^{-1} = \exp(-A)$

$\exp(A) \exp(B) = \exp(A+B)$ iff AB=BA   Beware !

$(\exp A)^t = \exp(A^t)$

$(\exp A)^* = \exp(A^*)$

$\det(\exp A) = \exp(Tr(A))$

The map $t \in \mathbb{R} \to \exp(tA)$ defines a 1-parameter group. The map is differentiable and :

$\frac{d}{dt}(\exp tA)|_{t=\tau} = (\exp \tau A) A = A \exp \tau A$

$\frac{d}{dt}(\exp tA)|_{t=0} = A$

Conversely if $f : \mathbb{R}_+ \to \mathbb{C}(r)$ is a continuous homomorphism then $\exists A \in \mathbb{C}(r) : f(t) = \exp tA$

Warning ! The map $t \in \mathbb{R} \to \exp A(t)$ where the matrix A(t) depends on t has no simple derivative. We *do not* have

$\frac{d}{dt}(\exp A(t)) = A'(t) \exp A(t)$

Theorem 465 *(Taylor 1 p.19) Let A be a nxn complex matrix, [v] a nx1 matrix, then :*

$\forall t \in \mathbb{R} : (\exp t[A])[v] = \sum_{j=1}^n (\exp \lambda_j t)[w_j(t)]$

where : λ_j *are the eigen values of A,* $[w_j(t)]$ *is a polynomial in t, valued in* $\mathbb{C}(n,1)$

If A is diagonalizable then the $[w_j(t)] = Cte$

Theorem 466 *Integral formulation: If all the eigen value of A are in the open disc $|z| < r$ then $\exp A = \dfrac{1}{2i\pi} \int_C (zI - A)^{-1} e^z dz$ with C any closed curve around the origin and included in the disc*

The inverse function of exp is the logarithm : $\exp\left(\log\left((A)\right)\right) = A$. It is usally an multivalued function (as for the complex numbers).

$\log(BAB^{-1}) = B(\log A)B^{-1}$

$\log(A^{-1}) = -\log A$

If A has no zero or negative eigen values : $\log A = \int_{-\infty}^0 [(s-A)^{-1} - (s-1)^{-1}]ds$

Cartan's decomposition : Any invertible matrix $A \in \mathbb{C}(r)$ can be uniquely written : $A = P \exp Q$ with :

$P = A \exp(-Q); P^t P = I$

$Q = \dfrac{1}{2} \log(A^t A); Q^t = Q$

P,Q are real if A is real

Analytic functions

Theorem 467 *Let $f : \mathbb{C} \to \mathbb{C}$ any holomorphic function on an open disc $|z| < r$ then : $f(z) = \sum_{n=0}^\infty a_n z^n$ and the series : $f(A) = \sum_{n=0}^\infty a_n A^n$ converges for $\|A\| < r$*

With the Cauchy's integral formula, for any closed curve C circling x and contained within the disc, it holds:

$f(x) = \frac{1}{2i\pi} \int_C \frac{f(z)}{z-x} dz$ then : $f(A) = \frac{1}{2i\pi} \int_C f(z)(zI - A)^{-1} dz$ where C is any closed curve enclosing all the eigen values of A and contained within the disc

If $\|A\| < 1 : \sum_{p=0}^\infty (-1)^p A^p = (I + A)^{-1}$

Derivative

They are useful formulas for the derivative of functions of a matrix depending on a variable.

1. Determinant:

Theorem 468 *let $A = [a_{ij}] \in \mathbb{R}(n)$, then $\dfrac{d \det A}{d a_{ij}} = (-1)^{i+j} \det \left[A_{\{1...n\backslash j\}}^{\{1...n\backslash i\}} \right] = \left[A^{-1} \right]_i^j \det A$*

Proof. we have $\left[A^{-1} \right]_{ij} = \frac{1}{\det A}(-1)^{i+j} \det [A_{ji}]$ where $\left[A^{-1} \right]_{ij}$ is the element of A^{-1} and $\det[A_{ji}]$ the minor. ∎

Beware reversed indices!

Theorem 469 *If $\mathbb{R} \to \mathbb{R}(n) :: A(x) = [a_{ij}(x)]$, A invertible then $\dfrac{d \det A}{dx} = (\det A) Tr \left(\frac{dA}{dx} \left(A^{-1} \right) \right)$*

Proof. Schur's decomposition : $A = UTU^*, UU^* = I, T$ triangular

let be : $A' = U'TU^* + UT'U^* + UT\left(U^*\right)'$

the derivative of U : $\left[u_j^i(x)\right] = U^* \to \left(U^*\right)' = \overline{u}_i^j(x)' \to \left(U^*\right)' = \left(U'\right)^*$

$A'A^{-1} = U'TU^*UT^{-1}U^* + UT'U^*UT^{-1}U^* + UT\left(U^*\right)'UT^{-1}U^*$

$= U'U^* + UT'T^{-1}U^* + UT\left(U^*\right)'UT^{-1}U^*$

$Tr(A'A^{-1}) = Tr\left(U'U^*\right) + Tr\left(T'T^{-1}\right) + Tr\left(\left(UT\left(U^*\right)'\right)\left(UT^{-1}U^*\right)\right)$

$Tr\left(\left(UT\left(U^*\right)'\right)\left(UT^{-1}U^*\right)\right) = Tr\left(\left(UT^{-1}U^*\right)\left(UT\left(U^*\right)'\right)\right) = Tr\left(U\left(U^*\right)'\right)$

$UU^* = I \Rightarrow U'U^* + U\left(U^*\right)' = 0$

$Tr(A'A^{-1}) = Tr\left(T'T^{-1}\right)$

$\Theta = T^{-1}$ is triangular with diagonal such that : $\theta_k^i t_j^k = \delta_j^i \Rightarrow \theta_k^i t_i^k = 1 = \sum_{k=i}^n \theta_k^i t_i^k = \theta_i^i t_i^k$

so $\theta_i^i = 1/$eigen values of A

$$Tr(A'A^{-1}) = Tr\left(T'T^{-1}\right) = \sum_{i=1}^n \frac{\lambda_i'}{\lambda_i} = \sum_i \left(\ln \lambda_i\right)' = \left(\sum_i \ln \lambda_i\right)' = \left(\ln \prod_i \lambda_i\right)' =$$

$\left(\ln \det A\right)'$ ∎

2. Inverse:

Theorem 470 *If* $K = [k_{pq}] \in \mathbb{R}(n)$*, is an invertible matrix, then* $\frac{dk_{pq}}{dj_{rs}} = -k_{pr}k_{sq}$ *with* $J = K^{-1} = [j_{pq}]$

Proof. Use : $K_\lambda^\gamma J_\mu^\lambda = \delta_\mu^\gamma$

$\Rightarrow \left(\frac{\partial}{\partial J_\beta^\alpha}K_\lambda^\gamma\right) J_\mu^\lambda + K_\lambda^\gamma \left(\frac{\partial}{\partial J_\beta^\alpha}J_\mu^\lambda\right) = 0 = \left(\frac{\partial}{\partial J_\beta^\alpha}K_\lambda^\gamma\right) J_\mu^\lambda + K_\lambda^\gamma \delta_\alpha^\lambda \delta_\beta^\mu = \left(\frac{\partial}{\partial J_\beta^\alpha}K_\lambda^\gamma\right) J_\mu^\lambda + K_\alpha^\gamma \delta_\beta^\mu$

$0 = \left(\frac{\partial}{\partial J_\beta^\alpha}K_\lambda^\gamma\right) J_\mu^\lambda K_\nu^\mu + K_\alpha^\gamma \delta_\beta^\mu K_\nu^\mu = \left(\frac{\partial}{\partial J_\beta^\alpha}K_\nu^\gamma\right) + K_\alpha^\gamma K_\nu^\beta$

$\Rightarrow \frac{\partial}{\partial J_l^k}K_j^i = -K_k^i K_j^l$ ∎

As $\mathbb{C}(r)$ is a normed algebra the derivative with respect to a matrix (and not only with respect to its elements) is defined :

$\varphi : \mathbb{C}(r) \to \mathbb{C}(r) :: \varphi(A) = (I_r + A)^{-1}$ then $\frac{d\varphi}{dA} = -A$

Matrices of $SO(\mathbb{R}, p, q)$

(See also Lie groups - classical groups)

These matrices are of some importance in physics, because the Lorentz group of Relativity is just $SO(\mathbb{R}, 3, 1)$.

$SO(\mathbb{R},p,q)$ is the group of nxn real matrices with n=p+q such that :

detM = 1

$A^t [I_{p,q}] A = I_{n \times n}$ where $[I_{p,q}] = \begin{bmatrix} I_{p \times p} & 0 \\ 0 & -I_{q \times q} \end{bmatrix}$

Any matrix of $SO(\mathbb{R}, p, q)$ has a Cartan decomposition, so can be uniquely written as :

$A = [\exp p][\exp l]$ with $[p] = \begin{bmatrix} 0 & P_{p \times q} \\ P_{q \times p}^t & 0 \end{bmatrix}, l = \begin{bmatrix} M_{p \times p} & 0 \\ 0 & N_{q \times q} \end{bmatrix}, M = -M^t, N = -N^t$

(or as $A = [\exp l'][\exp p']$ with similar p',l' matrices).

The matrix $[l]$ is block diagonal antisymmetric.

This theorem is new.

Theorem 471 $\exp p = \begin{bmatrix} I_p & 0 \\ 0 & I_q \end{bmatrix} + \begin{bmatrix} H\left(\cosh D - I_q\right)H^t & H(\sinh D)U^t \\ U(\sinh D)H^t & U(\cosh D - I_q)U^t \end{bmatrix}$
with $H_{p\times q}$ such that : $H^t H = I_q, P = HDU^t$ where D is a real diagonal qxq matrix and U is a $q\times q$ real orthogonal matrix.

Proof. We assume that p > q

The demonstration is based upon the decomposition of $[P]_{p\times q}$ using the singular values decomposition. P reads :

$P = VQU^t$ where :

$Q = \begin{bmatrix} D_{q\times q} \\ 0_{(p-q)\times q} \end{bmatrix}_{p\times q} ; D = diag(d_k)_{k=1...q}; d_k \geq 0$

$[V]_{p\times p}, [U]_{q\times q}$ are orthogonal

Thus : $PP^t = V \begin{bmatrix} D^2 \\ & 0_{(p-q)\times q} \end{bmatrix} V^t; P^t P = UD^2 U^t$

The eigen values of PP^t are $\{d_1^2, ...d_q^2, 0, .., 0\}$ and of $P^t P : \{d_1^2, ...d_q^2\}$. The decomposition is not unique.

Notice that we are free to choose the sign of d_k, the choice $d_k \geq 0$ is just a convenience.

So :

$[p] = \begin{bmatrix} 0 & P \\ P^t & 0 \end{bmatrix} = \begin{bmatrix} V & 0 \\ 0 & U \end{bmatrix} \begin{bmatrix} 0 & Q \\ Q^t & 0 \end{bmatrix} \begin{bmatrix} V^t & 0 \\ 0 & U^t \end{bmatrix} = [k] \begin{bmatrix} 0 & Q \\ Q^t & 0 \end{bmatrix} [k]^t$

with : $k = \begin{bmatrix} V & 0 \\ 0 & U \end{bmatrix} : [k][k]^t = I_{p\times p}$

and :

$\exp [p] = [k] \left(\exp \begin{bmatrix} 0 & Q \\ Q^t & 0 \end{bmatrix} \right) [k]^t$

$\begin{bmatrix} 0 & Q \\ Q^t & 0 \end{bmatrix}^{2m} = \begin{bmatrix} D^{2m} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & D^{2m} \end{bmatrix} ; m > 0$

$\begin{bmatrix} 0 & Q \\ Q^t & 0 \end{bmatrix}^{2m+1} = \begin{bmatrix} 0 & 0 & D^{2m+1} \\ 0 & 0 & 0 \\ D^{2m+1} & 0 & 0 \end{bmatrix}$

thus :

$\exp \begin{bmatrix} 0 & Q \\ Q^t & 0 \end{bmatrix} = I_{p+q} + \sum_{m=0}^{\infty} \frac{1}{(2m+1)!} \begin{bmatrix} 0 & 0 & D^{2m+1} \\ 0 & 0 & 0 \\ D^{2m+1} & 0 & 0 \end{bmatrix} + \sum_{m=1}^{\infty} \frac{1}{(2m)!} \begin{bmatrix} D^{2m} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & D^{2m} \end{bmatrix}$

$= I_{p+q} + \begin{bmatrix} 0 & 0 & \sinh D \\ 0 & 0 & 0 \\ \sinh D & 0 & 0 \end{bmatrix} + \begin{bmatrix} \cosh D & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \cosh D \end{bmatrix} - \begin{bmatrix} I_q & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & I_q \end{bmatrix}$

with : $\cosh D = diag(\cosh d_k); \sinh D = diag(\sinh d_k)$

And :

$\exp p = \begin{bmatrix} V & 0 \\ 0 & U \end{bmatrix} \begin{bmatrix} \cosh D & 0 & \sinh D \\ 0 & I_{p-q} & 0 \\ \sinh D & 0 & \cosh D \end{bmatrix} \begin{bmatrix} V^t & 0 \\ 0 & U^t \end{bmatrix}$

In order to have some unique decomposition write :

$$\exp p = \begin{bmatrix} V \begin{bmatrix} \cosh D & 0 \\ 0 & I_{p-q} \end{bmatrix} V^t & V \begin{bmatrix} \sinh D \\ 0 \end{bmatrix} U^t \\ U \begin{bmatrix} \sinh D & 0 \end{bmatrix} V^t & U(\cosh D)U^t \end{bmatrix}$$

Thus with the block matrices V_1 (q,q) and V_3 (p-q,q)

$$V = \begin{bmatrix} V_1 & V_2 \\ V_3 & V_4 \end{bmatrix} \in O(\mathbb{R}, p)$$

$$V^t V = V V^t = I_p$$

$$\begin{bmatrix} V_1 V_1^t + V_2 V_2^t & V_1 V_3^t + V_2 V_4^t \\ V_3 V_1^t + V_4 V_2^t & V_3 V_3^t + V_4 V_4^t \end{bmatrix} = \begin{bmatrix} V_1^t V_1 + V_3^t V_3 & V_2^t V_1 + V_4^t V_3 \\ V_1^t V_2 + V_3^t V_4 & V_2^t V_2 + V_4^t V_4 \end{bmatrix} = \begin{bmatrix} I_q & 0 \\ 0 & I_{p-q} \end{bmatrix}$$

So :

$$V \begin{bmatrix} \cosh D + I_q & 0 \\ 0 & I_{p-q} \end{bmatrix} V^t = \begin{bmatrix} V_2 V_2^t + V_1(\cosh D)V_1^t & V_2 V_4^t + V_1(\cosh D)V_3^t \\ V_4 V_2^t + V_3(\cosh D)V_1^t & V_4 V_4^t + V_3(\cosh D)V_3^t \end{bmatrix}$$

$$= I_p + \begin{bmatrix} V_1 \\ V_3 \end{bmatrix} (\cosh D - I_q) \begin{bmatrix} V_1^t & V_3^t \end{bmatrix}$$

$$V \begin{bmatrix} \sinh D \\ 0 \end{bmatrix} U^t = \begin{bmatrix} V_1(\sinh D)U^t \\ V_3(\sinh D)U^t \end{bmatrix} = \begin{bmatrix} V_1 \\ V_3 \end{bmatrix} (\sinh D)U^t$$

$$U \begin{bmatrix} \sinh D & 0 \end{bmatrix} V^t = U(\sinh D) \begin{bmatrix} V_1^t & V_3^t \end{bmatrix}$$

$$\exp p = \begin{bmatrix} I_p + \begin{bmatrix} V_1 \\ V_3 \end{bmatrix} (\cosh D - I_q) \begin{bmatrix} V_1^t & V_3^t \end{bmatrix} & \begin{bmatrix} V_1 \\ V_3 \end{bmatrix} (\sinh D)U^t \\ U(\sinh D) \begin{bmatrix} V_1^t & V_3^t \end{bmatrix} & U(\cosh D)U^t \end{bmatrix}$$

Let us denote $H = \begin{bmatrix} V_1 \\ V_3 \end{bmatrix}$

H is a pxq matrix with rank q : indeed if not the matrix V would not be regular. Moreover :

$$V^t V = V V^t = I_p \Rightarrow V_1^t V_1 + V_3^t V_3 = I_q \Leftrightarrow H^t H = I_q$$

And :

$$\exp p = \begin{bmatrix} I_p + H\left(\cosh D - I_q\right) H^t & H(\sinh D)U^t \\ U(\sinh D)H^t & U(\cosh D)U^t \end{bmatrix}$$

The number of parameters are here just pq and as the Cartan decomposition is a diffeomorphism the decomposition is unique.

H, D and U are related to P and p by :

$$P = V \begin{bmatrix} D \\ 0 \end{bmatrix} U^t = HDU^t$$

$$p = \begin{bmatrix} 0 & P \\ P^t & 0 \end{bmatrix} = \begin{bmatrix} H & 0 \\ 0 & U \end{bmatrix}_{(p+q,2q)} \begin{bmatrix} 0 & D \\ D & 0 \end{bmatrix}_{(2q,2q)} \begin{bmatrix} H & 0 \\ 0 & U \end{bmatrix}^t_{(2q,p+q)} \quad \blacksquare$$

With this decomposition it is easy to compute the powers of exp(p)

$$k \in Z : (\exp p)^k = \exp(kp) = \begin{bmatrix} I_p + H\left(\cosh kD - I_q\right) H^t & H(\sinh kD)U^t \\ U(\sinh kD)H^t & U(\cosh kD)U^t \end{bmatrix}$$

Notice that : $\exp(kp) = \exp\left(\begin{bmatrix} 0 & kP \\ kP^t & 0 \end{bmatrix}\right)$

so with the same singular values decomposition the matrix D' :

$$(kP)^t (kP) = D'^2 = k^2 D,$$

$$kP = (\tfrac{k}{|k|}V)D'U^t = (\epsilon V)(|k| D) U^t$$

$$(\exp p)^k = \exp(kp) = \begin{bmatrix} I_p + H\left(\cosh kD - I_q\right)H^t & H(\sinh kD)U^t \\ U(\sinh kD)H^t & U(\cosh kD)U^t \end{bmatrix}$$

In particular with k -1 :

$$(\exp p)^{-1} = \begin{bmatrix} I_p + H\left(\cosh D\right)H^t - HH^t & -H\left(\sinh D\right)U^t \\ -U\left(\sinh D\right)H^t & U\left(\cosh D\right)U^t \end{bmatrix})$$

For the Lorentz group the decomposition reads :

H is a vector 3x1 matrix : $H^t H = 1$, D is a scalar, U=[1],

$$l = \begin{bmatrix} M_{3\times 3} & 0 \\ 0 & 0 \end{bmatrix}, M = -M^t \text{ thus } \exp l = \begin{bmatrix} R & 0 \\ 0 & 1 \end{bmatrix} \text{ where } R \in SO\left(\mathbb{R}, 3\right)$$

$$A \in SO\left(3, 1, \mathbb{R}\right) : A = \exp p \exp l = \begin{bmatrix} I_3 + \left(\cosh D - 1\right)HH^t & \left(\sinh D\right)H \\ \left(\sinh D\right)H^t & \cosh D \end{bmatrix} \begin{bmatrix} R & 0 \\ 0 & 1 \end{bmatrix}$$

9 CLIFFORD ALGEBRA

Mathematical objects such as "spinors" and spin representations are frequently met in physics. The great variety of definitions, sometimes clever but varying greatly and too focused on a pretense of simplicity, gives a confusing idea of this field. In fact the unifying concept which is the base of all these mathematical objects is the Clifford algebra. This is a special structure, involving a vector space, a symmetric bilinear form and a field, which is more than an algebra and distinct from a Lie algebra. It introduces a new operation - the product of vectors - which can be seen as disconcerting at first, but when the structure is built in a coherent way, step by step, we feel much more comfortable with all its uses in the other fields, such as representation theory of groups, fiber bundles and functional analysis. So we will proceed as usual in the more general settings, because it is no more difficult and underlines the key ideas which sustain the structure.

9.1 Main operations in a Clifford algebra

9.1.1 Definition of the Clifford algebra

This is the most general definition of a Clifford algebra.

Definition 472 *Let F be a vector space over the field K (of characteristic $\neq 2$) endowed with a symmetric bilinear non degenerate form g (valued in the field K). The **Clifford algebra** Cl(F,g) and the canonical map $\imath : F \to Cl(F,g)$ are defined by the following universal property : for any associative algebra A over K (with internal product x and unit e) and K-linear map $f : F \to A$ such that :*
  $\forall v, w \in F : f(v) \times f(w) + f(w) \times f(v) = 2g(v,w) \times e$
  there exists a unique algebra morphism : $\varphi : Cl(F,g) \to A$ such that $f = \varphi \circ \imath$

    The Clifford algebra includes the scalar K and the vectors F (so we identify $\imath(u)$ with $u \in F$ and $\imath(k)$ with $k \in K$)
    Remarks :
    i) There is also the definition $f(v) \times f(w) + f(w) \times f(v) + 2g(v,w) \times e = 0$ which sums up to take the opposite for g (careful about the signature which is important)
    ii) F can be a real or a complex vector space, but *g must be symmetric*, meaning that a hermitian sesquilinear form does not work.
    iii) It is common to define a Clifford algebra through a quadratic form : any quadratic form gives a bilinear symmetric form by polarization, and as a bilinear symmetric form is necessary for most of the applications, we can easily jump over this step.
    iv) This is an algebraic definition, which encompasses the case of infinite dimensional vector spaces. However, as usual when working with infinite dimensional vector space, additional structure over V should be required, V should be be a Banach vector space, the form g linear and consistent with the norm, so we would have a Hilbert space.

A definition is not a proof of existence. Happily :

Theorem 473 *There is always a Clifford algebra, isomorphic, as vector space, to the algebra ΛF of antisymmetric tensors with the exterior product.*

Proof. Cl(F,g)=$(\otimes T)/I(V,g)$ where I(V,g) is the two sided ideal generated by elements of the form $v \otimes v - g(v,v)1$

The isomorphism follows the determination of the bases (see below) ∎

9.1.2 Algebra structure

1. Internal product:

Definition 474 *The **internal product** of Cl(F,g) is denoted by a dot $\cdot$. It is such that : $\forall v, w \in F : v \cdot w + w \cdot v = 2g(v,w)$*

Theorem 475 *With this internal product $(Cl(F,g), \cdot)$ is a unital algebra on the field K, with unity element the scalar $1 \in K$*

Notice that a Clifford algebra is an algebra but is more than that because of this fundamental relation (valid only for vectors of F, not for any element of the Clifford algebra).

Two useful relations :

$\forall u, v \in F : u \cdot v \cdot u = g(u,u) v - 2g(u,v) u \in F$

Proof. $u \cdot v \cdot u = u \cdot (u \cdot v - 2g(u,v)) = g(u,u) v - 2g(u,v) u$ ∎

$e_p \cdot e_q \cdot e_i - e_i \cdot e_p \cdot e_q = 2(\eta_{iq} e_p - \eta_p e_q)$

Proof. $e_i \cdot e_p \cdot e_q = (-e_p \cdot e_i + 2\eta_{ip}) \cdot e_q = -e_p \cdot e_i \cdot e_q + 2\eta_{ip} e_q = -e_p \cdot (-e_q \cdot e_i + 2\eta_{iq}) + 2\eta_{ip} e_q$

$= e_p \cdot e_q \cdot e_i - 2\eta_{iq} e_p + 2\eta_{ip} e_q$

$e_p \cdot e_q \cdot e_i - e_i \cdot e_p \cdot e_q = 2(\eta_{iq} e_p - \eta_{ip} e_q)$ ∎

2. Homogeneous elements:

Definition 476 *The **homogeneous** elements of degree r of Cl(F,g) are elements which can be written as product of r vectors of F*

$w = u_1 \cdot u_2 ... \cdot u_r$. The homogeneous elements of degree n=dimF are called pseudoscalars (there are also many denominations for various degrees and dimensions, but they are only complications)

3. Basis of Cl(F,g):

Theorem 477 *(Fulton p.302) The set of elements :*

$\left\{ 1, e_{i_1} \cdot ... \cdot e_{i_k}, 1 \le i_1 < i_2 ... < i_k \le \dim F, k = 1...2^{\dim F} \right\}$ *where $(e_i)_{i=1}^{\dim F}$ is an orthonormal basis of F, is a basis of the Clifford algebra Cl(F,g) which is a vector space over K of dimCl(F,g)=$2^{\dim F}$*

Notice that the basis of Cl(F,g) must have the basis vector 1 to account for the scalars.

4. Fundamental identity

Theorem 478 *For an orthonormal basis* $(e_i) : e_i \cdot e_j + e_j \cdot e_i = 2\eta_{ij}$ *where* $\eta_{ij} = g(e_i, e_j) = 0, \pm 1$

so
$i \neq j : e_i \cdot e_j = -e_j \cdot e_i$
$e_i \cdot e_i = \pm 1$
If σ is a permutation of the *ordered* set of indices :

Theorem 479 $\{i_1, ..., i_n\} : e_{\sigma(i_1)} \cdot e_{\sigma(i_2)} .. \cdot e_{\sigma(i_r)} = \epsilon(\sigma) e_{i_1} \cdot e_{i_2} .. \cdot e_{i_r}$

Warning ! it works for orthogonal vectors, not for any vector and the indices must be different

A bilinear symmetric form is fully defined by an orthonormal basis. They will always be used in a Clifford algebra.

So any element of Cl(F,g) can be expressed as :
$w = \sum_{k=0}^{2^{\dim F}} \sum_{\{i_1,...,i_k\}} w_{\{i_1,...,i_k\}} e_{i_1} \cdot ... \cdot e_{i_k} = \sum_{k=0}^{2^{\dim F}} \sum_{I_k} w_{I_k} e_{i_1} \cdot ... \cdot e_{i_k}$
Notice that $w_0 \in K$

5. Isomorphism with the exterior algebra:

There is the isomorphism of vector spaces (but not of algebra: the product $\cdot$ does not correspond to the product $\wedge$) :
$e_{i_1} \cdot e_{i_2} \cdot .. e_{i_k} \in Cl(F, g) \leftrightarrow e_{i_1} \wedge e_{i_2} \wedge .. e_{i_k} \in \wedge F$
This isomorphism does not depend of the choice of the orthonormal basis

9.1.3 Involutions

1. Principal involution α

Definition 480 *The **principal involution** in Cl(F,g) denoted* $\alpha : Cl(F, g) \rightarrow Cl(F, g)$ *acts on homogeneous elements as :* $\alpha(v_1 \cdot v_2 .. \cdot v_r) = (-1)^r (v_1 \cdot v_2 .. \cdot v_r)$

$\alpha \left(\sum_{k=0}^{2^{\dim F}} \sum_{\{i_1,...,i_k\}} w_{\{i_1,...,i_k\}} e_{i_1} \cdot ... \cdot e_{i_k} \right)$
$= \sum_{k=0}^{2^{\dim F}} \sum_{\{i_1,...,i_k\}} (-1)^k w_{\{i_1,...,i_k\}} e_{i_1} \cdot ... \cdot e_{i_k}$
It has the properties:
$\alpha \circ \alpha = Id,$
$\forall w, w' \in Cl(F, g) : \alpha(w \cdot w') = \alpha(w) \cdot \alpha(w')$

2. Decomposition of Cl(F,g)
It follows that Cl(F,g) is the direct sum of the two eigen spaces with eigen value ± 1 for α.

146

Definition 481 *The set $Cl_0(F,g)$ of elements of a Clifford algebra $Cl(F,g)$ which are invariant by the principal involution is a subalgebra and a Clifford algebra.*

$Cl_0(F,g) = \{w \in Cl(F,g) : \alpha(w) = w\}$

Its elements are the sum of homogeneous elements which are themselves product of an even number of vectors.

As a vector space its basis is $1, e_{i_1} \cdot e_{i_2} \cdot ..e_{i_{2k}} : i_1 < i_2... < i_{2k}$

Theorem 482 *The set $Cl_1(F,g)$ of elements w of a Clifford algebra $Cl(F,g)$ such that $\alpha(w) = -w$ is a vector subspace of $Cl(F,g)$*

$Cl_1(F,g) = \{w \in Cl(F,g) : \alpha(w) = -w\}$

It is not a subalgebra. As a vector space its basis is $e_{i_1} \cdot e_{i_2} \cdot ..e_{i_{2k+1}} : i_1 < i_2... < i_{2k+1}$

$Cl_0 \cdot Cl_0 \subset Cl_0, Cl_0 \cdot Cl_1 \subset Cl_1, Cl_1 \cdot Cl_0 \subset Cl_1, Cl_1 \cdot Cl_1 \subset Cl_0$ so Cl(F,g) is a $\mathbb{Z}/2$ graded algebra.

3. Transposition

Definition 483 *The **transposition** on Cl(F,g) is the involution which acts on homogeneous elements by : $(v_1 \cdot v_2... \cdot v_r)^t = (v_r \cdot v_{r-1}... \cdot v_1)$.*

$$\left(\sum_{k=0}^{2^{\dim F}} \sum_{\{i_1,...,i_k\}} w_{\{i_1,...,i_k\}} e_{i_1} \cdot ... \cdot e_{i_k}\right)^t = \sum_{k=0}^{2^{\dim F}} \sum_{\{i_1,...,i_k\}} (-1)^{\frac{k(k-1)}{2}} w_{\{i_1,...,i_k\}} e_{i_1} \cdot ... \cdot e_{i_k}.$$

9.1.4 Scalar product on the Clifford algebra

Theorem 484 *A non degenerate bilinear symmetric form g on a vector space F can be extended in a non degenerate bilinear symmetric form G on $Cl(F,g)$.*

Consider a basis of Cl(F,g) deduced from an orthonormal basis of F. Define G by :

$i_1 < i_2.. < i_k, j_1 < j_2.. < j_k : G(e_{i_1} \cdot e_{i_2} \cdot ..e_{i_k}, e_{j_1} \cdot e_{j_2} \cdot ..e_{j_l}) = \delta_{kl} g(e_{i_1}, e_{j_1}) \times$
$...g(e_{i_k}, e_{j_k}) = \delta_{kl} \eta_{i_1 j_1} ... \eta_{i_k j_k}$

A basis of Cl(F,g) is an orthonormal basis for G. G does not depend on the choice of the basis. It is not degenerate.

k,l$\in K : G(k,l) = -kl$

$u, v \in F : G(u,v) = g(u,v)$

$w = \sum_{i<j} w_{ij} e_i \cdot e_j, w' = \sum_{i<j} w'_{ij} e_i \cdot e_j : G(w,w') = \sum_{i<j} w_{ij} w'_{ij} \eta_{ii} \eta_{jj}$

a,u,v$\in Cl(F,g) : G(u,v) = \langle u \cdot v^t \rangle$ where $\langle u \cdot v^t \rangle$ is the scalar component of $u \cdot v^t$

The transpose is the adjoint of the left and right Clifford product in the meaning :

$G(a \cdot u, v) = G(u, a^t \cdot v); G(u \cdot a, v) = G(u, v \cdot a^t)$

9.1.5 Volume element

Volume element

Definition 485 *A volume element of the Clifford algebra Cl(F,g) is an element ϖ such that $\varpi \cdot \varpi = 1$*

Let F be n dimensional and $(e_i)_{i=1}^n$ an orthonormal basis of (F,g) with K=R or C. The element : $e_0 = e_1 \cdot e_2... \cdot e_n \in Cl(F,g)$ does not depend on the choice of the orthonormal basis. It has the properties :

$e_0 \cdot e_0 = (-1)^{\frac{n(n-1)}{2}+q} = \pm 1$
$e_0 \cdot e_0 = +1$ if p-q=0,1 mod 4
$e_0 \cdot e_0 = -1$ if p-q=2,3 mod 4
where p,q is the signature of g if K=R. If K=C then q=0 and p=n

Thus if K=$\mathbb{C}$ there is always a volume element ϖ of Cl(F,g),which does not depend of a basis. It is defined up to sign by: $\varpi = e_0$ if $e_0 \cdot e_0 = 1$, and $\varpi = ie_0$ if $e_0 \cdot e_0 = -1$

If K=$\mathbb{R}$ and $e_0 \cdot e_0 = 1$ there is always a volume element ϖ of Cl(F,g),which does not depend of a basis, such that $\varpi \cdot \varpi = 1$. It is defined up to sign by: $\varpi = e_0$ if $e_0 \cdot e_0 = 1$. If $e_0 \cdot e_0 = -1$ Cl(F,g) can be extended to its complexified Cl_c(F,g) (see below).

So in the following we assume that such a volume element ϖ has been defined in Cl(F,g) or Cl_c(F,g) and we consider the complex case.

Decomposition of Cl(F,g)

Theorem 486 *The Clifford subalgebra $Cl_0(F,g) = Cl_0^+(F,g) \oplus Cl_0^-(F,g)$ where $Cl_0^+(F,g), Cl_0^-(F,g)$ are two isomorphic subalgebras and $\forall w \in Cl_0^+(F,g), w' \in Cl_0^-(F,g) : w \cdot w' = 0$*
The vector space $Cl_1(F,C) = Cl_1^+(F,C) \oplus Cl_1^-(F,C)$

Proof. The map : $w \to \varpi \cdot w$ is a linear map on Cl(F,g) and $\varpi \cdot (\varpi \cdot w) = w$ so it has ± 1 as eigen values.

Let be the two eigen spaces :
$Cl^+(F,g) = \{w \in Cl(F,g) : \varpi \cdot w = w\}$,
$Cl^-(F,g) = \{w \in Cl(F,g) : \varpi \cdot w = -w\}$
We have : $Cl(F,g) = Cl^+(F,g) \oplus Cl^-(F,g)$ as eigen spaces for different eigen values

$Cl_0^+(F,g) = Cl_0(F,g) \cap Cl^+(F,g)$ and $Cl_0^-(F,g) = Cl_0(F,g) \cap Cl^-(F,g)$ are subspace of $Cl_0(F,g)$

$Cl_0(F,g), Cl^+(F,g)$ are subalgebras, so is $Cl_0^+(F,g)$

if $w, w' \in Cl_0^-(4,C), \varpi \cdot w \cdot w' = -w \cdot w' \Leftrightarrow w \cdot w' \in Cl_0^-(F,g) : Cl_0^-(F,g)$ is a subalgebra

the only element common to the two subalgebras is 0, thus $Cl_0(F,g) = Cl_0^+(F,g) \oplus Cl_0^-(F,g)$

ϖ commute with any element of $Cl_0(F, g)$, and anticommute with all elements of $Cl_1(F, g)$ so

If $w \in Cl_0^+(F, g)$, $w' \in Cl_0^-(F, g) : \varpi \cdot w = w, \varpi \cdot w' = -w'$

$\varpi \cdot w \cdot \varpi \cdot w' = w \cdot \varpi \cdot \varpi \cdot w' = w \cdot w' = -\varpi \cdot w \cdot w' = -w \cdot w' \Rightarrow w \cdot w' = 0$

Similarly : $Cl_1(F, C) = Cl_1^+(F, C) \oplus Cl_1^-(F, C)$ (but they are not subalgebras) ∎

So any element w of $Cl(F, g)$ can be written : $w = w^+ + w^-$ with $w^+ \in Cl^+(F, g)$, $w^- \in Cl^-(F, g)$

Creation and annihiliation operators

Definition 487 *With the volume element ϖ in Cl(F,g)*

*The **creation operator** is $p_+ = \frac{1}{2}(1 + \varpi)$*

*The **annihiliation operator** is $p_+ = \frac{1}{2}(1 - \varpi)$*

Let be : $p_\epsilon = \frac{1}{2}(1 + \epsilon\varpi)$ with $\epsilon = \pm 1$

Identities :

$p_\epsilon^2 = p_\epsilon, p_+ \cdot p_- = p_- \cdot p_+ = 0, p_+ + p_- = 1$

For any $v \in F : p_\epsilon v = v p_{-\epsilon}$

$p_\epsilon \cdot w = w \cdot p_\epsilon = \frac{1}{2}((1 + \epsilon) w^+ + (1 - \epsilon) w^-)$

So for any $w = w^+ + w^- \in Cl(F, g)$:

$p_+ \cdot w = w^+, p_- \cdot w = w^-; p_+ \cdot w^+ = w^+,$

$p_- \cdot w^+ = 0; p_- \cdot w^- = w^-, p_- \cdot w^+ = 0$

$w^+.p_+ = w^+, w^-.p_+ = 0, w^+.p_- = 0, w^-.p_- = w^-$

$w^+.p_+ = 0, w^-.p_+ = w^-, w^+.p_- = w^+, w^-.p_- = 0$

9.2 Pin and Spin groups

9.2.1 Adjoint map

1. Inverse of an element of Cl(F,g)

Theorem 488 *In a Clifford algebra any element which is the product of non null norm vectors has an inverse for $\cdot$:*

$$(u_1 \cdot ... \cdot u_k)^{-1} = \alpha\left((u_1 \cdot ... \cdot u_k)^t\right) / \prod_{r=1}^{k} g(u_r, u_r)$$

So the set $(GCl(F, g), \cdot)$ of invertible elements of Cl(F,g) is a group (but not a vector space).

2. Adjoint map:

Definition 489 *The **adjoint map**, denoted **Ad** , is the map :*

Ad $: GCl(F, g) \times Cl(F, g) \to Cl(F, g) :: \mathbf{Ad}_w u = \alpha(w) \cdot u \cdot w^{-1}$

Where GCl(F,g) is the group of invertible elements of the Clifford algebra Cl(F,g)

Theorem 490 *The adjoint map* **Ad** *is a* $(GCl(F,g), \cdot)$ *group automorphism*

If $w, w' \in GCl(F,g) : \mathbf{Ad}_w \circ \mathbf{Ad}_{w'} = \mathbf{Ad}_{w \cdot w'}$

3. Clifford group :
The Clifford group is the set : $P = \{w \in GCl(F,g) : \mathbf{Ad}_w(F) \subset F\}$

Theorem 491 *The map :* **Ad** $:P{\to}O(F,g)$ *is a surjective morphism of groups*

Proof. If $w \in P$ then $\forall u, v \in F : g(\mathbf{Ad}_w u, \mathbf{Ad}_w v) = g(u,v)$ so $\mathbf{Ad}_w \in O(F,g)$ the group of orthogonal linear maps with g

Over (F,g) a reflexion of vector u with g(u,u)$\neq$ 0 is the orthogonal map :
$R(u) : F \to F :: R(u)x = x - 2\frac{g(x,u)}{g(u,u)}u$ and $x, u \in F : \mathbf{Ad}_u x = R(u)x$

Any orthogonal linear map over a n-dimensional vector space can be written as the product of at most 2n reflexions. Which reads : $\forall h \in O(F,g), \exists u_1, ...u_k \in F, k \leq 2\dim F : $ h=$R(u_1)\circ...R(u_k) = \mathbf{Ad}_{u_1}\circ...\circ\mathbf{Ad}_{u_r} = \mathbf{Ad}_{u_1 \cdot ... \cdot u_k} = \mathbf{Ad}_w, w \in P$

So $\mathbf{Ad}_{w \cdot w'} = \mathbf{Ad}_{u_1 \cdot ... \cdot u_k} \circ \mathbf{Ad}_{u_1' \cdot ... \cdot u_l'} = \mathbf{Ad}_{u_1} \circ ... \circ \mathbf{Ad}_{u_r} \circ \mathbf{Ad}_{u_1'} \circ ... \circ \mathbf{Ad}_{u_l'} = h \circ h'$

Thus the map : **Ad** $:P{\to}O(F,g)$ is a surjective homomorphism and $(\mathbf{Ad})^{-1}(O(F,g))$ is the subset of P comprised of homogeneous elements of Cl(F,g), products of vectors u_k with g(u_k,u_k)$\neq$ 0 ∎

9.2.2 Pin group

1. Definition

Definition 492 *The* **Pin group** *of Cl(F,g) is the set :*
$Pin(F,g) = \{w \in Cl(F,g), w = w_1 \cdot ... \cdot w_r, g(w_k, w_k) = 1\}$ *with* $\cdot$

If $w \in Pin(F,g)$ then :
$\alpha(w) = (-1)^r w$ and $w^t = w^{-1}$
$\forall u, v \in F : g(\mathbf{Ad}_w u, \mathbf{Ad}_w v) = g(u,v)$

Theorem 493 $(Pin(F,g), \cdot)$ *is a subgroup of the Clifford group*

Proof. $(w_1 \cdot ..w_k)^{-1} = \alpha\left((w_1... \cdot w_k)^t\right) = (-1)^k w_k \cdot w_{k-1} \cdot ... \cdot w_1 \in Pin(F,g)$
$\forall v \in F : \mathbf{Ad}_w v = u_1 \cdot ...u_k \cdot v \cdot u_k \cdot u_{k-1}... \cdot u_1$
$u_k \cdot v \cdot u_k = -2v + 2\left(\sum_j \eta_{jj} u_{kj} v_{kj}\right) u_k \in F$ ∎

2. Morphism with O(F,g):

Theorem 494 **Ad** *is a surjective group morphism :* **Ad** $:(Pin(F,g), \cdot) \to (O(F,g), \circ)$ *and O(F,g) is isomorphic to Pin(F,g)/{+1, −1}*

Proof. It is the restriction of the map $\mathbf{Ad}$ $:P{\to}O\left(F,g\right)$ to Pin(F,g)

For any $h \in O\left(F,g\right)$ there are two elements $(w,-w)$ of Pin(F,g) such that :
$\mathbf{Ad}_w = h$ ∎

So there is an action of O(F,g) on Cl(F,g) : $\lambda\,:\,O(F,g)\,\times\,Cl(F,g)\,\to\,$
$Cl(F,g) :: \lambda\left(h,w\right) = Ad_s w$ where $s \in Pin(F,g) : Ad_s = h$

3. Action of Pin(F,g) on Cl(F,g):

Theorem 495 *(Cl(F,g),$\mathbf{Ad}$) is a representation of Pin(F,g):*

Proof. For any s in Pin(F,g) the map $\mathbf{Ad}_s$ is linear on Cl(F,g) : $\mathbf{Ad}_s\left(kw + k'w'\right) =$
$\alpha\left(s\right){\cdot}(kw + k'w'){\cdot}s^{-1} = k\alpha\left(s\right){\cdot}w{\cdot}s^{-1}+k'\alpha\left(s\right){\cdot}w'{\cdot}s^{-1}$ and $\mathbf{Ad}_s\mathbf{Ad}_{s'} =\mathbf{Ad}_{ss'}$,$\mathbf{Ad}_1 =$
Id_F ∎

Theorem 496 *(F,$\mathbf{Ad}$) is a representation of Pin(F,g)*

This is the restriction of the representation on Cl(F,g)
(see Representation of groups).

9.2.3 Spin group

1. Definition:

Definition 497 *The **Spin group** of Cl(F,g) is the set :$Spin\left(F,g\right) = \{w \in Cl\left(F,g\right) : w = w_1 \cdot ... \cdot w_{2r}, g(w_k, w_k$*
with ·

So $Spin\left(F,g\right) = Pin\left(F,g\right) \cap Cl_0\left(F,g\right)$
If $w \in Spin(F,g)$ then :
$\alpha\left(w\right) = w$ and $w^t = w^{-1}$
$\forall u,v \in F : g\left(\mathbf{Ad}_w u, \mathbf{Ad}_w v\right) = g\left(u,v\right)$

2. Morphism with SO(F,g):

Theorem 498 *$\mathbf{Ad}$ is a surjective group morphism :* $\mathbf{Ad}:\left(Spin(F,g),\cdot\right) \to$
$(SO\left(F,g\right),\circ)$ *and SO(F,g) is isomorphic to Spin(F,g)/$\{+1,-1\}$*

Proof. It is the restriction of the map $\mathbf{Ad}$ $:P{\to}\mathbf{SO}\left(F,g\right)$ to Spin(F,g)

For any $h \in SO\left(F,g\right)$ there are two elements $(w,-w)$ of Spin(F,g) such that
: $\mathbf{Ad}_w = h$ ∎

3. Actions over Cl(F,g):

Theorem 499 *There is an action of SO(F,g) on Cl(F,g) :*
$\lambda\,:\,SO(F,g)\,\times\,Cl(F,g)\,\to\,Cl(F,g)\,::\,\lambda\left(h,u\right)\,=\,w\,\cdot\,u\,\cdot\,w^{-1}$ *where* $w \in$
$Spin(F,g) : \mathbf{Ad}_w = h$

Theorem 500 *(Cl(F,g),$\mathbf{Ad}$) is a representation of Spin(F,g):* $\mathbf{Ad}_s w = s \cdot w \cdot$
$s^{-1} = s \cdot w \cdot s^t$

Proof. This is the restriction of the representation of Pin(F,g) ∎

Theorem 501 *(F,$\mathbf{Ad}$) is a representation of Spin(F,g)*

This is the restriction of the representation on Cl(F,g)

9.2.4 Characterization of Spin(F,g) and Pin(F,g)

We develop here properties of the Pin(F,g), Spin(F,g) which are useful in several other parts of the book (mainly Fiber bundles and Functional analysis). We need results which can be found in the Part Lie groups, but it seems better to deal with these topics here. We assume that the vector space F is finite n dimensional.

Lie Groups

  1. Lie groups

Theorem 502 *The groups Pin(F,g) and Spin(F,g) are Lie groups*

Proof. O(F,g) is a Lie group and Pin(F,g)=O(F,g)$\times \{+1, -1\}$ ∎

    Any element of Pin(F,g) reads in an orthonormal basis of F:

$s = \sum_{k=0}^{n} \sum_{\{i_1,..i_k\}} S_{i_1...i_{2k}} e_{i_1} \cdot e_{i_2} \cdot ..e_{i_k} = \sum_{k=0}^{n} \sum_{I_k} S_{I_k} E_{I_k}$ with $S_{I_k} \in K$,

where the components S_{I_k} are not independant because the generator vectors must have norm 1.

    Any element of Spin(F,g) reads :

$s = \sum_{k=0}^{N} \sum_{\{i_1,..i_{2k}\}} S_{i_1...i_{2k}} e_{i_1} \cdot e_{i_2} \cdot ..e_{i_{2k}} = \sum_{k=0}^{N} \sum_{I_k} S_{I_k} E_{I_k}$ with $S_{I_k} \in K, N \leq n/2$

with the same remark.

    So Pin(F,g) and Spin(F,g) are not vector spaces, but manifolds embedded in the vector space Cl(F,g) : they are Lie groups.

    Pin(F,g) and Spin(F,g) are respectively a double cover, as manifold, of O(F,g) and SO(F,g). However the latter two groups may be not connected and in these cases Pin(F,g) and Spin(F,g) are not a double cover as Lie group.

  2. Lie algebra of the group

Theorem 503 *The Lie algebra $T_1 Pin(F, g)$ is isomorphic to the Lie algebra o(F,g) of O(F,g)*

The Lie algebra $T_1 Spin(F, g)$ is isomorphic to the Lie algebra so(F,g) of SO(F,g)

Proof. O(F,g) is isomorphic to Pin(F,g)/$\{+1, -1\}$. The subgroup $\{+1, -1\}$ is a normal, abelian subgroup of Pin(F,g). So the derivative of the map $h : Pin(F, g) \to O(F, g)$ is a morphism of Lie algebra with kernel the Lie algebra of $\{+1, -1\}$ which is 0 because the group is abelian. So h'(1) is an isomorphism (see Lie groups). Similarly for $T_1 Spin(F, g)$ ∎

Component expressions of the Lie algebras

Theorem 504 *The Lie algebra of Pin(F,g) is a subset of Cl(F,g).*

Proof. With the formula above, for any map $s : \mathbb{R} \to Pin(F,g) : s(t) = \sum_{k=0}^{n} \sum_{\{i_1 .. i_k\}} S_{i_1 ... i_{2k}}(t) e_{i_1} \cdot e_{i_2} \cdot .. e_{i_k}$ and its derivative reads $\frac{d}{dt} s(t)|_{t=0} = \sum_{k=0}^{n} \sum_{\{i_1 .. i_k\}} \frac{d}{dt} S_{i_1 ... i_{2k}}(t)|_{t=0} e_{i_1} \cdot e_{i_2} \cdot .. e_{i_k}$ that is an element of Cl(F,g) $\blacksquare$

Because $h'(1) : T_1 Pin(F,g) \to o(F,g)$ is an isomorphism, for any vector $\overrightarrow{\kappa} \in o(F,g)$ there is an element $\sigma(\overrightarrow{\kappa}) = h'(1)^{-1} \overrightarrow{\kappa}$ of Cl(F,g). Our objective here is to find the expression of $\sigma(\overrightarrow{\kappa})$ in the basis of Cl(F,g).

Lemma 505 $\forall u \in F : \sigma(\overrightarrow{\kappa}_a) \cdot u - u \cdot \sigma(\overrightarrow{\kappa}_a) = J_a u$

Proof. i) In the standard representation (F,ȷ) of SO(p,q) an element h(s) of SO(p,q) reads ȷ(h(s)). And in the orthonormal basis of F the formula : $\mathbf{Ad}_s u = s \cdot u \cdot s^{-1} = h(s) u$ reads : $\mathbf{Ad}_s u = s \cdot u \cdot s^{-1} = ȷ(h(s)) u$ where u,s are expressed in their components with respect to the basis. By derivation with respect to s at s=1 $(\mathbf{Ad})'|_{s=1} : T_1 Pin(F,g) \to o(F,g)$ reads : $(\mathbf{Ad})'|_{s=1} \sigma(\overrightarrow{\kappa}) = ȷ(h'(1)) \overrightarrow{\kappa}$

With a basis $(\overrightarrow{\kappa}_a)_{a=1}^{m}$ of o(F,g) $\overrightarrow{\kappa} = \sum_{a=1}^{m} \kappa^a \overrightarrow{\kappa}_a$ and $ȷ(h'(1)) \overrightarrow{\kappa} = \sum_{a=1}^{m} \kappa^a J_a$ with $J_a = ȷ(h'(1))(\overrightarrow{\kappa}_a)$ where J_a is a nxn matrix such that : $[\eta][J_a]^t + [J_a][\eta] = 0$

ii) The derivation of the product with respect to s at s=t : $\mathbf{Ad}_s u = s \cdot u \cdot s^{-1}$ gives :

$(\mathbf{Ad}_s u)'|_{s=t} \xi_t = \xi_t \cdot u \cdot t^{-1} - t \cdot u \cdot t^{-1} \cdot \xi_t \cdot t^{-1}$

For t=1 and $\xi_t = \sigma(\overrightarrow{\kappa}) : (\mathbf{Ad}_s u)'|_{s=1} \sigma(\overrightarrow{\kappa}) = \sigma(\overrightarrow{\kappa}) \cdot u - u \cdot \sigma(\overrightarrow{\kappa})$

iii) The relation $(\mathbf{Ad})'|_{s=1} \sigma(\overrightarrow{\kappa}) = h'(1) \overrightarrow{\kappa}$ reads : $\sigma(\overrightarrow{\kappa}) \cdot u - u \cdot \sigma(\overrightarrow{\kappa}) = \sum_{a=1}^{m} \kappa^a J_a u$ and because σ is linear:

$\sum_{a=1}^{m} \kappa_a (\sigma(\overrightarrow{\kappa}_a) \cdot u - u \cdot \sigma(\overrightarrow{\kappa}_a)) = \sum_{a=1}^{m} \kappa^a J_a u$

that is : $\forall u \in F : \sigma(\overrightarrow{\kappa}_a) \cdot u - u \cdot \sigma(\overrightarrow{\kappa}_a) = J_a u$ $\blacksquare$

From there one can get a more explicit expression for the elements of the Lie algebra so(F,g).

Theorem 506 *The vector $\overrightarrow{\kappa}$ of the Lie algebra so(F,g) can be written in Cl(F,g) as : $\sigma(\overrightarrow{\kappa}) = \sum_{ij} [\sigma]_j^i e_i \cdot e_j$ with $[\sigma] = \frac{1}{4}[J][\eta]$ where $[J]$ is the matrix of $\overrightarrow{\kappa}$ in the standard representation of so(F,g)*

Proof. We have also $\sigma(\overrightarrow{\kappa}_a) = \sum_{k=0}^{N} \sum_{I_k} s_{aI_k} E_{I_k}$ where s_{aI_k} are fixed scalars (depending on the bases).

Thus : $\sum_{k=0}^{N} \sum_{I_k} s_{aI_k} (E_{I_k} \cdot u - u \cdot E_{I_k}) = J_a(u)$ and taking $u = e_i$:

$\forall i = 1..n : \sum_{k=0}^{N} \sum_{I_k} s_{aI_k} (E_{I_k} \cdot e_i - e_i \cdot E_{I_k}) = J_a(e_i) = \sum_{j=1}^{n} [J_a]_i^j e_j$

$I_k = \{i_1, ..., i_{2k}\}$:

$(E_{I_k} \cdot e_i - e_i \cdot E_{I_k}) = e_{i_1} \cdot e_{i_2} \cdot .. e_{i_{2k}} \cdot e_i - e_i \cdot e_{i_1} \cdot e_{i_2} \cdot .. e_{i_{2k}}$

If $i \notin I_k : E_{I_k} \cdot e_i - e_i \cdot E_{I_k} = 2(-1)^{l+1} e_{i_1} \cdot e_{i_2} \cdot .. e_{i_{2k}} \cdot e_i$

If $i \in I_k, i = i_i : E_{I_k} \cdot e_i = (-1)^{2k-l} \eta_{ii} e_{i_1} \cdot e_{i_2} .. \cdot \widehat{e_{i_l}} .. e_{i_{2k}}, e_i \cdot E_{I_k} = (-1)^{l-1} \eta_{ii} e_{i_1} \cdot e_{i_2} .. \cdot \widehat{e_{i_l}} .. e_{i_{2k}}$

so : $E_{I_k} \cdot e_i - e_i \cdot E_{I_k} = 2(-1)^l \eta_{ii} e_{i_1} \cdot e_{i_2} .. \cdot \widehat{e_{i_l}} .. e_{i_{2k}}$

So : $s_{aI_k} = 0$ for $k \neq 1$ and for k=1 : $I_{1pq} = \{e_p, e_q\}, p < q$:

$\sum_{p<q} s_{apq} (e_p \cdot e_q \cdot e_i - e_i \cdot e_p \cdot e_q) = \sum_{i<j} (-2 s_{aij} \eta_{ii} e_j) = \sum_{j=1}^{n} [J_a]_i^j e_j$

i<j : $s_{aij} = -\frac{1}{2} \eta_{ii} [J_a]_i^j$

$\kappa = -\frac{1}{2} \sum_{a=1}^{m} \kappa_a \sum_{i<j} \eta_{ii} [J_a]_i^j e_i \cdot e_j$

$[\eta] [J_a]^t + [J_a] [\eta] = 0 \Rightarrow [J_a]_i^j = -\eta_{ii} \eta_{jj} [J_a]_j^i$ so the formula is consistent if we replace i by j :

$\sigma(\overrightarrow{\kappa}_a) = -\frac{1}{2} \sum_{j<i} \eta_{jj} [J_a]_j^i e_j \cdot e_i$

$\sigma(\overrightarrow{\kappa}_a) = -\frac{1}{4} \left(\sum_{i<j} \eta_{ii} [J_a]_i^j e_i \cdot e_j + \sum_{j<i} \eta_{jj} [J_a]_j^i e_j \cdot e_i \right) = -\frac{1}{4} \left(\sum_{i<j} \eta_{ii} [J_a]_i^j e_i \cdot e_j - \sum_{j<i} \eta_{ii} [J_a]_i^j e_j \cdot e_i \right)$

$= -\frac{1}{4} \left(\sum_{i<j} \eta_{ii} [J_a]_i^j e_i \cdot e_j + \sum_{j<i} \eta_{ii} [J_a]_i^j e_i \cdot e_j \right) = -\frac{1}{4} \left(\sum_{i,j} \eta_{ii} [J_a]_i^j e_i \cdot e_j - \sum_i \eta_{ii} [J_a]_i^i e_i \cdot e_i \right)$

$= -\frac{1}{4} \left(\sum_{i,j} \eta_{ii} [J_a]_i^j e_i \cdot e_j - Tr([J_a]) \right) = -\frac{1}{4} \left(\sum_{i,j} \eta_{ii} [J_a]_i^j e_i \cdot e_j \right)$ because J is traceless.

$\sigma(\overrightarrow{\kappa}) = -\frac{1}{4} \left(\sum_{i,j} \eta_{ii} [J]_i^j e_i \cdot e_j \right)$

If we represent the components of $\sigma(\overrightarrow{\kappa})$ in a matrix $[\sigma]$ nxn : $\sigma(\overrightarrow{\kappa}) = \sum_{ij}$

$[\sigma]_j^i e_i \cdot e_j = -\frac{1}{4} \left(\sum_{i,j} \eta_{ii} [J]_i^j e_i \cdot e_j \right)$

$[\sigma]_j^i = -\frac{1}{4} ([J][\eta])_i^j \Leftrightarrow [\sigma] = -\frac{1}{4} ([J][\eta])^t = -\frac{1}{4} [\eta][J]^t = \frac{1}{4} [J][\eta]$ ∎

Theorem 507 *The action of Spin(F,g) on o(F,g) is :* $\mathbf{Ad}_s \sigma(\overrightarrow{\kappa}) = \sigma \left(\jmath^{-1} Conj_{[\jmath h(s)]} [\jmath(\overrightarrow{\kappa})] \right)$

Proof. $\mathbf{Ad}_s \sigma(\kappa) = s \cdot \sum_{ij} [\sigma]_j^i e_i \cdot e_j \cdot s^{-1} = \sum_{ij} [\sigma]_j^i \mathbf{Ad}_s e_i \cdot \mathbf{Ad}_s e_j = \sum_{ij}$
$[\sigma]_j^i [\jmath h(s)]_i^k e_k \cdot [\jmath h(s)]_j^l e_l$

$= \sum [\jmath h(s)]_i^k [\sigma]_j^i [\jmath h(s)]_j^l e_k \cdot e_l = \sum_{kl} \left([\jmath h(s)][\sigma][\jmath h(s)]^t \right) e_k \cdot e_l$

$\mathbf{Ad}_s \sigma(\kappa) = \sum_{kl} [\widetilde{\sigma}]_l^k e_k \cdot e_l = \sum_{kl} \left([\jmath h(s)][\sigma][\jmath h(s)]^t \right) e_k \cdot e_l$

$[\widetilde{\sigma}] = [\jmath h(s)][\sigma][\jmath h(s)]^t$

$[\widetilde{J}][\eta] = [\jmath h(s)][J][\eta][\jmath h(s)]^t$

but : $[\jmath h(s)]^t [\eta][\jmath h(s)] = [\eta]$

$[\widetilde{J}][\eta] = [\jmath h(s)][J][\eta][\eta][\jmath h(s)]^{-1}[\eta] = [\jmath h(s)][J][\jmath h(s)]^{-1}[\eta]$ ∎

Derivatives of the translation and adjoint map

1. Translations:

The translations on Pin(F,g) are : $s, t \in Pin(F, g) : L_s t = s \cdot t, R_s t = t \cdot s$

The derivatives with respect to t are : $L'_s t(\xi_t) = s \cdot \xi_t, R'_s t(\xi_t) = \xi_t \cdot s$ with $\xi_t \in T_t Pin(F, g)$

With : $\xi_t = L'_t(1) \sigma(\overrightarrow{\kappa}) = R'_t(1) \sigma(\overrightarrow{\kappa}) = t \cdot \sigma(\overrightarrow{\kappa}) = \sigma(\overrightarrow{\kappa}) \cdot t$

2. Adjoint map :

As a Lie group the adjoint map is the derivative of $s \cdot x \cdot s^{-1}$ with respect to x at x=1:

$Ad : T_1 Pin(F, g) \to \mathcal{L}(T_1 Pin(F, g); T_1 Pin(F, g)) :: Ad_s = (s \cdot x \cdot s^{-1})'|_{x=1} = L'_s(s^{-1}) \circ R'_{s^{-1}}(1) = R'_{s^{-1}}(s) \circ L'_s(1)$

$Ad_s \sigma(\overrightarrow{\kappa}) = s \cdot \sigma(\overrightarrow{\kappa}) \cdot s^{-1} = \mathbf{Ad}_s \sigma(\overrightarrow{\kappa})$

3. Using : $(\mathbf{Ad}_s u)'|_{s=t} \xi_t = \xi_t \cdot u \cdot t^{-1} - t \cdot u \cdot t^{-1} \cdot \xi_t \cdot t^{-1}$ and $\xi_t = L'_t(1) \sigma(\overrightarrow{\kappa}) = t \cdot \sigma(\overrightarrow{\kappa})$

$(\mathbf{Ad}_s u)'|_{s=t} t \cdot \sigma(\overrightarrow{\kappa}) = \mathbf{Ad}_t (\sigma(\overrightarrow{\kappa}) \cdot u - u \cdot \sigma(\overrightarrow{\kappa}))$

On the other hand (F, jh) is a representation of Spin(F,g) so $(jh\,(s))' \,|_{s=t} = jh(t) \circ jh'(1)L'_{t-1}t$

$(jh\,(s))' \,|_{s=t} L'_t\,(1)\,\overrightarrow{\kappa} = jh(t) \circ jh'(1)\overrightarrow{\kappa}$

9.3 Classification of Clifford algebras

Clifford algebras are very rich structures, so it is not too surprising that they all look alike : there are not too possible many Clifford algebras. Thus the idea of classifying the Clifford algebras and, as usual, this starts with morphisms of Clifford algebras, meaning maps between Clifford algebras which preserve all the defining features of these structures. The second step is to look after simpler sets, which can be viewed as "workable" proxy for Clifford algebras. This leads, always along the same path, to the representation theory of Clifford algebra. It looks like, but is not totally identical, to the usual representation theory of algebras and groups.

9.3.1 Morphisms of Clifford algebras

Definition

Definition 508 *A **Clifford algebra morphism** between the Clifford algebras $Cl(F_1, g_1), Cl\,(F_2, g_2)$ on the same field K is an algebra morphism F : $Cl\,(F_1, g_1) \rightarrow Cl\,(F_2, g_2)$*

Which means that :
$\forall w, w' \in F_1, \forall k, k' \in K \;:\; F\,(kw + k'w') = kF\,(w) + k'F(w'), F\,(1) = 1, F\,(w \cdot w') = F\,(w) \cdot F\,(w')$

It entails that :
$F\,(u \cdot v + v \cdot u) = F\,(u) \cdot F\,(v) + F\,(v) \cdot F\,(u) = 2g_2\,(F\,(u), F\,(v)) = F\,(2g_1\,(u, v)) = 2g_1\,(u, v)$

so F must preserve the scalar product.

Categories

Theorem 509 *Clifford algebras on a field K and their morphisms constitute a category $\mathfrak{Cl}_K$.*

The product of Clifford algebras morphisms is a Clifford algebra morphism.

Vector spaces (V,g) on the same field K endowed with a symmetric bilinear form g, and linear maps f which preserve this form, constitute a category, denoted $\mathfrak{V}_B$

$f \in \hom_{\mathfrak{V}_B}\,((F_1, g_1), (F_2, g_2)) \Leftrightarrow f \in L\,(V_1; V_2), \forall u, v \in F_1 : g_2\,(f\,(u), f\,(v)) = g_1\,(u, v)$

We define the functor : $\mathfrak{TCl} : \mathfrak{V}_B \mapsto \mathfrak{Cl}_K$ which associates :

to each object (F,g) of $\mathfrak{V}_B$ its Clifford algebra Cl(F,g) : $\mathfrak{TCl} : (F, g) \mapsto Cl\,(F, g)$

to each morphism of vector spaces a morphism of Clifford algebras :

$\mathfrak{TCl} : f \in \hom_{\mathfrak{V}_B}\left(\left(F_1, g_1\right), \left(F_2, g_2\right)\right) \mapsto F \in \hom_{\mathfrak{Cl}_K}\left(\left(F_1, g_1\right), \left(F_2, g_2\right)\right)$

$F : Cl(F_1, g_1) \to Cl\left(F_2, g_2\right)$ is defined as follows :

$\forall k, k' \in K, \forall u, v \in F_1 : F\left(k\right) = k, F\left(u\right) = f\left(u\right), F\left(ku + k'v\right) = kf\left(u\right) + k'f\left(v\right), F\left(u \cdot v\right) = f\left(u\right) \cdot f\left(v\right)$

and as a consequence :

$F\left(u \cdot v + v \cdot u\right) = f\left(u\right) \cdot f\left(v\right) + f\left(v\right) \cdot f\left(u\right) = 2g_2\left(f\left(u\right), f\left(v\right)\right) = 2g_1\left(u, v\right) = F\left(2g_1\left(u, v\right)\right)$

Theorem 510 *Linear maps $f \in L(F_1; F_2)$ preserving the scalar product can be extended to morphisms F over Clifford algebras such that the diagram commutes :*

$$
\begin{array}{ccc}
(F_1, g_1) & \overset{Cl}{\to} & Cl\left(F_1, g_1\right) \\
\downarrow & & \downarrow \\
\downarrow f & & \downarrow F \\
\downarrow & & \downarrow \\
(F_2, g_2) & \overset{Cl}{\to} & Cl\left(F_2, g_2\right)
\end{array}
$$

Theorem 511 $\mathfrak{TCl} : \mathfrak{V}_B \mapsto \mathfrak{Cl}_K$ *is a functor from the category of vector spaces over K endowed with a symmetric bilinear form, to the category of Clifford algebras over K.*

Fundamental isomorphisms

As usual an isomorphism is a morphism which is also a bijective map. Two Clifford algebras which are linked by an isomorphism are said to be isomorphic. An automorphism of Clifford algebra is a Clifford isomorphism on the same Clifford algebra. The only Clifford automorphisms of finite dimensional Clifford algebras are the changes of orthonormal basis, with matrix A such that : $[A]^t\left[\eta\right]\left[A\right] = \left[\eta\right]$.

Theorem 512 *All Clifford algebras Cl(F,g) where F is a complex n dimensional vector space are isomorphic.*

Theorem 513 *All Clifford algebras Cl(F,g) where F is a real n dimensional vector space and g have the same signature, are isomorphic.*

Notation 514 $Cl\left(\mathbb{C}, n\right)$ *is the common structure of Clifford algebras over a n dimensional complex vector space*

$Cl\left(\mathbb{R}, p, q\right)$ *is the common structure of Clifford algebras over a real vector space endowed with a bilinear symmetric form of signature (+ p, - q).*

The common structure of $Cl\left(\mathbb{C}, n\right)$ is the Clifford algebra $\left(\mathbb{C}^n, g\right)$ over $\mathbb{C}$ endowed with the canonical bilinear form : $g\left(u, v\right) = \sum_{i=1}^{n} \left(u_i\right)^2, u_i \in \mathbb{C}$

$Cl_0\left(\mathbb{C}, n\right) \simeq Cl(\mathbb{C}, n-1)$

The common structure of $Cl\left(\mathbb{R}, p, q\right)$ is the Clifford algebra $\left(\mathbb{R}^n, g\right)$ over $\mathbb{R}$ with p+q=n endowed with the canonical bilinear form :

$g\left(u, v\right) = \sum_{i=1}^{p} \left(u_i\right)^2 - \sum_{i=p+1}^{n} \left(u_i\right)^2, u_i \in \mathbb{R}$

Warning ! The algebras Cl($\mathbb{R}$,p,q) and Cl($\mathbb{R}$,q,p) are *not* isomorphic if $p \neq q$. However $Cl(\mathbb{R}, 0, n) \simeq Cl(\mathbb{R}, n, 0)$

Pin and Spin groups are subsets of the Clifford algebras so, as such, are involved in the previous morphisms. However in their cases it is more logical to focus on their group structure, and consider group morphisms (see below).

9.3.2   Representation of a Clifford algebra

The previous theorems gives to the endeavour of classification a tautological flavour. So if we want to go further we have to give up a bit on the requirements for the morphism. This leads to the idea of representation, which is different and quite extensive. The representation of a Clifford algebra is a more subtle topic than it seems. To make this topic clearer we distinguish two kinds of representations, related but different, the algebraic representation and the geometric representation.

Definitions
1. Algebraic representations:

Definition 515 *An algebraic representation of a Clifford algebra Cl(F,g) over a field K is a couple (A, ρ) of an algebra $(A, \circ)$ on the field K and a map : $\rho : Cl\left(F, g\right) \to A$ which is an algebra morphism :*

$\forall X, Y \in Cl(F, g), k, k' \in K :$
$\rho\left(kX + k'Y\right) = k\rho(X) + k'\rho(Y),$
$\rho\left(X \cdot Y\right) = \rho(X) \circ \rho(Y), \rho\left(1\right) = I_A$

(with $\circ$ as internal operation in A, and A is required to be unital with unity element I)

$(Cl(F, g), \tau)$ where τ is any automorphism is an algebraic representation.

When Cl(F,g) is finite dimensional the algebra is usually a set of matrices, or of couple of matrices, as it will be seen in the next subsection.

If Cl(F,g) is a real Clifford algebra and A a complex algebra with a real structure : $A = A_{\mathbb{R}} \oplus iA_{\mathbb{R}}$, this is a real representation where elements X$\in A_{\mathbb{R}}$ and iX$\in iA_{\mathbb{R}}$ are deemed different.

If Cl(F,g) is a complex algebra A must be complex, possibly through a complex structure on A (usually by complexification : $A \to A_{\mathbb{C}} = A \oplus iA$).

Notice that if there is an algebra A isomorphic, as algebra, to Cl(F,g), there is not always the possibility to define a Clifford algebra structure on A (take the square matrices) and so a Clifford algebra morphism is more than a simple algebra morphism.

2. Geometric representation:

Definition 516 *A geometric representation of a Clifford algebra Cl(F,g) over a field K is a couple (V, ρ) of a vector space V on the field K and a map : $\rho : Cl(F, g) \rightarrow L(V; V)$ which is an algebra morphism :*

$\forall X, Y \in Cl(F, g), k, k' \in K : \rho(kX + k'Y) = k\rho(X) + k'\rho(Y),$
$\rho(X \cdot Y) = \rho(X) \circ \rho(Y), \rho(1) = Id_V$

Notice that the internal operation in L(V;V) is the composition of maps, and L(V;V) is always unital.

A geometric representation is a special algebraic representation, where a vector space V has been specified. When the algebra A is a set of mxm matrices, then the corresponding "standard geometric representation" is just $V = K^m$ and matrices act on the left on columns mx1 matrices. If Cl(F,g), V are finite dimensional then practically the geometric representation is a representation on an algebra of matrices, and for all purposes this is an algebraic representation. However the distinction is necessary for two reasons :

i) some of the irreducible algebraic representations of Clifford algebras are on sets of couples of matrices, possibly on another field K', for which there is no clear geometric interpretation.

ii) from the strict point of view of the representation theory, the "true nature" of the space vector V does not matter, and can be taken as K^n, this is the standard representation. But quite often, and notably in physics, we want to add some properties to V (such that a scalar product) and then the choice of V matters.

Definition 517 *An algebraic representation (A, ρ) of a Clifford algebra Cl(F,g) over a field K is **faithful** if ρ is bijective.*

Definition 518 *If (A, ρ) is an algebraic representation (A, ρ) of a Clifford algebra Cl(F,g) over a field K, a subalgebra A' of A is **invariant** if $\forall w \in Cl(F, g), \forall a \in A' : \rho(w) a \in A'$*

Definition 519 *An algebraic representation (A, ρ) of a Clifford algebra Cl(F,g) over a field K is **irreducible** if there is no subalgebra A' of A which is invariant by ρ*

Equivalence of representations
1. Composition of representation and morphisms :

The algebras A on the same field and their morphisms constitute a category. So composition of algebras morphisms are morphisms. Clifford algebras morphisms are algebras morphisms. So composition of Clifford algebras morphisms and algebras morphisms are still algebras morphisms.

Whenever there is an automorphism of Clifford algebra τ on Cl(F,g), and a morphism of algebra $\mu : A \rightarrow A'$, for any given representation (A, ρ), then $(A, \rho \circ \tau)$ or $(A', \mu \circ \rho)$ is still a representation of the Clifford algebra Cl(F,g).

But we need to know if this is still "the same" representation of Cl(F,g).

2. First it seems logical to say that a change of orthonormal basis in the Clifford algebra still gives the same representation. All automorphisms on a Clifford algebra Cl(F,g) are induced by a change of orthonormal basis in F, so :

Definition 520 *If (A, ρ) is an algebraic representation of a Clifford algebra Cl(F,g), τ an automorphism of Clifford algebra on Cl(F,g), then $(A, \rho \circ \tau)$ is an equivalent algebraic representation of Cl(F,g)*

So the representations $(Cl(F, g), \tau)$ of Cl(F,g) on itself are equivalent.

3. Algebraic representations :

Definition 521 *Two algebraic representations $(A_1, \rho_1), (A_2, \rho_2)$ of a Clifford algebra Cl(F,g) are said to be **equivalent** if there is a bijective algebra morphism $\phi : A_1 \rightarrow A_2$ such that : $\phi \circ \rho_1 = \rho_2$*

For a geometric representation a morphism such that : $\phi : L(V_1; V_1) \rightarrow L(V_2; V_2)$ is not very informative. This leads to:

4. Geometric representation:

Definition 522 *An **interwiner** between two geometric representations $(V_1, \rho_1), (V_2, \rho_2)$ of a Clifford algebra Cl(F,g) is a linear map $\phi : V_1 \rightarrow V_2$ such that $\forall w \in Cl(F, g) : \phi \circ \rho_1(w) = \rho_2(w) \circ \phi \in L(V_1; V_2)$*

Definition 523 *Two geometric representations of a Clifford algebra Cl(F,g) are said to be **equivalent** if there is a bijective interwiner.*

In two equivalent geometric representations $(V_1, \rho_1), (V_2, \rho_2)$ the vector spaces must have the same dimension. Conversely two Banach vector spaces with the same dimension (possibly infinite) on the same field are isomorphic so (V_1, ρ_1) give the equivalent representation (V_2, ρ_2) by : $\rho_2(w) = \phi \circ \rho_1(w) \circ \phi^{-1}$

If (V, ρ) is a geometric representation of Cl(F,g), μ and automorphism of V, then (V, ρ_2) is an equivalent representation with $\rho_2(w) = \phi \circ \rho_1(w) \circ \phi^{-1}$. Conjugation : $Conj_\mu \rho(w) = \mu \circ \rho(w) \circ \mu^{-1}$ is a morphism on L(V;V) so $(L(V; V), Conj_\mu \rho)$ is an algebraic representation equivalent to $(L(V; V), \rho)$.

The generators of a representation

The key point in a representation of a Clifford algebra Cl(F,g) is the representation of an orthonormal basis (e_i) of F, which can be seen as the generators of the algebra itself.

1. Operations on $\rho(Cl(F, g))$

If F is n dimensional, with orthonormal basis $(e_i)_{i=1}^n$, denote : $\gamma_i = \rho(e_i), i = 1..n, \gamma_0 = \rho(1)$.

ρ is injective (said faithful) iff all the γ_i are distincts.

As consequences of the morphism :

$$\rho\left(\sum_{k=0}^{2^{\dim F}} \sum_{\{i_1,...,i_k\}} w_{\{i_1,...,i_k\}} e_{i_1}\cdot...\cdot e_{i_k}\right) = \sum_{k=0}^{2^{\dim F}} \sum_{\{i_1,...,i_k\}} w_{\{i_1,...,i_k\}} \gamma_{i_1}...\gamma_{i_k}$$

$$\forall v,w \in F : \rho(v)\rho(w) + \rho(w)\rho(v) = 2g(v,w)I$$

$$\gamma_i\gamma_j + \gamma_j\gamma_i = 2\eta_{ij}\gamma_0$$

If $u \in Cl(F,g)$ is invertible then $\rho(u)$ is invertible and $\rho\left(u^{-1}\right) = \rho(u)^{-1} = \rho\left(-u/g(u,u)\right) = -\rho(u)/g(u,u)$

$$-\eta_{ii}\gamma_i = \gamma_i^{-1}$$

The images $\rho(Pin(F,g)), \rho(Spin(F,g))$ are subgroups of the group of invertible elements of A, and ρ is a group morphism.

2. so(F,g) Lie algebra :

An element $\overrightarrow{\kappa}$ of the Lie algebra so(F,g) and the corresponding matrices $[J]$ in its standard representation (F,j) is represented by $\sigma(\overrightarrow{\kappa}) = \frac{1}{4}\sum_{ij}\left([J][\eta]\right)_j^i e_i \cdot e_j$ in Cl(F,g) and by : $\rho(\kappa) = \frac{1}{4}\sum_{ij}\left([J][\eta]\right)_j^i \gamma_i\gamma_j$ in A

(V,ρ) is usually not a Lie algebra representation of the Lie algebra so(F,g).

3. Generators :

The image $\rho(Cl(F,g))$ is a subalgebra of A, generated by the set $(\rho(e_i))_{i=1}^n$ and the internal operations of A : linear combination on K and product (denoted without symbol). So one can consider the restriction of ρ to this set, on which it is surjective (but not necessarily injective if all the generators are not distinct).

Whenever there is an automorphism of Clifford algebra τ on Cl(F,g), and a morphism of algebra $\mu : A \to A'$, for any given representation (A,ρ), then $(A,\rho\circ\tau)$ or $(A',\mu\circ\rho)$ is still an equivalent representation of the Clifford algebra Cl(F,g), but the generators will not be the same.

For example :

if (A,ρ) is an algebraic representation, a change of orthonormal basis in F with matrix M is an automorphism τ of Clifford algebra. It gives the equivalent representation $(A,\rho\circ\tau)$ with new generators : $\gamma_i' = \sum_{j=1}^n M_j^i\gamma_j$. A must be such that : $[M]^t[\eta][M] = [\eta]$.

If (V,ρ) is a geometric representation then the map : $^t : L(V;V) \to L(V^*;V^*)$ is a morphism and (V^*,ρ^t) is still an equivalent geometric representation with $[\gamma_i'] = [\gamma_i]^t$.

If (V,ρ) is a geometric representation on a finite dimensional vector space, then a change of basis in V, with matrix M,is an automorphism μ and $(V,\mu\circ\rho)$ is still an equivalent geometric representation with generators $[\gamma_i'] = [M][\gamma_i][M]^{-1}$

But if A is a complex algebra with a real structure, conjugation is an anti-linear map, so cannot define a morphism.

Usually the problem is to find a set of generators which have also some nice properties, such being symmetric or hermitian. So the problem is to find a representation equivalent to a given representation (A,ρ) such that the generators have the required properties. This problem has not always a solution.

4. Conversely, given an algebra A on the field K, one can define an algebraic representation (A,ρ) of Cl(F,g) if we have a set of generators $(\gamma_i)_{i=0,...n}$ through the identities :

$\rho\left(e_i\right) = \gamma_i, \rho\left(1\right) = \gamma_0$ and $\rho\left(ke_i + k'e_j\right) = k\gamma_i + k'\gamma_j, \rho\left(e_i \cdot e_j\right) = \gamma_i\gamma_j$

The generators γ_i picked in A must meet some conditions (1) :

i) they must be invertible

ii) $\forall i, j : \gamma_i\gamma_j + \gamma_j\gamma_i = 2\eta_{ij}\gamma_0$

The condition $-\eta_{ii}\gamma_i = \gamma_i^{-1}$ is a consequence of the previous

The choice of the n+1 elements $\left(\gamma_i\right)_{i=0,...n}$ is not unique.

If, starting from a given algebra A of matrices, one looks for a set of generators which meet the conditions (1), but have also some nice properties such being symmetric or hermitian, there is no general guarantee that it can be done. However for the most usual representations one can choose symmetric or hermitian matrices as shown below. But then the representation is fixed by the choice of the generators.

5. All irreducible representations of Clifford algebras are on sets of rxr matrices with $r = 2^k$. So a practical way to solve the problem is to start with 2x2 matrices and extend the scope by Kronecker product :

Pick four 2x2 matrices E_j such that : $E_iE_j + E_jE_i = 2\eta_{ij}I_2, E_0 = I_2$ (the Dirac matrices or likes are usually adequate)

Compute : $F_{ij} = E_i \otimes E_j$

Then : $F_{ij}F_{kl} = E_iE_k \otimes E_jE_l$

With some good choices of combination by recursion one gets the right γ_i

The Kronecker product preserves the symmetry and the hermicity, so if one starts with E_j having these properties the γ_i will have it.

A classic representation

1. A Clifford algebra Cl(F,g) has a geometric representation on the algebra ΛF^* of linear forms on F

Consider the maps with $u \in V$:

$\lambda\left(u\right) : \Lambda_r F^* \to \Lambda_{r+1} F^* :: \lambda\left(u\right)\mu = u \wedge \mu$

$i_u : \Lambda_r F^* \to \Lambda_{r-1} F^* :: i_u\left(\mu\right) = \mu\left(u\right)$

The map : $\Lambda F^* \to \Lambda F^* :: \widetilde{\rho}\left(u\right) = \lambda\left(u\right) - i_u$ is such that :

$\widetilde{\rho}\left(u\right) \circ \widetilde{\rho}\left(v\right) + \widetilde{\rho}\left(v\right) \circ \widetilde{\rho}\left(u\right) = 2g\left(u, v\right)Id$

thus there is a map : $\rho : Cl(F, g) \to \Lambda F^*$ such that : $\rho \cdot \imath = \widetilde{\rho}$ and $(\Lambda F^*, \rho)$ is a geometric representation of Cl(F,g). It is reducible.

2. If F is an even dimensional real vector space, this construct can be extended to a complex representation.

Assume that there is a complex structure F_J on V with $J \in L(V; V) : J^2 = -Id$

Define the hermitian scalar product on F_J : $\langle u, v \rangle = g\left(u, v\right) + ig\left(u, J\left(v\right)\right)$

On the complex algebra ΛF_J^* define the map : $\widehat{\rho}\left(u\right) = -i\left(\lambda\left(u\right) - i_u\right)$ for $u \in F_J$

$(\Lambda F_J^*, \widehat{\rho} \cdot \imath)$ is a complex representation of $Cl(F, g)$ and a complex representation of the complexified $Cl_{\mathbb{C}}(F, g)$

9.3.3 Classification of Clifford algebras

All finite dimensional Clifford algebras, which have common structures, have faithful irreducible algebraic representations. There are 2 cases according to the field K over which F is defined.

Complex algebras

Theorem 524 *The unique faithful irreducible algebraic representation of the complex Clifford algebra Cl($\mathbb{C}$, n) is over a group of matrices of complex numbers*

The algebra A depends on n :

If n=2m : $A = \mathbb{C}(2^m)$: the square matrices $2^m \times 2^m$ (we get the dimension 2^{2m} as vector space)

If n=2m+1 : $A = \mathbb{C}(2^m) \oplus \mathbb{C}(2^m) \simeq \mathbb{C}(2^m) \times \mathbb{C}(2^m)$: couples (A,B) of square matrices $2^m \times 2^m$ (the vector space has the dimension 2^{2m+1}). A and B are two independant matrices.

The representation is faithful so there is a bijective correspondance between elements of the Clifford algebra and matrices.

The internal operations on A are the addition, multiplication by a complex scalar and product of matrices. When there is a couple of matrices each operation is performed independantly on each component (as in the product of a vector space):

$\forall ([A], [B]), ([A'], [B']) \in A, k \in \mathbb{C}$

$([A], [B]) + ([A'], [B']) = ([A] + [A'], [B] + [B'])$

$k([A], [B]) = (k[A], k[B])$

The map ρ is an isomorphism of algebras : $\forall w, w' \in Cl(\mathbb{C}, n), z, z' \in \mathbb{C}$:

$\rho(w) = [A]$ or $\rho(w) = ([A], [B])$

$\rho(zw + z'w') = z\rho(w) + z'\rho(w') = z[A] + z'[A']$ or $= (z[A] + z'[A'], z[B] + z'[B'])$

$\rho(w \cdot w') = \rho(w) \cdot \rho(w') = [A][B]$ or $= ([A][A'], [B][B'])$

In particular :

$Cl(\mathbb{C}, 0) \simeq \mathbb{C}; Cl(\mathbb{C}, 1) \simeq \mathbb{C} \oplus \mathbb{C}; Cl(\mathbb{C}, 2) \simeq \mathbb{C}(4)$

Real Clifford algebras

Theorem 525 *The unique faithful irreducible algebraic representation of the Clifford algebra Cl($\mathbb{R}$, p, q) is over an algebra of matrices*

(Husemoller p.161) The matrices algebras are over a field K' ($\mathbb{C}, \mathbb{R}$) or the division ring H of quaternions with the following rules :

$$\begin{bmatrix} (p-q)\bmod 8 & Matrices & (p-q)\bmod 8 & Matrices \\ 0 & \mathbb{R}\left(2^m\right) & 0 & \mathbb{R}\left(2^m\right) \\ 1 & \mathbb{R}\left(2^m\right)\oplus\mathbb{R}\left(2^m\right) & -1 & \mathbb{C}\left(2^m\right) \\ 2 & \mathbb{R}\left(2^m\right) & -2 & H\left(2^{m-1}\right) \\ 3 & \mathbb{C}\left(2^m\right) & -3 & H\left(2^{m-1}\right)\oplus H\left(2^{m-1}\right) \\ 4 & H\left(2^{m-1}\right) & -4 & H\left(2^{m-1}\right) \\ 5 & H\left(2^{m-1}\right)\oplus H\left(2^{m-1}\right) & -5 & \mathbb{C}\left(2^m\right) \\ 6 & H\left(2^{m-1}\right) & -6 & \mathbb{R}\left(2^m\right) \\ 7 & \mathbb{C}\left(2^m\right) & -7 & \mathbb{R}\left(2^m\right)\oplus\mathbb{R}\left(2^m\right) \end{bmatrix}$$

On H matrices are defined similarly as over a field, with the non commutativity of product.

Remark : the division ring of quaternions can be built as $Cl_0\left(\mathbb{R},0,3\right)$

$H\oplus H,\mathbb{R}\oplus\mathbb{R}$: take couples of matrices as above.

The representation is faithful so there is a bijective correspondance between elements of the Clifford algebra and of matrices. The dimension of the matrices in the table must be adjusted to n=2m or 2m+1 so that $\dim_{\mathbb{R}}A=2^n$

The internal operations on A are performed as above when A is a direct product of group of matrices.

ρ is a real isomorphism, meaning that $\rho\left(kw\right)=k\rho\left(w\right)$ only if $k\in\mathbb{R}$ even if the matrices are complex.

There are the following isomorphisms of algebras :

$Cl(\mathbb{R},0)\simeq\mathbb{R};Cl(\mathbb{R},1,0)\simeq\mathbb{R}\oplus\mathbb{R};Cl(\mathbb{R},0,1)\simeq\mathbb{C}$
$Cl(\mathbb{R},3,1)\simeq\mathbb{R}\left(4\right),Cl(\mathbb{R},1,3)\simeq H\left(2\right)$

9.3.4 Classification of Pin and Spin groups

Spin groups are important as they give non standard representations of the orthogonal groups SO(n) and SO(p,q). See more in the Lie groups part.

Pin and Spin are subset of the respective Clifford algebras, so the previous algebras morphisms entail group morphisms with the invertible elements of the algebras. Moreover, groups of matrices are well known and themselves classified. So what matters here is the group morphism with these "classical groups". The respective classical groups involved are the orthogonal groups O(K,p,q) for Pin and the special orthogonal groups SO(K,p,q) for Spin. A key point is that to one element of O(K,p,q) or SO(K,p,q) correspond two elements of Pin or Spin. This topic is addressed through the formalism of "cover" of a manifold (see Differential geometry) and the results about the representations of the Pin and Spin groups are presented in the Lie group part .

Complex case

Theorem 526 *All Pin(F,g) groups over complex vector spaces of same dimension are group isomorphic. The same for the Spin(F,g) groups.*

Notation 527 *$Pin\,(\mathbb{C}, n)$ is the common group structure of $Pin(Cl\,(\mathbb{C}, n))$*
$Spin\,(\mathbb{C}, n)$ is the common group structure of $Spin(Cl\,(\mathbb{C}, n))$

Each of the previous isomorphisms induces an isomorphism of groups:
Warning ! it does not extend to the multiplication by a scalar or a sum !
These groups are matrices groups (linear group of matrices).
$Spin\,(\mathbb{C}, n)$ is simply connected and is the universal double cover of SO($\mathbb{C}$,n):
SO($\mathbb{C}, n$)=Spin($\mathbb{C}$,n)/($\pm I$)

Real Case

Theorem 528 *All Pin(F,g) groups over real vector spaces of same dimension endowed with bilinear symmetric form of same signature are group isomorphic. The same for the Spin(F,g) groups.*

Notation 529 *$Pin\,(\mathbb{R}, p, q)$ is the common group structure of $Pin(Cl\,(\mathbb{R}, p, q))$*
$Spin\,(\mathbb{R}, p, q)$ is the common group structure of $Spin(Cl\,(\mathbb{R}, p, q))$

Each of the previous isomorphisms induces an isomorphism of groups.
These groups are matrices groups (linear group of matrices) and Lie groups
a) If p or q = 0:
Pin($\mathbb{R}$,0,n), Pin($\mathbb{R}$,n,0) are not isomorphic, they are not connected.
Spin($\mathbb{R}$,n)=Spin($\mathbb{R}$,0,n) and Spin($\mathbb{R}$,n,0) are isomorphic and are the unique double cover of SO($\mathbb{R}$,n)
For n>2 Spin($\mathbb{R}$,n) is simply connected and is the universal cover of SO($\mathbb{R}$,n)
b) If p,q are >0 :
Pin($\mathbb{R}$,p,q),Pin($\mathbb{R}$,q,p) are not isomorphic if $p \neq q$
Pin($\mathbb{R}$,p,q) is not connected, it maps to O($\mathbb{R}$,p,q) but the map is not surjective
Spin($\mathbb{R}$,p,q) and Spin($\mathbb{R}$,q,p) are isomorphic
If n>2 Spin($\mathbb{R}$,p,q) is a double cover of SO_0($\mathbb{R}$,p,q), the connected component of the identity of the group SO($\mathbb{R}$,p,q).

9.3.5   Complexification of a Clifford algebra

It is possible to extend any real vector space F to a complex vector space F_c and g can be extended by defining g_c(iu,v)=g_c(u,iv)=ig(u,v), which gives a complex Clifford algebra $Cl(F_c, g_c)$.

On the other hand the Clifford algebra can be complexified by extension :
$Cl(G, g) \rightarrow Cl_c\,(F, g) = Cl(F, g) \otimes C$
The two procedures give the same result : $Cl(F_c, g_c) = Cl_c\,(F, g)$
In this process $Cl(p, q) \rightarrow Cl(p + q, C) = Cl_c(p, q)$

The group $Spin_c\,(F, g)$ is the subgroup of $Cl_c\,(F, g)$ comprised of elements :
$S = zs$ where z is a module 1 complex scalar, and s$\in Spin(F, g)$. It is a subgroup of the group $Spin\,(F_c, g_c)$.

Part III
PART 3 : ANALYSIS

Analysis is a very large area of mathematics. It adds to the structures and operations of algebra the concepts of "proximity" and "limit". Its key ingredient is topology, a way to introduce these concepts in a very general but rigorous manner, to which is dedicated the first section. It is mainly a long, but by far not exhaustive, list of definitions and results which are necessary for a basic understanding of the rest of the book. The second section is dedicated to the theory of measure, which is the basic tool for integrals, with a minimum survey of probability theory. The third and fourth sections are dedicated to analysis on sets endowed with a vector space structure, mainly Banach spaces and algebras, which lead to Hilbert spaces and the spectral theory. The review is more detailed on these latter difficult and important topics.

10   GENERAL TOPOLOGY

Topology can be understood with two different, related, meanings. Initially it has been an extension of geometry, starting with Euler, Listing and pursued by Poincaré, to study "qualitative" properties of objects without referring to a vector space structure. Today this is understood as algebraic topology, of which some basic elements are presented below.

The second meaning, called "general topology", is the mathematical way to define "proximity" and "limit", and is the main object of this section. It has been developed in the beginning of the XX° century by Cantor, as an extension of the set theory, and developped with metrics over a set by Fréchet, Hausdorff and many others. General topology is still often introduced through metric spaces. But, when the basic tools such as open, compact,... have been understood, they are often easier to use, with a much larger scope. So we start with these general concepts. Metric spaces bring additional properties. Here also it has been usual to focus on definite positive metrics, but many results still hold with semi-metrics which are common.

This is a vast area, so there are many definitions, depending on the authors and the topic studied. We give only the most usual, which can be useful, and often a prerequisite, in advanced mathematics. We follow mainly Wilansky, Gamelin and Schwartz (tome 1). The reader can also usefully consult the tables of theorems in Wilansky.

10.1 Topological space

In this subsection topological concepts are introduced without any metric. They all come from the definition of a special collection of subsets, the open subsets.

10.1.1 Topology

Open subsets

Definition 530 *A **topological space** is a set E, endowed with a collection $\Omega \subset 2^E$ of subsets called **open** subsets such that :*

$\quad E \in \Omega, \varnothing \in \Omega$

$\quad \forall I : O_i \in \Omega, \cup_{i \in I} O_i \in \Omega$

$\quad \forall I, card I < \infty : O_i \in \Omega, \cap_{i \in I} O_i \in \Omega$

The key points are that every (even infinite) union of open sets is open, and every *finite intersection* of open sets is open.

The power set 2^E is the set of subsets of E, so $\Omega \subset 2^E$. Quite often the open sets are not defined by a family of sets, meaning a map : $I \rightarrow 2^E$

Example : in $\mathbb{R}$ the open sets are generated by the open intervals]a,b[(a and b excluded).

Topology

The **topology** on E is just the collection Ω of its open subsets, and a topological space will be denoted (E, Ω). Different collections define different topologies (but they can be equivalent : see below). There are many different topologies on the same set : there is always $\Omega_0 = \{\varnothing, E\}$ and $\Omega_\infty = 2^E$ (called the **discrete topology**).

When $\Omega_1 \subset \Omega_2$ the topology defined by Ω_1 is said to be "thinner" (or stronger) then Ω_2, and Ω_2 "coarser" (or weaker) than Ω_1. The issue is usually to find the "right" topology, meaning a collection of open subsets which is not too large, but large enough to bring interesting properties.

Closed subsets

Definition 531 *A subset A of a topological space (E, Ω) is **closed** if A^c is open.*

So :

Theorem 532 *In a topological space :*

$\quad \varnothing, E$ *are closed,*

$\quad$ *any intersection of closed subsets is closed,*

$\quad$ *any finite union of closed subsets is closed.*

A topology can be similarly defined by a collection of closed subsets.

Relative topology

Definition 533 *If X is a subset of the topological space (E, Ω) the **relative topology** (or induced topology) in X inherited from E is defined by taking as open subsets of $X : \Omega_X = \{O \cap X, O \in \Omega\}$. Then (X, Ω_X) is a topological space, and the subsets of Ω_X are said to be **relatively open** in X.*

But they are not necessarily open in E : indeed X can be any subset and one cannot know if $O \cap X$ is open or not in E.

10.1.2 Neighborhood

Topology is the natural way to define what is "close" to a point.

1. Neighborhood:

Definition 534 *A **neighborhood** of a point x in a topological space (E,Ω) is a subset n(x) of E which contains an open subset containing x: $\exists O \in \Omega : O \subset n(x), x \in O$*

Indeed a neighborhood is just a convenient, and abbreviated, way to say : "a subset which contains on open subset which contains x".

Notation 535 *n(x) is a neighborhood of a point x of the topological space (E,Ω)*

Definition 536 *A point x of a subset X in a topological space (E,Ω) is **isolated** in X if there is a neighborhood n(x) of x such that $n(x) \cap X = \{x\}$*

2. Interior, exterior:

Definition 537 *A point x is an **interior point** of a subset X of the topological space (E,Ω) if X is a neighborhood of x. The **interior** $\overset{\circ}{X}$ of X is the set of its interior points, or equivalently,the largest open subset contained in X (the union of all open sets contained in X) . The **exterior** $(\overset{\circ}{X^c})$ of X is the interior of its complement, or equivalently,the largest open subset which does not intersect X (the union of all open sets which do not intersect X)*

Notation 538 *$\overset{\circ}{X}$ is the interior of the set X*

Theorem 539 *$\overset{\circ}{X}$ is an open subset : $\overset{\circ}{X} \sqsubseteq X$ and $\overset{\circ}{X} = X$ iff X is open.*

3. Closure:

Definition 540 *A point x is **adherent** to a subset X of the topological space (E,Ω) if each of its neighborhoods meets X. The **closure** $\overline{X}$ of X is the set of the points which are adherent to X or, equivalently, the smallest closed subset which contains X (the intersection of all closed subsets which contains X)*

Notation 541 $\overline{X}$ *is the closure of the subset X*

Theorem 542 $\overline{X}$ *is a closed subset : $X \sqsubseteq \overline{X}$ and $\overline{X} = X$ iff X is closed.*

Definition 543 *A subset X of the topological space (E,Ω) is **dense** in E if its closure is $E : \overline{X} = E$*

$$\Leftrightarrow \forall \varpi \in \Omega, \varpi \cap X \neq \varnothing$$

4. Border:

Definition 544 *A point x is a **boundary point** of a subset X of the topological space (E,Ω) if each of its neighborhoods meets both X and X^c. The **border** ∂X of X is the set of its boundary points.*

Theorem 545 ∂X *is a closed subset*

Notation 546 ∂X *is the border (or boundary) of the set X*

Another common notation is $\dot{X} = \partial X$

5. The relation between interior, border, exterior and closure is summed up in the following theorem:

Theorem 547 *If X is a subset of a topological space (E,Ω) then :*
$$\overline{X} = \overset{\circ}{X} \cup \partial X = \left(\left(\overset{\circ}{X^c} \right) \right)^c$$
$$\overset{\circ}{X} \cap \partial X = \varnothing$$
$$(\overset{\circ}{X^c}) \cap \partial X = \varnothing$$
$$\partial X = \overline{X} \cap \overline{(X^c)} = \partial (X^c)$$

10.1.3 Base of a topology

A topology is not necessarily defined by a family of subsets. The base of a topology is just a way to define a topology through a family of subsets, and it gives the possibility to precise the thinness of the topology by the cardinality of the family.

Base of a topology

Definition 548 *A **base** of a topological space (E, Ω) is a family $(B_i)_{i \in I}$ of subsets of E such that : $\forall O \in \Omega, \exists J \subset I : O = \cup_{j \in J} B_j$*

Theorem 549 *(Gamelin p.70) A family $(B_i)_{i \in I}$ of subsets of E is a base of the topological space (E, Ω) iff $\forall x \in E, \exists i \in I : x \in B_i$ and $\forall i, j \in I : x \in B_i \cap B_j \Rightarrow \exists k \in I : x \in B_k, B_k \subset B_i \cap B_j$*

Theorem 550 *(Gamelin p.70) A family $(B_i)_{i \in I}$ of open subsets of Ω is a base of the topological space (E, Ω) iff $\forall x \in E, \forall n(x)$ neighborhood of x, $\exists i \in I : x \in B_i, B_i \subset n(x)$*

Countable spaces

The word "countable" in the following can lead to some misunderstanding. It does not refer to the number of elements of the topological space but to the cardinality of a base used to define the open subsets. It is clear that a topology is stronger if it has more open subsets, but too many opens make difficult to deal with them. Usually the "right size" is a countable base.

1. Basic definitions:

Definition 551 *A topological space is*
*   ***first countable** if each of its points has a neighborhood with a countable base.*
*   ***second countable** if it has a countable base.*

Second countable$\Rightarrow$First countable

In a second countable topological space there is a family $(B_n)_{n \in \mathbb{N}}$ of subsets which gives, by union and finite intersection, all the open subsets of Ω.

2. Open cover:

The "countable" property appears quite often through the use of open covers, where it is useful to restrict their size.

Definition 552 *An **open cover** of a topological space (E, Ω) is a family $(O_i)_{i \in I}, O_i \subset \Omega$ of open subsets whose union is E. A **subcover** is a subfamily of an open cover which is still an open cover. A **refinement** of an open cover is a family $(F_j)_{j \in J}$ of subsets of E whose union is E and such that each member is contained in one of the subset of the cover : $\forall j \in J, \exists i \in I : F_j \sqsubseteq O_i$*

Theorem 553 *Lindelöf (Gamelin p.71) If a topological space is second countable then every open cover has a countable open subcover.*

3. Another useful property of second countable spaces is that it is often possible to extend results obtained on a subset of E. The procedure uses dense subspaces.

Definition 554 *A topological space* (E, Ω) *is* **separable** *if there is a countable subset of E which is dense in E.*

Theorem 555 *(Gamelin p.71) A second countable topological space is separable*

10.1.4 Separation

It is useful to have not too many open subsets, but it is also necessary to have not too few in order to be able to "distinguish" points. They are different definitions of this concept. By far the most common is the "Hausdorff" property.

Definitions

They are often labeled by a T from the german "Trennung"=separation.

Definition 556 *(Gamelin p.73) A topological space* (E, Ω) *is*

Hausdorff *(or T2) if for any pair x,y of distinct points of E there are open subsets O,O' such that* $x \in O, y \in O', O \cap O' = \varnothing$

regular *if for any pair of a closed subset X and a point* $y \notin X$ *there are open subsets O,O' such that* $X \subset O, y \in O', O \cap O' = \varnothing$

normal *if for any pair of closed disjoint subsets X,Y* $X \cap Y = \varnothing$ *there are open subsets O,O' such that* $X \subset O, Y \subset O', O \cap O' = \varnothing$

T1 *if a point is a closed set.*

T3 *if it is T1 and regular*

T4 *if it is T1 and normal*

The definitions for regular and normal can vary in the litterature (but Hausdorff is standard). See Wilansky p.46 for more.

Theorems

1. Relations between the definitions

Theorem 557 *(Gamelin p.73) T4* $\Rightarrow$ *T3* $\Rightarrow$ *T2* $\Rightarrow$ *T1*

Theorem 558 *(Gamelin p.74) A topological space* (E, Ω) *is normal iff for any closed subset X and open set O containing X there is an open subset O' such that* $\overline{O'} \subset O$ *and* $X \subset O'$

Theorem 559 *(Thill p.84) A topological space is regular iff it is homeomorphic to a subspace of a compact Hausdorff space*

2. Separation implies the possibility to define continuous characteristic maps. Here the main thorems:

Theorem 560 *Urysohn (Gamelin p.75): If X,Y are disjoint subsets of a normal topological space* (E, Ω) *there is a continuous function* $f : E \rightarrow [0, 1] \subset \mathbb{R}$ *such that f(x)=0 on X and f(x)=1 on Y.*

Theorem 561 *Tietze (Gamelin p.76): If F is a closed subset of a normal topological space (E, Ω) $\varphi : F \to \mathbb{R}$ bounded continuous, then there is $\Phi : E \to \mathbb{R}$ bounded continuous such that $\Phi = \varphi$ over F.*

Remarks :

1) "separable" is a concept which is not related to separation (see base of a topology).

2) it could seem strange to consider non Hausdorff space. In fact usually this is the converse which happens : one wishes to consider as "equal" two different objects which share basic properties (for instance functions which are almost everywhere equal) : thus one looks for a topology that does not distinguish these objects. Another classic solution is to build a quotient space through an equivalence relation.

10.1.5 Compact

Compact is a topological mean to say that a set is not "too large". The other useful concept is locally compact, which means that "bounded" subsets are compact.

Definitions

Definition 562 *A topological space (E, Ω) is :*
    compact if for any open cover there is a finite open subcover.
    countably compact if for any countable open cover there is a finite open subcover.
    locally compact if each point has a compact neighborhood
    compactly generated if a subset X of E is closed in E iff $X \cap K$ is closed for any compact K in E. We have the equivalent for open subsets.

In a second countable space an open cover has a countable subcover (Lindelöf theorem). Here it is finite.

Definition 563 *A subset X of topological space (E, Ω) is :*
    compact in E if for any open cover of X there is a finite subcover of X
    relatively compact if its closure is compact

Definition 564 *A **Baire space** is a topological space where the intersection of any sequence of dense subsets is dense*

Theorems
    Compact $\Rightarrow$ countably compact
    Compact $\Rightarrow$ locally compact
    Compact, locally compact, first countable spaces are compactly generated.

Theorem 565 *(Gamelin p.83) Any compact topological space is locally compact. Any discrete set is locally compact. Any non empty open subset of $\mathbb{R}^n$ is locally compact.*

Theorem 566 *(Gamelin p.79) Any finite union of compact subsets is compact.*

Theorem 567 *(Gamelin p.79) A closed subset of a compact topological space is compact.*

Theorem 568 *(Wilansky p.81) A topological space (E, Ω) is compact iff for any family $(X_i)_{i \in I}$ of subsets for wich $\cap_{i \in I} X_i = \varnothing$ there is a finite subfamily J for which $\cap_{i \in J} X_i = \varnothing$*

Theorem 569 *(Wilansky p.82) The image of a compact subset by a continuous map is compact*

Theorem 570 *(Gamelin p.80) If X is a compact subset of a Hausdorff topological space (E, Ω) :*
*   - X is closed*
*   - $\forall y \notin X$ there are open subsets O,O' such that : $y \in O, X \subset O', O \cap O' = \varnothing$*

Theorem 571 *(Wilansky p.83) A compact Hausdorff space is normal and regular.*

Theorem 572 *(Gamelin p.85) A locally compact Hausdorff space is regular*

Theorem 573 *(Wilansky p.180) A locally compact, regular topological space is a Baire space*

Compactification
   (Gamelin p.84)
   This is a general method to build a compact space from a locally compact Hausdorff topological space (E, Ω). Define F=E$\cup \{\infty\}$ where ∞ is any point (not in E). There is a unique topology for F such that F is compact and the topology inherited in E from F coincides with the topology of E. The open subsets in F are either open subsets of E or O such that $\infty \in O$ and E\O is compact in E.

10.1.6   Paracompact spaces

The most important property of paracompact spaces is that they admit a **partition of unity**, with which it is possible to extend local constructions to global constructions over E. This is a mandatory tool in differential geometry.

Definitions

Definition 574 *A family $(X_i)_{i \in I}$ of subsets of a topological space (E, Ω) is :*
*   **locally finite** if every point has a neighborhood which intersects only finitely many elements of the family.*
* σ-**locally finite** if it is the union of countably many locally finite families.*

Definition 575 *A topological space (E, Ω) is **paracompact** if any open cover of E has a refinement which is locally finite.*

Theorems

Theorem 576 *(Wilansky p.191) The union of a locally finite family of closed sets is closed*

Theorem 577 *(Bourbaki) Every compact space is paracompact. Every closed subspace of a paracompact space is paracompact.*

Theorem 578 *(Bourbaki) Every paracompact space is normal*

Warning ! an infinite dimensional Banach space may not be paracompact

Existence of a partition of unity

Theorem 579 *(Nakahara p.206) For any paracompact Hausdorff topological space (E, Ω) and open cover $(O_i)_{i \in I}$, there is a family $(f_j)_{j \in J}$ of continuous functions $f_j : E \to [0, 1] \subset \mathbb{R}$ such that :*
*   - $\forall j \in J, \exists i \in I : support(f_i) \subset O_j$*
*   - $\forall x \in E, \exists n(x), \exists K \subset J, card(K) < \infty : \forall y \in n(x) : \forall j \in J \backslash K : f_j(y) = 0, \sum_{j \in K} f_j(y) = 1$*

10.1.7   Connected space

Connectedness is related to the concepts of "broken into several parts". This is a global property, which is involved in many theorems about unicity of a result.

Definitions

Definition 580 *(Schwartz I p.87) A topological space (E, Ω) is **connected** if it does not admit a partition into two subsets (other than E and $\varnothing$) which are both closed or both open, or equivalently if there are no subspace (other than E and $\varnothing$) which are both closed and open. A subset X of a topological space (E, Ω) is connected if it is connected in the induced topology.*

So if X is not connected (say disconnected) in E if there are two subspaces of E, both open or closed in E, such that $X = (X \cap A) \cup (X \cap B)$, $A \cap B \cap X = \varnothing$.

Definition 581 *A topological space (E, Ω) is **locally connected** if for each point x and each open subset O which contains x, there is a connected open subset O' such that $x \in O', O' \sqsubseteq O$.*

Definition 582 *The **connected component** C(x) of a point x of E is the union of all the connected subsets which contains x.*

It is the largest connected subset of E which contains x. So $x \sim y$ if $C(x) = C(y)$ is a relation of equivalence which defines a partition of E. The classes of equivalence of this relation are the connected components of E. They are disjoint, connected subsets of E and their union is E. Notice that the components are not necessarily open or closed. If E is connected it has only one component.

Theorems

Theorem 583 *The only connected subsets of $\mathbb{R}$ are the intervals $[a, b]$, $]a, b[$, $[a, b[$, $]a, b]$ (usually denoted $|a, b|$) a and b can be $\pm\infty$*

Theorem 584 *(Gamelin p.86) The union of disjointed connected subsets is connected*

Theorem 585 *(Gamelin p.86) The image of a connected subset by a continuous map is connected*

Theorem 586 *(Wilansky p.70) If X is connected in E, then its closure $\overline{X}$ is connected in E*

Theorem 587 *(Schwartz I p.91) If X is a connected subset of a topological space (E, Ω) , Y a subset of E such that $X \cap \overset{\circ}{Y} \neq \varnothing$ and $X \cap (\overline{Y})^c \neq \varnothing$ then $X \cap \partial Y \neq \varnothing$*

Theorem 588 *(Gamelin p.88) Each connected component of a topological space is closed. Each connected component of a locally connected space is both open and closed.*

10.1.8   Path connectedness

Path connectedness is a stronger form of connectedness.

Definitions

Definition 589 *A **path** on a topological space E is a continuous map : $c : J \rightarrow$ E from a connected subset J of $\mathbb{R}$ to E. The codomain C={$c(t), t \in J$} of c is a subset of E, which is a **curve**.*

The same curve can be described using different paths, called **parametrisation**. Take $f : J\prime \rightarrow J$ where J' is another interval of $\mathbb{R}$ and f is any bijective continuous map, then : $c' = c \circ f : J' \rightarrow E$ is another path with image C.

A path from a point x of E to a point y of E is a path such that $x \in C, y \in C$

Definition 590 *Two points x,y of a topological space (E, Ω) are **path-connected** (or arc-connected) if there is a path from x to y.*

A subset X of E is path-connected if any pair of its points are path-connected.

*The **path-connected component** of a point x of E is the set of the points of E which are path-connected to x.*

$x \sim y$ if x and y are path-connected is a relation of equivalence which defines a partition of E. The classes of equivalence of this relation are the path-connected components of E.

Definition 591 *A topological space (E, Ω) is **locally path-connected** if for any point x and neighborhood n(x) of x there is a neighborhood n'(x) included in n(x) which is path-connected.*

Theorems

Theorem 592 *(Schwartz I p.91) if X is a subset of a topological space (E, Ω), any path from $a \in \overset{\circ}{X}$ to $b \in (\overset{\circ}{X^c})$ meets ∂X*

Theorem 593 *(Gamelin p.90) If a subset X of a topological space (E, Ω) is path connected then it is connected.*

Theorem 594 *(Gamelin p.90) Each connected component of a topological space (E, Ω) is the union of path-connected components of E.*

Theorem 595 *(Schwartz I p.97) A path-connected topological space is locally path-connected. A connected, locally path-connected topological space, is path-connected. The connected components of a locally path-connected topological space are both open and closed, and are path connected.*

10.1.9 Limit of a sequence

Definitions

Definition 596 *A point $x \in E$ is an **accumulation point** (or cluster) of the sequence $(x_n)_{n \in \mathbb{N}}$ in the topological space (E, Ω) if for any neighborhood $n(x)$ and any N there is $p > N$ such that $x_p \in n(x)$*

A neighborhood of x contains infinitely many x_n

Definition 597 *A point $x \in E$ is a **limit** of the sequence $(x_n)_{n \in \mathbb{N}}$ in the topological space (E, Ω) if for any neighborhood $n(x)$ of x there is N such that $\forall n \geq N : x_n \in n(x)$. Then $(x_n)_{n \in \mathbb{N}}$ **converges** to x and this property is denoted $x = \lim_{n \to \infty} x_n$.*

There is a neighborhood of x which contains all the x_n for n>N

Definition 598 *A sequence $(x_n)_{n \in \mathbb{N}}$ in the topological space (E, Ω) is **convergent** if it admits at least one limit.*

So a limit is an accumulation point, but the converse is not aways true. And a limit is not necessarily unique.

Theorems

Theorem 599 *(Wilansky p.47) The limit of a convergent sequence in a Hausdorff topological space (E, Ω) is unique. Conversely if the limit of any convergent sequence in a topological space (E, Ω) is unique then (E, Ω) is Hausdorff.*

Theorem 600 *(Wilansky p.27) The limit (s) of a convergent sequence $(x_n)_{n \in \mathbb{N}}$ in the subset X of a topological space (E, Ω) belong to the closure of $X : \lim_{n \to \infty} x_n \in \overline{X}$. Conversely if the topological space (E, Ω) is first-countable then any point adherent to a subset X of E is the limit of a sequence in X.*

As a consequence :

Theorem 601 *A subset X of the topological space (E, Ω) is closed if the limit of any convergent sequence in X belongs to X*

This is the usual way to prove that a subset is closed. Notice that the condition is sufficient and not necessary if E is not first countable.

Theorem 602 *Weierstrass-Bolzano (Schwartz I p.75): In a compact topological space every sequence has a accumulation point.*

Theorem 603 *(Schwartz I p.77) A sequence in a compact topological space converges to a iff a is its unique accumulation point.*

10.1.10   Product topology

Definition

Theorem 604 *(Gamelin p.100) If $(E_i, \Omega_i)_{i \in I}$ is a family of topological spaces, the **product topology** on $E = \prod\limits_{i \in I} E_i$ is defined by the collection of open sets :*

If I is finite : $\Omega = \prod\limits_{i \in I} \Omega_i$

If I is infinite : $\Omega = \prod\limits_{i \in I} \varpi_i$, $\varpi_i \subset E_i$ such that $\exists J$ finite $\subset I : i \in J : \varpi_i \subset \Omega_i$

So the open sets of E are the product of a finite number of open sets, and the other components are any subsets.

The projections are the maps : $\pi_i : E \to E_i$

The product topology is the smallest Ω for which the projections are continuous maps

Theorems

Theorem 605 *(Gamelin p.100-103) If $(E_i)_{i \in I}$ is a family of topological spaces (E_i, Ω_i) and $E = \prod\limits_{i \in I} E_i$ their product endowed with the product topology, then:*

i) E is Hausdorff iff the (E_i, Ω_i) are Hausdorff

ii) E is connected iff the (E_i, Ω_i) are connected

iii) E is compact iff the (E_i, Ω_i) are compact (Tychonoff's theorem)

iv) If I is finite, then E is regular iff the (E_i, Ω_i) are regular

v) If I is finite, then E is normal iff the (E_i, Ω_i) are normal

vi) If I is finite, then a sequence in E is convergent iff each of its component is convergent

vii) If I is finite and the (E_i, Ω_i) are secound countable then E is secound countable

Theorem 606 *(Wilansky p.101) An uncountable product of non discrete space cannot be first countable.*

Remark : the topology defined by taking only products of open subsets in all E_i (called the box topology) gives too many open sets if I is infinite and the previous results are no longer true.

10.1.11   Quotient topology

Quotient spaces are very common. so it is very useful to understand how it works. An equivalence relation on a space E is just a partition of E, and the quotient set E'=E/$\sim$ the set of of its classes of equivalence (so each element is itself a subset). The key point is that E' is not necessarily Hausdorff, and it happens only if the classes of equivalence are closed subsets of E.

Definition

Definition 607 *(Gamelin p.105) Let (E, Ω) be a topological space, $\sim$ and an equivalence relation on E, $\pi : E \to E'$ the projection on the quotient set $E'=E/\sim$. The* **quotient topology** *on E' is defined by taking as open sets Ω' in $E' : \Omega' = \left\{ O' \subset E' : \pi^{-1}(O') \in \Omega \right\}$*

So π is continuous and this is the largest (meaning the largest Ω') topology for which π is continuous.

Theorems
   The property iv) is used quite often.

Theorem 608 *(Gamelin p.107) The quotient set E' of a topological space (E,Ω) endowed with the quotient topology is :*
   i) connected if E is connected
   ii) path-connected if E is path-connected
   iii) compact if E is compact
   iv) Hausdorff iff E is Hausdorff and each equivalence class is closed in E

Theorem 609 *(Gamelin p.105) Let (E, Ω) be a topological space, $E'=E/\sim$ the quotient set endowed with the quotient topology, $\pi : E \to E'$ the projection, F a topological space*
   i) a map $\varphi : E' \to F$ is continuous iff $\varphi \circ \pi$ is continuous
   ii) If a continuous map $f : E \to F$ is such that f is constant on each equivalence class, then there is a continuous map : $\varphi : E' \to F$ such that $f = \varphi \circ \pi$

A map $f : E \to F$ is called a quotient map if F is endowed with the quotient topology (Wilansky p.103).
   Let $f : E \to F$ be a continuous map between compact, Hausdorff, topological spaces E,F. Then $a \sim b \Leftrightarrow f(a) = f(b)$ is an equivalence relation over E and $E/\sim$ is homeomorphic to F.
   Remark : the quotient topology is the final topology with respect to the projection (see below).

10.2   Maps on topological spaces

10.2.1   Support of a function

Definition 610 *The* **support** *of the function $f : E \to K$ from a topological space (E,Ω) to a field K is the subset of E : Supp(f)=$\overline{\{x \in E : f(x) \neq 0\}}$ or equivalently the complement of the largest open set where f(x) is zero.*

Notation 611 *Supp(f) is the support of the function f. This is a closed subset of the domain of f*

Warning ! f(x) can be zero in the support, it is necessarily zero outside the support.

10.2.2 Continuous map

Definitions

Definition 612 *A map $f : E \to F$ between two topological spaces $(E, \Omega), (F, \Omega')$:*
i) ***converges*** *to $b \in F$ when x converges to $a \in E$ if for any open O' in F such that $b \in O'$ there is an open O in E such that $a \in O$ and $\forall x \in O : f(x) \in O'$*
ii).is ***continuous in a*** $\in E$ *if for any open O' in F such that $f(a) \in O'$ there is an open O in E such that $a \in O$ and $\forall x \in O : f(x) \in O'$*
iii) is ***continuous over a subset X of E*** *if it is continuous in any point of X*

f converges to a is denoted : $f(x) \to b$ when $x \to a$ or equivalently : $\lim_{x \to a} f(x) = b$
if f is continuous in a, it converges towards b=f(a), and conversely if f converges towards b then one can define by continuity f in a by f(a)=b.

Notation 613 $C_0(E; F)$ *is the set of continuous maps from E to F*

Continuity is completed by some definitions which are useful :

Definition 614 *A map $f : X \to F$ from a closed subset X of a topological space E to a topological space F is* ***semi-continuous*** *in $a \in \partial X$ if, for any open O' in F such that $f(a) \in O'$, there is an open O in E such that $a \in O$ and $\forall x \in O \cap X : f(x) \in O'$*

Which is, in the language of topology, the usual f$\to b$ when x$\to a_+$

Definition 615 *A map $f : E \to \mathbb{C}$ from a topological space (E, Ω) to $\mathbb{C}$* ***vanishes at infinity*** *if : $\forall \varepsilon > 0, \exists K$ compact $: \forall x \in K : |f(x)| < \varepsilon$*

Which is, in the language of topology, the usual f$\to 0$ when x$\to \infty$

Properties of continuous maps
1. Category

Theorem 616 *The composition of continuous maps is a continuous map*

if $f : E \to F$,$g : F \to G$ then $g \circ f$ is continuous.

Theorem 617 *The topological spaces and continuous maps constitute a category*

2. Continuity and convergence of sequences:

179

Theorem 618 *If the map $f : E \to F$ between two topological spaces is continuous in a, then for any sequence $(x_n)_{n \in \mathbb{N}}$ in E which converges to a : $f(x_n) \to f(a)$*

The converse is true only if E is first countable. Then f is continuous in a iff for any sequence $(x_n)_{n \in \mathbb{N}}$ in E which converges to a : $f(x_n) \to f(a)$.

3. Fundamental property of continuous maps:

Theorem 619 *The map $f : E \to F$ between two topological spaces is **continuous over E** if the preimage of any open subset of F is an open subset of E :*
$$\forall O' \in \Omega', f^{-1}(O') \in \Omega$$

4. Other properties of continuous maps :

Theorem 620 *If the map $f : E \to F$ between two topological spaces $(E, \Omega), (F, \Omega')$ is continuous over E then :*
i) if $X \subset E$ is compact in E, then f(X) is a compact in F
ii) if $X \subset E$ is connected in E, then f(X) is connected in F
iii) if $X \subset E$ is path-connected in E, then f(X) is path-connected in F
iv) if E is separable, then f(E) is separable
v) if Y is open in F, then $f^{-1}(Y)$ is open in E
vi) if Y is closed in F, then $f^{-1}(Y)$ is closed in E
vii) if X is dense in E and f surjective, then f(X) is dense in F
viii) the graph of f=$\{(x, f(x)), x \in E\}$ is closed in ExF

Theorem 621 *If $f \in C_0(E; \mathbb{R})$ and E is a non empty, compact topological space, then f has a maximum and a minimum.*

Theorem 622 *(Wilansky p.57) If $f, g \in C_0(E; F)$ E,F Hausdorff topological spaces and f(x)=g(x) for any x in a dense subset X of E, then f=g in E.*

Theorem 623 *(Gamelin p.100) If $(E_i)_{i \in I}$ is a family of topological spaces (E_i, Ω_i) and $E = \prod_{i \in I} E_i$ their product endowed with the product topology, then:*
i) The projections : $\pi_i : E \to E_i$ are continuous
ii) If F is a topological space, a map $\varphi : E \to F$ is continuous iff $\forall i \in I, \pi_i \circ \varphi$ is continuous

Theorem 624 *(Wilansky p.53) A map $f : E \to F$ between two topological spaces E,F is continuous iff $\forall X \subset E : f(\overline{X}) \subset \overline{f(X)}$*

Theorem 625 *(Wilansky p.57) The characteristic function of a subset which is both open and closed is continuous*

Algebraic topological spaces

Whenever there is some algebraic structure on a set E, and a topology on E, the two structures are said to be consistent is the operations defined over E in the algebraic structure are continuous. So we have topological groups, topological vector spaces,...which themselves define Categories.

Example : a group $(G, \cdot)$ is a topological group if : $\cdot : G \times G \to G$, $G \to G$ :: g^{-1} are continuous

10.2.3   Topologies defined by maps

Compact-open topology

Definition 626 *(Husemoller p.4) The **compact-open topology** on the set $C_0(E; F)$ of all continuous maps between two topological spaces (E, Ω) and (F, Ω') is defined by the base of open subsets : $\{\varphi : \varphi \in C_0(E; F), \varphi(K) \subset O'\}$ where K is a compact subset of E and O' an open subset of F.*

Weak, final topology

This is the implementation in topology of a usual mathematical trick : to pull back or to push forward a structure from a space to another. These two procedures are inverse from each other. They are common in functional analysis.

1. Weak topology:

Definition 627 *Let E be a set, Φ a family $(\varphi_i)_{i \in I}$ of maps : $\varphi_i : E \to F_i$ where (F_i, Ω_i) is a topological space. The **weak topology** on E with respect to Φ is defined by the collection of open subsets in E : $\Omega = \cup_{i \in I} \left\{ \varphi_i^{-1}(\varpi_i), \varpi_i \in \Omega_i \right\}$*

So the topology on $(F_i)_{i \in I}$ is "pulled-back" on E.

2. Final topology:

Definition 628 *Let F be a set, Φ a family $(\varphi_i)_{i \in I}$ of maps : $\varphi_i : E_i \to F$ where (E_i, Ω_i) is a topological space. The **final topology** on F with respect to Φ is defined the collection of open subsets in F : $\Omega' = \cup_{i \in I} \left\{ \varphi_i(\varpi_i), \varpi_i \in \Omega_i \right\}$*

So the topology on $(E_i)_{i \in I}$ is "pushed-forward" on F.

3. Continuity:

In both cases, this is the coarsest topology for which all the maps φ_i are continuous.

They have the universal property :

Weak topology : given a topological space G, a map $g : G \to E$ is continuous iff all the maps $\varphi_i \circ g$ are continuous (Thill p.251)

Final topology : given a topological space G, a map $g : F \to G$ is continuous iff all the maps $g \circ \varphi_i$ are continuous.

4. Convergence:

Theorem 629 *(Thill p.251) If E is endowed by the weak topology induced by the family $(\varphi_i)_{i \in I}$ of maps : $\varphi_i : E \to F_i$, a sequence $(x_n)_{n \in \mathbb{N}}$ in E converges to x iff $\forall i \in I : f_i(x_n) \to f(x)$*

5. Hausdorff property

Theorem 630 *(Wilansky p.94) The weak topology is Hausdorff iff Φ is separating over E.*

Which means $\forall x \neq y, \exists i \in I : \varphi_i(x) \neq \varphi_i(y)$

6. Metrizability:

Theorem 631 *(Wilansky p.94) The weak topology is semi-metrizable if Φ is a sequence of maps to semi-metrizable spaces. The weak topology is metrizable iff Φ is a sequence of maps to metrizable spaces which is separating over E*

10.2.4   Homeomorphism

Definition 632 *A **homeomorphism** is a bijective and continuous map $f : E \to F$ between two topological spaces E,F such that its inverse f^{-1} is continuous.*

Definition 633 *A **local homeomorphism** is a map $f : E \to F$ between two topological spaces E,F such that for each $a \in E$ there is a neighborhood n(a) and a neighborhood n(b) of b=f(a) and the restriction of $f : n(a) \to n(b)$ is a homemomorphism.*

The homeomorphisms are the isomorphisms of the category of topological spaces.

Definition 634 *Two topological spaces are **homeomorphic** if there is an homeomorphism between them.*

Homeomorphic spaces share the same topological properties. Equivalently a topological property is a property which is preserved by homeomorphism. Any property than can be expressed in terms of open and closed sets is topological. Examples : if E and F are homeomorphic, E is connected iff F is connected, E is compact iff F is compact, E is Hausdorff iff F is Hausdorff,...
Warning ! this is true for a global homeomorphism, not a local homeomorphism

Definition 635 *The topologies defined by the collections of open subsets Ω, Ω' on the same set E are **equivalent** if there is an homeomorphism between (E, Ω) and (E, Ω').*

So, for all topological purposes, it is equivalent to take (E, Ω) or (E, Ω')

Theorem 636 *(Wilansky p.83) If $f \in C_0(E; F)$ is one to one, E compact, F Hausdorff then f is a homeomorphism of E and f(E)*

Theorem 637 *(Wilansky p.68) Any two non empty convex open sets of $\mathbb{R}^m$ are homeomorphic*

10.2.5 Open and closed maps

It would be handy if the image of an open set by a map would be an open set, but this is the contrary which happens with a continuous map. This leads to the following definitions :

Definition 638 *A map $f : E \to F$ between two topological spaces is :*
    *an **open map**, if the image of an open subset is open*
    *a **closed map**, if the image of a closed subset is closed*

The two properties are distinct : a map can be open and not closed (and vice versa).
Every homeomorphism is open and closed.

Theorem 639 *(Wilansky p.58) A bijective map is open iff its inverse is continuous.*

Theorem 640 *The composition of two open maps is open; the composition of two closed maps is closed.*

Theorem 641 *(Schwartz II p.190) A local homeomorphism is an open map.*

Theorem 642 *A map $f : E \to F$ between two topological spaces is :*
    open iff $\forall X \subset E : f(\overset{\circ}{X}) \sqsubseteq (f(\overset{\circ}{X}))$
    closed iff $\forall X \subset E : \overline{f(X)} \subset f(\overline{X})$

Theorem 643 *(Wilansky p.103) Any continuous open surjective map $f : E \to F$ is a quotient map. Any continuous closed surjective map $f : E \to F$ is a quotient map.*

meaning that F has the quotient topology. They are the closest thing to a homeomorphism.

Theorem 644 *(Thill p.253) If $f : E \to F$ is a continuous closed map from a compact space E to a Hausdorff space, if f is injective f is an embedding, if is bijective f is a homeomorphism.*

10.2.6 Proper maps

This is the same purpose as above : remedy to the defect of continuous maps that the image of a compact space is compact.

Definition 645 *A map $f : E \to F$ between two topological spaces is a* ***proper map*** *(also called a compact map) is the preimage of a compact subset of F is a compact subset of E.*

Theorem 646 *A continuous map $f \in C_0(E; F)$ is proper if it is a closed map and the pre-image of every point in F is compact.*

Theorem 647 *Closed map lemma: Every continuous map $f \in C_0(E; F)$ from a compact space E to a Hausdorff space F is closed and proper.*

Theorem 648 *A continuous function between locally compact Hausdorff spaces which is proper is also closed.*

Theorem 649 *A topological space is compact iff the maps from that space to a single point are proper.*

Theorem 650 *If $f \in C_0(E; F)$ is a proper continuous map and F is a compactly generated Hausdorff space, then f is closed.*

this includes Hausdorff spaces which are either first-countable or locally compact

10.3 Metric and Semi-metric spaces

The existence of a metric on a set is an easy way to define a topology and, indeed, this is still the way it is taught usually. Anyway a metric brings more properties

10.3.1 Metric and Semi-metric spaces

Semi-metric, Metric

Definition 651 *A **semi-metric** (or pseudometric) on a set E is a map : $d : E \times E \to \mathbb{R}$ which is symmetric, positive and such that :*
$d(x,x)=0, \; \forall x, y, z \in E : d(x, z) \leq d(x, y) + d(y, z)$

Definition 652 *A **metric** on a set E is a definite positive semi-metric : $d(x, y) = 0 \Leftrightarrow x = y$*

Examples :
i) on a real vector space a bilinear definite positive form defines a metric :
$d(x, y) = g(x - y, x - y)$

ii) a real affine space whose underlying vector space is endowed with a bilinear definite positive form :

$$d\left(A,B\right) = g\left(\overrightarrow{AB}, \overrightarrow{AB},\right)$$

iii) on any set there is the discrete metric : d(x,y)=0 if x=y, d(x,y)=1 otherwise

Definition 653 *If the set E is endowed with a semi-metric d:*
  *a **Ball** is the set $B\left(a,r\right) = \{x \in E : d\left(a,x\right) < r\}$ with $r > 0$*
  *the **diameter** of a subset X of is diam $= \sup_{x,y \in X} d(x,y)$*
  *the **distance** between a subset X and a point a is : $\delta\left(a,X\right) = \inf d(x,a)_{x \in X}$*
  the distance between 2 subsets X,Y is : $\delta(A,B) = \inf d(x,y)_{x \in X, y \in Y}$

Definition 654 *If the set E is endowed with a semi-metric d, a subset X of E is:*
  ***bounded** if $\exists R \subset \mathbb{R} : \forall x,y \in X : d\left(x,y\right) \leq R \Leftrightarrow diam(X) < \infty$*
  ***totally bounded** if $\forall r > 0$ there is a finite number of balls of radius r which cover X.*

totally bounded $\Rightarrow$ bounded

Topology on a semi-metric space
  One of the key differences between semi metric and metric spaces is that a semi metric space is usually not Hausdorff.

  1. Topology :

Theorem 655 *A semi-metric on a set E induces a topology whose base are the open balls : $B\left(a,r\right) = \{x \in E : d\left(a,x\right) < r\}$ with r>0*

The open subsets of E are generated by the balls, through union and finite intersection.

Definition 656 *A **semi-metric space** (E,d) is a set E endowed with the topology denoted (E,d) defined by its semi-metric. It is a **metric space** if d is a metric.*

  2. Neighborhood:

Theorem 657 *A neighborhood of the point x of a semi-metric space (E,d) is any subset of E that contains an open ball B(x,r).*

Theorem 658 *(Wilansky p.19) If X is a subset of the semi-metric space (E,d), then $x \in \overline{X}$ iff $\delta(x,X)=0$*

  3. Equivalent topology: the same topology can be induced by different metrics, and conversely different metrics can induce the same topology.

Theorem 659 *(Gamelin p.27) The topology defined on a set E by two semi-metrics d,d' are equivalent iff the identity map $(E, d) \rightarrow (E, d')$ is an homeomorphism*

Theorem 660 *A semi-metric d induces in any subset X of E an equivalent topology defined by the restriction of d to X.*

Example : If d is a semi metric, $\min(d, 1)$ is a semi metric equivalent to d.

4. Base of the topology

Theorem 661 *(Gamelin p.72) A metric space is first countable*

Theorem 662 *(Gamelin p.24, Wilansky p.76) A metric or semi-metric space is separable iff it is second countable.*

Theorem 663 *(Gamelin p.23) A subset of a separable metric space is separable*

Theorem 664 *(Gamelin p.23) A totally bounded metric space is separable*

Theorem 665 *(Gamelin p.25) A compact metric space is separable and second countable*

Theorem 666 *(Wilansky p.128) A totally bounded semi-metric space is second countable and so is separable*

Theorem 667 *(Kobayashi I p.268) A connected, locally compact, metric space is second countable and separable*

4. Separability:

Theorem 668 *(Gamelin p.74) A metric space is a T4 topological space, so it is a normal, regular, T1 and Hausdorff topological space*

Theorem 669 *(Wilansky p.62) A semi-metric space is normal and regular*

5. Compactness

Theorem 670 *(Wilansky p.83) A compact subset of a semi-metric space is bounded*

Theorem 671 *(Wilansky p.127) A countably compact semi-metric space is totally bounded*

Theorem 672 *(Gamelin p.20) In a metric space E, the following properties are equivalent for any subset X of E :*
i) X is compact
ii) X is closed and totally bounded
iii) every sequence in X has an accumulation point (Weierstrass-Bolzano)
iv) every sequence in X has a convergent subsequence

Warning ! in a metric space a subset closed and bounded is not necessarily compact

Theorem 673 *Heine-Borel: A subset X of $\mathbb{R}^m$ is closed and bounded iff it is compact*

Theorem 674 *(Gamelin p.28) A metric space (E,d) is compact iff every continuous function $f : E \to \mathbb{R}$ is bounded*

6. Paracompactness:

Theorem 675 *(Wilansky p.193) A semi-metric space has a $\sigma-$locally finite base for its topology.*

Theorem 676 *(Bourbaki, Lang p.34) A metric space is paracompact*

6. Convergence of a sequence

Theorem 677 *A sequence $(x_n)_{n\in\mathbb{N}}$ in a semi-metric space (E,d) converges to the limit x iff $\forall \varepsilon > 0, \exists N \in \mathbb{N} : \forall n > N : d(x_n, x) < \varepsilon$*

Theorem 678 *(Schwartz I p.77) In a metric space a sequence is convergent iff it has a unique point of accumulation*

The limit is unique *if d is a metric.*

7.Product of semi-metric spaces:
There are different ways to define a metric on the product of a finite number of metric spaces $E = E_1 \times E_2 \times ... \times E_n$

The most usual ones for $x = (x_1, ..., x_n)$ are : the euclidean metric : $d(x,y) = \left(\sum_{i=1}^{n} d_i(x_i, y_i)^2 \right)$ and the max metric : $d(x,y) = \max d_i(x_i, y_i)$

With these metrics (E,d) is endowed with the product topology (cf.above).

Semi-metrizable and metrizable spaces
1. Definitions:

Definition 679 *A topological space (E, Ω) is said to be* **semi-metrizable** *if there is a semi-metric d on E such that the topologies $(E, \Omega), (E, d)$ are equivalent. A topological space (E, Ω) is said to be* **metrizable** *if there is a metric d on E such that the topologies $(E, \Omega), (E, d)$ are equivalent.*

2. Conditions for semi-metrizability:

Theorem 680 *Nagata-Smirnov(Wilansky p.198) A topological space is semi-metrizable iff it is regular and has a $\sigma-$locally finite base.*

Theorem 681 *Urysohn (Wilansky p.185): A second countable regular topological space is semi-metrizable*

Theorem 682 *(Wilansky p.186) A separable topological space is semi-metrizable iff it is second countable and regular.*

Theorem 683 *(Schwartz III p.428) A compact or locally compact topological space is semi-metrizable.*

Theorem 684 *(Schwartz III p.427) A topological space (E, Ω) is semi-metrizable iff :*
$$\forall a \in E, \forall n\,(a)\,, \exists f \in C_0(E; \mathbb{R}^+) : f(a) > 0, x \in n(a)^c : f(x) = 0$$

3. Conditions for metrizability:

Theorem 685 *(Wilansky p.186) A second countable T3 topological space is metrizable*

Theorem 686 *(Wilansky p.187) A compact Hausdorff space is metrizable iff it is second-countable*

Theorem 687 *Urysohn (Wilansky p.187) :A topological space is separable and metrizable iff it is T3 and second-countable.*

Theorem 688 *Nagata-Smirnov: A topological space is metrizable iff it is regular, Hausdorff and has a σ-locally finite base.*

A σ-locally finite base is a base which is a union of countably many locally finite collections of open sets.

Pseudo-metric spaces
Some sets (such that the Fréchet spaces) are endowed with a family of semi-metrics, which have some specific properties. In particular they can be Hausdorff.

1. Definition:

Definition 689 *A **pseudo-metric space** is a set endowed with a family $(d_i)_{i \in I}$ such that each d_i is a semi-metric on E and*
$$\forall J \subset I, \exists k \in I : \forall j \in J : d_k \geq d_j$$

2. Topology:

Theorem 690 *(Schwartz III p.426) On a pseudo-metric space $(E, (d_i)_{i \in I})$, the collection Ω of open sets*
$$O \in \Omega \Leftrightarrow \forall x \in O, \exists r > 0, \exists i \in I : B_i\,(x, r) \subset O \text{ where } B_i(a, r) = \{x \in E : d_i(a, x) < r\}$$
is the base for a topology.

Theorem 691 *(Schwartz III p.427) A pseudometric space $\left(E,(d_i)_{i \in I}\right)$ is Hausdorff iff $\forall x \neq y \in E, \exists i \in I : d_i(x,y) > 0$*

3. Continuity:

Theorem 692 *(Schwartz III p.440) A map $f : E \to F$ from a topological space (E,Ω) to a pseudo-metric space $\left(F,(d_i)_{i \in I}\right)$ is continuous at a$\in E$ if :*
$\forall \varepsilon > 0, \exists \varpi \in E : \forall x \in \varpi, \forall i \in I : d_i\left(f(x),f(a)\right) < \varepsilon$

Theorem 693 *Ascoli (Schwartz III p.450) A family $(f_k)_{k \in K}$ of maps : $f_k :$*
$E \to F$ from a topological space (E,Ω) to a pseudo-metric space $\left(F,(d_i)_{i \in I}\right)$
*is **equicontinuous** at a$\in E$ if : $\forall \varepsilon > 0, \exists \varpi \in \Omega : \forall x \in \varpi, \forall i \in I, \forall k \in K :$*
$d_i\left(f_k(x),f(a)\right) < \varepsilon$

Then the closure F of $(f_k)_{k \in K}$ in F^E (with the topology of simple convergence) is still equicontinuous at a.All maps in F are continuous at a, the limit of every convergent sequence of maps in F is continuous at a.

If a sequence $(f_n)_{n \in \mathbb{N}}$ of continuous maps on E, is equicontinuous and converges to a continuous map f on a dense subset of E, then it converges to f in E and uniformly on any compact of E.

4. Pseudo-metrizable topological space:

Definition 694 *A topological space (E,Ω) is pseudo-metrizable if it is homeomorphic to a space endowed with a family of pseudometrics*

Theorem 695 *(Schwartz III p.433) A pseudo-metric space $\left(E,(d_i)_{i \in I}\right)$ such that the set I is countable is metrizable.*

10.3.2 Maps on a semi-metric space

Continuity

Theorem 696 *A map $f : E \to F$ between semi-metric space (E,d),(F,d') is continuous in $a \in E$ iff $\forall \varepsilon > 0, \exists \eta > 0 : \forall d(x,a) < \eta, d'\left(f(x),f(a)\right) < \varepsilon$*

Theorem 697 *On a semi-metric space (E,d) the map $d : E \times E \to \mathbb{R}$ is continuous*

Uniform continuity

Definition 698 *A map $f : E \to F$ between the semi-metric spaces (E,d),(F,d')*
*is **uniformly continuous** if $\forall \varepsilon > 0, \exists \eta > 0 : \forall x,y \in E : d(x,y) < \eta, d'\left(f(x),f(y)\right) < \varepsilon$*

Theorem 699 *(Wilansky p.59) A map f uniformly continuous is continuous (but the converse is not true)*

Theorem 700 *(Wilansky p.219) A subset X of a semi-metric space is bounded iff any uniformly continuous real function on X is bounded*

Theorem 701 *(Gamelin p.26, Schwartz III p.429) A continuous map $f : E \to F$ between the semi-metric spaces E,F where E is compact if uniformly continuous*

Theorem 702 *(Gamelin p.27) On a semi-metric space (E,d), $\forall a \in E$ the map $d(a,.) : E \to \mathbb{R}$ is uniformly continuous*

Uniform convergence of sequence of maps

Definition 703 *The sequence of maps : $f_n : E \to F$ where (F,d) is a semi-metric space **converges uniformly** to $f : E \to F$ if :*
$$\forall \varepsilon > 0, \exists N \in \mathbb{N} : \forall x \in E, \forall n > N : d(f_n(x), f(x)) < \varepsilon$$

Convergence uniform $\Rightarrow$ convergence but the converse is not true

Theorem 704 *(Wilansky p.55) If the sequence of maps : $f_n : E \to F$,where E is a topological space and F is a semi-metric space, **converges uniformly** to f then :*
i) if the maps f_n are continuous at a, then f is continuous at a.
ii) If the maps f_n are continuous in E, then f is continuous in E

Lipschitz map

Definition 705 *A map $f : E \to F$ between the semi-metric spaces (E,d),(F,d')*
is
*i) a **globally Lipschitz** (or Hölder continuous) map of order a>0 if*
$\exists k \geq 0 : \forall x, y \in E : d'(f(x), f(y)) \leq k (d(x,y))^a$
*ii) a **locally Lipschitz** map of order a>0 if*
$\forall x \in E, \exists n(x), \exists k \geq 0 : \forall y \in n(x) : d'(f(x), f(y)) \leq k (d(x,y))^a$
*iii) a **contraction** if*
$\exists k, 0 < k < 1 : \forall x, y \in E : d(f(x), f(y)) \leq k d(x,y)$
*iv) an **isometry** if*
$\forall x, y \in E : d'(f(x), f(y)) = d(x,y)$

Embedding of a subset
It is a way to say that a subset contains enough information so that a function can be continuously extended from it.

Definition 706 *(Wilansky p.155) A subset X of a topological set E is said to be **C-embedded** in E if every continuous real function on X can be extended to a real continuous function on E.*

Theorem 707 *(Wilansky p.156) Every closed subset of a normal topological space E is C-embedded.*

Theorem 708 *(Schwartz 2 p.443) Let E be a metric space, X a closed subset of E, $f : X \to \mathbb{R}$ a continuous map on X, then there is a map $F : E \to \mathbb{R}$ continuous on E, such that : $\forall x \in X : F(x) = f(x), \sup_{x \in E} F(x) = \sup_{y \in X} f(y), \inf_{x \in E} F(x) = \inf_{y \in X} f(y)$*

10.3.3 Completeness

Completeness is an important property for infinite dimensional vector spaces as it is the only way to assure some fundamental results (such that the inversion of maps) through the fixed point theorem.

Cauchy sequence

Definition 709 *A sequence $(x_n)_{n \in \mathbb{N}}$ in a semi-metric space (E,d) is a **Cauchy sequence** if :*
$$\forall \varepsilon > 0, \exists N \in \mathbb{N} : \forall n, m > N : d(x_n, x_m) < \varepsilon$$

Any convergent sequence is a Cauchy sequence. But the converse is not always true.

Similarly a sequence of maps $f_n : E \to F$ where (F,d) is a semi-metric space, is a Cauchy sequence of maps if :
$$\forall \varepsilon > 0, \exists N \in \mathbb{N} : \forall x \in E, \forall n, m > N : d(f_n(x), f_m(x)) < \varepsilon$$

Theorem 710 *(Wilansky p.171) A Cauchy sequence which has a convergent subsequence is convergent*

Theorem 711 *(Gamelin p.22) Every sequence in a totally bounded metric space has a Cauchy subsequence*

Definition of complete semi-metric space

Definition 712 *A semi-metric space (E,d) is **complete** if any Cauchy sequence converges.*

Examples of complete metric spaces:
- Any finite dimensional vector space endowed with a metric
- The set of continuous, bounded real or complex valued functions over a metric space
- The set of linear continuous maps from a normed vector space E to a normed complete vector space F

Properties of complete semi-metric spaces

Theorem 713 *(Wilansky p.169) A semi-metric space is compact iff it is complete and totally bounded*

Theorem 714 *(Wilansky p.171) A closed subset of a complete metric space is complete. Conversely a complete subset of a metric space is closed.(untrue for semi-metric spaces)*

Theorem 715 *(Wilansky p.172) The countable product of complete spaces is complete*

Theorem 716 *(Schwartz I p.96) Every compact metric space is complete (the converse is not true)*

Theorem 717 *(Gamelin p.10) If $(f_n)_{n \in \mathbb{N}}$ is a Cauchy sequence of maps $f_n : E \to F$ in a complete matric space F, then there is a map : $f : E \to F$ such that f_n converges uniformly to f on E.*

Theorem 718 *Every increasing sequence on $\mathbb{R}$ with an upper bound converges
Every decreasing sequence on $\mathbb{R}$ with a lower bound converges*

Baire spaces

Theorem 719 *(Wilansky p.178) A complete semi metric space is a Baire space*

Theorem 720 *(Doob, p.4) If $f : X \to F$ is a uniformly continuous map on a dense subset X of a metric space E to a complete metric space F, then f has a unique uniformly continuous extension to E.*

Theorem 721 *Baire Category (Gamelin p.11): If $(X_n)_{n \in \mathbb{N}}$ is a family of dense open subsets of the complete metric space (E,d), then $\cap_{n=1}^{\infty} X_n$ is dense in E.*

Theorem 722 *A metric space (E,d) is complete iff every decreasing sequence of non-empty closed subsets of E, with diameters tending to 0, has a non-empty intersection.*

Fixed point

Theorem 723 *Contraction mapping theorem (Schwartz I p.101): If $f : E \to E$ is a contraction over a complete metric space then it has a unique **fixed point** a :$\exists a \in E : f(a) = a$
Furthermore if $f : E \times T \to E$ is continuous with respect to $t \in T$, a topological space, and
$\exists 1 > k > 0 : \forall x, y \in E, t \in T : d(f(x,t), f(y,t)) \leq k d(x,y)$
then there is a unique fixed point a(t) and $a : T \to E$ is continuous*

The point a can be found by iteration starting from any point b : $b_{n+1} = f(b_n) \Rightarrow a = \lim_{n \to \infty} b_n$ and we have the estimate : $d(b_n, a) \leq \frac{k^n}{k-1} d(b, f(b))$. So, if f is not a contraction, but if one of its iterated is a contraction, the theorem still holds.

This theorem is fundamental, for instance it is the key to prove the existence of solutions for differential equations, and it is one common way to compute solutions.

Theorem 724 *Brower: In $\mathbb{R}^n$, $n \geq 1$ any continuous map $f : B(0,1) \to B(0,1)$ (closed balls) has a fixed point.*

With the generalization : every continuous function from a convex compact subset K of a Banach space to K itself has a fixed point

Completion

Completeness *is not a topological property* : it is not preserved by homeomorphism. A topological space homeomorphic to a separable complete metric space is called a Polish space.

But a metric space which is not complete can be completed : it is enlarged so that, with the same metric, any Cauchy sequence converges.

Definition 725 *(Wilansky p.174) A completion of a semi-metric space (E,d) is a pair $(\overline{E}, \imath)$ of a complete semi-metric space $\overline{E}$ and an isometry $\imath$ from E to a dense subset of $\overline{E}$*

A completion of a metric space (E,d) is a pair $(\overline{E}, \imath)$ of a complete metric space $\overline{E}$ and an isometry $\imath$ from E to a dense subset of $\overline{E}$

Theorem 726 *(Wilansky p.175) A semi-metric space has a completion. A metric space has a completion, unique up to an isometry.*

The completion of a metric space $(\overline{E}, \imath)$ has the universal property that for any complete space (F,d') and uniformly continuous map $f : E \to F$ then there is a unique uniformly continuous function f' from $\overline{E}$ to F, which extends f.

The set of real number $\mathbb{R}$ is the completion of the set of rational numbers $\mathbb{Q}$. So $\mathbb{R}^n, \mathbb{C}^n$ are complete metric spaces for any fixed n, but not $\mathbb{Q}$.

If the completion procedure is applied to a normed vector space, the result is a Banach space containing the original space as a dense subspace, and if it is applied to an inner product space, the result is a Hilbert space containing the original space as a dense subspace.

10.4 Algebraic topology

Algebraic topology deals with the shape of objects, where two objects are deemed to have the same shape if one can pass from one to the other by a continuous deformation (so it is purely topological, without metric). The tools which have been developed for this purpose have found many other useful application in other fields. They highlight some fundamental properties of topological

spaces (topological invariants) so, whenever we look for some mathematical objects which "look alike" in some way, they give a quick way to restrict the scope of the search. For instance two manifolds which are not homotopic cannot be homeomorphic.

We will limit the scope at a short view of homotopy and covering spaces, with an addition for the Hopf-Rinow theorem. The main concept is that of simply connected spaces.

10.4.1 Homotopy

The basic idea of homotopy theory is that the kind of curves which can be drawn on a set, notably of loops, is in someway a characteristic of the set itself. It is studied by defining a group structure on loops which can be deformed continuously.

Homotopic paths
This construct is generalized below, but it is very common and useful to understand the concept.

1. A curve can be continuously deformed. Two curves are homotopic if they coincide in a continuous transformation. The precise definition is the following:

Definition 727 *Let (E, Ω) be a topological space, P the set P(a,b) of continuous maps $f \in C_0\left([0,1]\,;E\right) : f\left(0\right) = a, f\left(1\right) = b$*

*The paths $f_1, f_2 \in P\left(a,b\right)$ are **homotopic** if $\exists F \in C_0([0,1] \times [0,1]; E)$ such that :*

$\forall s \in [0,1] : F(s,0) = f_1(s), F(s,1) = f_2(s),$
$\forall t \in [0,1] : F(0,t) = a, F(1,t) = b$

2. $f_1 \sim f_2$ is an equivalence relation. It does not depend on the parameter :
$\forall \varphi \in C_0\left([0,1]\,;[0,1]\right), \varphi\left(0\right) = 0, \varphi\left(1\right) = 1 : f_1 \sim f_2 \Rightarrow f_1 \circ \varphi \sim f_2 \circ \varphi$
The quotient space P(a,b)/$\sim$ is denoted $[P\left(a,b\right)]$ and the classes of equivalences $[f]$.

3. Example : all the paths with same end points (a,b) in a convex subset of $\mathbb{R}^n$ are homotopic.

The key point is that not any curve can be similarly transformed in each other. In $\mathbb{R}^3$ curves which goes through a tore or envelop it are not homotopic.

Fundamental group
1. The set $[P\left(a,b\right)]$ is endowed with the operation $\cdot$:
If $a,b,c \in E, f \in P\left(a,b\right), g \in P\left(b,c\right)$ define the product $f \cdot g : P\left(a,b\right) \times P\left(b,c\right) \rightarrow P\left(a,c\right)$ by :
$0 \leq s \leq 1/2 \leq : f \cdot g(s) = f(2s),$
$1/2 \leq s \leq 1 : f \cdot g(s) = g(2s-1)$
This product is associative.
Define the inverse : $(f)^{-1}\left(s\right) = f\left(1 - s\right) \Rightarrow (f)^{-1} \circ f \in P\left(a,a\right)$

This product is defined over $[P(a,b)]$: If $f_1 \sim f_2$, $g_1 \sim g_2$ then $f_1 \cdot g_1 \sim f_2 \cdot g_2, (f_1)^{-1} \sim (f_2)^{-1}$

2. For homotopic loops:

Definition 728 *A **loop** is a path which begins and ends at the same point called **the base point**.*

The product of two loops with same base point is well defined, as is the inverse, and the identy element (denoted $[0]$) is the constant loop $f(t) = a$. So the set of loops with same base point is a group with $\cdot$. (it is not commutative).

Definition 729 *The **fundamental group** at a point a, denoted $\pi_1(E, a)$, of a topological space E, is the set of homotopic loops with base point a, endowed with the product of loops.*

$$\pi_1(E, a) = ([P(a, a)], \cdot)$$

3. Fundamental groups at different points are Isomorphic:
Let a,b$\in E$ such that there is a path f from a to b. Then :
$f_* : \pi_1(E, a) \rightarrow \pi_1(E, b) :: f_*([\gamma]) = [f] \cdot [\gamma] \cdot [f]^{-1}$
is a group isomorphism. So :

Theorem 730 *The fundamental groups $\pi_1(E, a)$ whose base point a belong to the same path-connected component of E are isomorphic.*

Definition 731 *The fundamental group of a path-connected topological space E, denoted $\pi_1(E)$, is the common group structure of its fundamental groups $\pi_1(E, a)$*

4. The fundamental group is a pure topological concept:

Theorem 732 *The fundamental groups of homeomorphic topological spaces are isomorphic.*

And this is a way to check the homeomorphism of topological spaces.
One of the consequences is the following :

Theorem 733 *(Gamelin p.123) If E,F are two topological spaces, $f : E \rightarrow F$ a homeomorphism such that $f(a) = b$, then there is an isomorphism $F : \pi_1(E, a) \rightarrow \pi_1(F, b)$*

Simply-connectedness

If $\pi_1(E) \sim [0]$ the group is said **trivial** : every loop can be continuously deformed to coincide with the point a.

Definition 734 *A path-connected topological group E is* **simply connected** *if its fundamental group is trivial :* $\pi_1(E) \sim [0]$

Roughly speaking a space is simply connected if there is no "hole" in it.

Definition 735 *A topological space E is* **locally simply connected** *if any point has a neighborhood which is simply connected*

Theorem 736 *(Gamelin p.121) The product of two simply connected spaces is simply connected*

Theorem 737 *A convex subset of* $\mathbb{R}^n$ *is simply connected. The sphere* S^n *(in* $\mathbb{R}^{n+1}$*) is simply connected for n>1 (the circle is not).*

Homotopy of maps

Homotopy can be generalized from paths to maps as follows:

Definition 738 *Two continuous maps* $f, g \in C_0(E; F)$ *between the topological spaces E,F are homotopic if there is a continuous map :* $F : E \times [0,1] \to F$ *such that :* $\forall x \in E : F(x,0) = f(x), F(x,1) = g(x)$

Homotopy of maps is an equivalence relation, which is compatible with the composition of maps.

Homotopy of spaces

1. Definition:

Definition 739 *Two topological spaces E,F are* **homotopic** *if there are maps* $f : E \to F, g : F \to E$*, such that* $f \circ g$ *is homotopic to the identity on E and* $g \circ f$ *is homotopic to the identity on F.*

Homeomorphic spaces are homotopic, but the converse is not always true.

Two spaces are homotopic if they can be transformed in each other by a continuous transformation : by bending, shrinking and expending.

Theorem 740 *If two topological spaces E,F are homotopic then if E is path-connected, F is path connected and their fundamental group are isomorphic. Thus if E is simply connected, F is simply connected*

The topologic spaces which are homotopic, with homotopic maps as morphisms, constitute a category.

2. Contractible spaces:

Definition 741 *A topological space is* **contractible** *if it is homotopic to a point*

The sphere is not contractible.

Theorem 742 *(Gamelin p.140) A contractible space is simply connected.*

3. Retraction of spaces:
More generally, a map $f \in C_0(E; X)$ between a topological space E and its subset X, is a continuous retract if $\forall x \in X : f(x) = x$ and then X is a retraction of E. E is retractible into X if there is a continuous retract (called a deformation retract) which is homotopic to the identity map on E.

If the subset X of the topological space E, is a continuous retraction of E and is simply connected, then E is simply connected.

Extension

Definition 743 *Two continuous maps $f, g \in C_0(E; F)$ between the topological spaces E,F are homotopic relative to the subset $X \subset E$ if there is a continuous map : $F : E \times [0, 1] \to F$ such that : $\forall x \in E : F(x, 0) = f(x), F(x, 1) = g(x)$ and $\forall t \in [0, 1], x \in X : F(x, t) = f(x) = g(x)$*

One gets back the homotopy of paths with E=$[0, 1]$, $X = \{a, b\}$.

This leads to the extension to homotopy of higher orders, by considering the homotopy of maps between n-cube $[0, 1]^r$ in $\mathbb{R}^r$ and a topological space E, with the fixed subset the boundary $\partial [0, 1]^r$ (all of its points such at least one $t_i = 0$ or 1). The homotopy groups of order $\pi_r(E, a)$ are defined by proceeding as above. They are abelian for r>1.

10.4.2   Covering spaces

A "fibered manifold" (see the Fibebundle part) is basically a pair of manifolds (M,E) where E is projected on M. Covering spaces can be seen as a generalization of this concept to topological spaces.

Definitions
  1. The definition varies according to the authors. This is the most general.

Definition 744 *Let $(E, \Omega), (M, \Theta)$ two topological spaces and a continuous surjective map : $\pi : E \to M$*

An open subset U of M is **evenly covered** *by E if :*

$\pi^{-1}(U)$ is the disjoint union of open subsets of E : $\pi^{-1}(U) = \cup_{i \in I} O_i; O_i \in \Omega; \forall i, j \in I : O_i \cap O_j = \varnothing$

and π is an homeomorphism on each $O_i \to \pi(O_i)$

The O_i are called the **sheets**. If U is connected they are the connected components of $\pi^{-1}(U)$

Definition 745 $E(M,\pi)$ *is a **covering space** if any point of M has a neighborhood which is evenly covered by E*

E is the **total space** , π the **covering map**, M the **base space**, $\pi^{-1}(x)$ the **fiber** over $x \in M$

Example : M=$\mathbb{R}, E = S_1$ the unit circle, $\pi : S_1 \to M :: \pi((\cos t, \sin t)) = t$

π is a local homeomorphism : each x in M has a neigborhood which is homeomorphic to a neighborhood $n\left(\pi^{-1}(x)\right)$. Thus E and M share all local topological properties : if M is locally connected so is E.

2. Order of a covering:

If M is connected every x in M has a neighborhood n(x) such that $\pi^{-1}(n(x))$ is homeomorphic to n(x)xV where V is a discrete space (Munkres). The cardinality of V is called the **degree r of the cover** : E is a double-cover of M if r=2. From the topological point of view E is r "copies" of M piled over M. This is in stark contrast with a fiber bundle E which is locally the "product" of M and a manifold V : so we can see a covering space a a fiber bundle with typical fiber a discrete space V (but of course the maps cannot be differentiable).

3. Isomorphims of fundamental groups:

Theorem 746 *Munkres: In a covering space $E(M,\pi),$ if M is connected and the order is r>1 then there is an isomorphism between the fundamental groups : $\widetilde{\pi} : \pi_1(E,a) \to \pi_1(M, \pi(a))$*

Fiber preserving maps

Definition 747 *A map : $f : E_1 \to E_2$ between two covering spaces $E_1(M,\pi_1), E_2(M,\pi_2)$ is **fiber preserving** if : $\pi_2 \circ f = \pi_1$*

$$
\begin{array}{ccccc}
E_1 & \to & f & \to & E_2 \\
\pi_1 & \searrow & & \swarrow & \pi_2 \\
& & E & &
\end{array}
$$

If f is an homeomorphisme then the covers are said **equivalent**.

Lifting property

1. Lift of a curve

Theorem 748 *(Munkres) If $\gamma : [0,1] \to M$ is a path then there exists a unique path $\Gamma : [0,1] \to E$ such that $\pi \circ \Gamma = \gamma$*

The path Γ is called the **lift** of γ.

If x and y are two points in E connected by a path, then that path furnishes a bijection between the fiber over x and the fiber over y via the lifting property.

2. Lift of a map

If $\varphi : N \to M$ is a continuous map in a simply connected topological space N, fix y$\in N, a \in \pi^{-1}(\varphi(a))$ in E, then there is a unique continuous map $\Phi : N \to E$ such that $\varphi = \pi \circ \Phi$.

Universal cover

Definition 749 *A covering space $E(M, \pi)$ is a **universal cover** if E is connected and simply connected*

If M is simply connected and E connected then π is bijective

The meaning is the following : let $E\prime(M, \pi')$ another covering space of M such that E' is connected. Then there is a map : $f : E \to E'$ such that $\pi = \pi' \circ f$

A universal cover is unique : if we fix a point x in M, there is a unique f such that $\pi(a) = x, \pi'(a') = x, \pi = \pi' \circ f$

10.4.3   Geodesics

This is a generalization of the topic studied on manifolds.

1. Let (E, d) be a metric space. A path on E is a continuous injective map from an interval $[a, b] \subset \mathbb{R}$ to E. If $[a, b]$ is bounded then the set $C_0([a, b]; E)$ is a compact connected subset. The curve generated by $p \in C_0([a, b]; E)$, denoted $p[a, b]$, is a connected, compact subset of E.

2. The **length of a curve** $p[a, b]$ is defined as : $\ell(p) = \sup \sum_{k=1}^{n} d(p(t_{k+1})), p(t_k))$ for any increasing sequence $(t_n)_{n \in \mathbb{N}}$ in [a,b]

The curve is said to be **rectifiable** if $\ell(p) < \infty$.

3. The length is inchanged by any change of parameter $p \to \widetilde{p} = p \circ \varphi$ where φ is order preserving.

The path is said to be at constant speed v if there is a real scalar v such that : $\forall t, t' \in [a, b] : \ell(p[t, t']) = v|t - t'|$

If the curve is rectifiable it is always possible to choose a path at constant speed 1 by :$\varphi(t) = \ell(p(t))$

4. A **geodesic** on E is a curve such that there is a path p$\in C_0(I; E)$, with I some interval of $\mathbb{R}$, such that :

$\forall t, t' \in I : d(p(t), p(t')) = |t' - t|$

5. A subset X is said **geodesically convex** if there is a geodesic which joins any pair of its points.

6. Define over E the new metric δ, called internally metric, by :

$\delta : E \times E \to \mathbb{R} :: \delta(x, y) = \inf_c \ell(c), p \in C_0([0, 1]; E) : p(0) = x, p(1) = y, \ell(c) < \infty$

$\delta \geq d$ and (E,d) is said to be an **internally metric** space if d=δ

A geodesically convex set is internally metric

A riemanian manifold is an internal metric space (with p$\in C_1([0,1]\,;E)$)

If $(E,d)\,,(F,d')$ are metric spaces and D is defined on ExF as

$D\left((x_1,y_1)\,,(x_2,y_2)\right) = \sqrt{d\left(x_1,x_2\right)^2 + d'\left(y_1,y_2\right)^2}$

then $(E \times F, D)$ is internally metric space iff E and F are internally metric spaces

A curve is a geodesic iff its projections are geodesic

7. The main result is the following:

Theorem 750 *Hopf-Rinow : If (E,d) is an internally metric, complete, locally compact space then:*

- every closed bounded subset is compact

- E is geodesically convex

Furthermore if, in the neighborhood of any point, any close curve is homotopic to a point (it is semi-locally simply connected) then every close curve is homotopic either to a point or a geodesic

It has been proven (Atkin) that the theorem is false for an infinite dimensional vector space (which is not, by the way, locally compact).

11 MEASURE

A measure is roughly the generalization of the concepts of "volume" or "surface" for a topological space. There are several ways to introduce measures :

    - the first, which is the most general and easiest to understand, relies on set functions So roughly a measure on a set E is a map $\mu : S \to \mathbb{R}$ where S is a set of subsets of E (a $\sigma-$algebra). We do not need a topology and the theory, based upon the ZFC model of sets, is quite straightforward. From a measure we can define integral $\int f\mu$, which are linear functional on the set C(E;$\mathbb{R}$).

    - the "Bourbaki way" goes the other way around, and is based upon Radon measures. It requires a topology, and, from my point of view, is more convoluted.

    So the will follow the first way. Definitions and results can be found in Doob and Schwartz (tome 2).

11.1 Measurable spaces

11.1.1 Limit of sequence of subsets

(Doob p.8)

Definition 751 *A sequence* $(A_n)_{n\in\mathbb{N}}$ *of subsets in E is :*
    monotone increasing if : $A_n \sqsubseteq A_{n+1}$
    monotone decreasing if : $A_{n+1} \sqsubseteq A_n$

Definition 752 *The **superior limit** of a sequence* $(A_n)_{n\in\mathbb{N}}$ *of subsets in E is the subset :*
    $\limsup A_n = \cap_{k=1}^{\infty} \cup_{j=k}^{\infty} A_n$

    It is the set of point in A_n for an infinite number of n

Definition 753 *The **inferior limit** of a sequence* $(A_n)_{n\in\mathbb{N}}$ *of subsets in E is the subset :*
    $\liminf A_n = \cup_{k=1}^{\infty} \cap_{j=k}^{\infty} A_n$

    It is the set of point in A_n for all but a finite number of n

Theorem 754 *Every sequence* $(A_n)_{n\in\mathbb{N}}$ *of subsets in E has a superior and an inferior limit and :*
    $\liminf A_n \sqsubseteq \limsup A_n$

Definition 755 *A sequence* $(A_n)_{n\in\mathbb{N}}$ *of subsets in E **converges** if its superior and inferior limit are equal and then its limit is:*
    $\lim_{n\to\infty} A_n = \limsup A_n = \liminf A_n$

Theorem 756 *A monotone increasing sequence of subsets converges to their union*

Theorem 757 *A monotone decreasing sequence of subsets converges to their intersection*

Theorem 758 *If B_p is a subsequence of a sequence $(A_n)_{n\in\mathbb{N}}$ then B_p converges iff $(A_n)_{n\in\mathbb{N}}$ converges*

Theorem 759 *If $(A_n)_{n\in\mathbb{N}}$ converges to A, then $(A_n^c)_{n\in\mathbb{N}}$ converges to A^c*

Theorem 760 *If $(A_n)_{n\in\mathbb{N}}, (B_n)_{n\in\mathbb{N}}$ converges respectively to A,B, then $(A_n\cup B_n)_{n\in\mathbb{N}}, (A_n\cap B_n)_{n\in\mathbb{N}}$ converges respectively to $A\cup B, A\cap B$*

Theorem 761 *$(A_n)_{n\in\mathbb{N}}$ converges to A iff the sequence of indicator functions $(1_{A_n})_{n\in\mathbb{N}}$ converges to 1_A*

Extension of $\mathbb{R}$
1. The compactification of $\mathbb{R}$ leads to define :
$\mathbb{R}_+ = \{r\in\mathbb{R}, r\geq 0\}, \overline{\mathbb{R}}_+ = \mathbb{R}_+\cup\{\infty\}, \overline{\mathbb{R}} = \mathbb{R}\cup\{+\infty, -\infty\}$
$\overline{\mathbb{R}}$ is compact .

2. Limit superior and limit inferior of a sequence:

Definition 762 *If $(x_n)_{n\in\mathbb{N}}$ is a sequence of real scalar on $\overline{\mathbb{R}}$*
*the **limit superior** is : $\limsup(x_n) = \lim_{n\to\infty}\sup_{p\geq n}(x_p)$*
*the **limit inferior** is :$\liminf(x_n) = \lim_{n\to\infty}\inf_{p\geq n}(x_p)$*

Theorem 763 $\liminf(x_n) \leq \limsup(x_n)$ *and are equal if the sequence converges (possibly at infinity).*

Warning ! this is different from the least upper bound :$\sup A = \min\{m\in E : \forall x\in E : m\geq x\}$ and the greatest lower bound $\inf A = \max\{m\in E : \forall x\in E : m\leq x\}$.

11.1.2 Measurable spaces

1. $\sigma-$algebras

Definition 764 *A collection S of subsets of E is an **algebra** if :*
$\varnothing\in S$
If $A\in S$ then $A^c\in S$ so $E\in S$
S is closed under finite union and finite intersection

Definition 765 *A $\sigma-$**algebra** is an algebra which contains the limit of any monotone sequence of its elements.*

The smallest $\sigma-$algebra is S=$\{\varnothing, E\}$, the largest is $S = 2^E$

2. Measurable space:

Definition 766 *A **measurable space** (E,S) is a set E endowed with a $\sigma-$algebra S. Every subset which belongs to S is said to be **measurable**.*

3. $\sigma-$algebras generated by a collection of sets:

Take any collection S of subsets of E, it is always possible to enlarge S in order to get a $\sigma-$algebra.

The smallest of the $\sigma-$algebras which include S will be denoted $\sigma(S)$.

If $(S_i)_{i=1}^n$ is a finite collection of subsets of 2^E then $\sigma(S_1 \times S_2.. \times S_n) = \sigma(\sigma(S_1) \times \sigma(S_2) \times ...\sigma(S_n))$

If (E_i, S_i) i=1..n are measurable spaces, then $(E_1 \times E_2 \times ...E_n, S)$ with $S = \sigma(S_1 \times S_2 \times ... \times S_n)$ is a measurable space

Warning ! $\sigma(S_1 \times S_2 \times ... \times S_n)$ is by far larger than $S_1 \times S_2 \times ... \times S_n$. If $E_1 = E_2 = \mathbb{R}$ S encompasses not only "rectangles" but almost any area in $\mathbb{R}^2$

4. Notice that in all these definitions there is no reference to a topology. However usually a $\sigma-$algebra is defined with respect to a given topology, meaning a collection of open subsets.

Definition 767 *A topological space (E, Ω) has a unique $\sigma-$algebra $\sigma(\Omega)$, called its **Borel algebra**, which is generated either by the open or the closed subsets.*

So a topological space can always be made a measurable space.

11.1.3 Measurable functions

A measurable function is different from an integrable function. They are really different concepts. Almost every map is measurable.

Defintion

Theorem 768 *If (F, S') a measurable space, f a map: $f : E \to F$ then the collection of subsets $(f^{-1}(A'), A' \in S')$ is a σ -algebra in E denoted $\sigma(f)$*

Definition 769 *A map $f : E \to F$ between the measurable spaces $(E, S), (F, S')$ is **measurable** if $\sigma(f) \sqsubseteq S$*

Definition 770 *A **Baire map** is a measurable map $f : E \to F$ between topological spaces endowed with their Borel algebras.*

Theorem 771 *Every continuous map is a Baire map.*

Properties
(Doob p.56)

Theorem 772 *The composed $f \circ g$ of measurable maps is a measurable map.*

The category of measurable spaces as for objects measurable spaces and for morphisms measurable maps

Theorem 773 *If $(f_n)_{n\in\mathbb{N}}$ is a sequence of measurable maps $f_n : E \to F$, with $(E, S), (F, S')$ measurable spaces, such that $\forall x \in E, \exists \lim_{n\to\infty} f_n(x) = f(x)$, then f is measurable*

Theorem 774 *If $(f_n)_{n\in\mathbb{N}}$ is a sequence of measurable functions : $f_n : E \to \overline{\mathbb{R}}$*
then the functions : $\limsup f_n = \inf_{j>i} \sup_{n>j} f_n ; \liminf_n f = \sup_{j>i} \inf f_n$
are measurable

Theorem 775 *If for i=1...n: $f_i : E \to F_i$ with $(E, S), (F_i, S'_i)$ measurables spaces then the map: $f = (f_1, f_2, ..f_n) : E \to F_1 \times F_2... \times F_n$ is measurable iff each f_i is measurable.*

Theorem 776 *If the map $f : E_1 \times E_2 \to F$, between measurable spaces is measurable, then for each x_1 fixed the map : $f_{x_1} : x_1 \times E_2 \to F :: f_{x_1}(x_2) = f(x_1, x_2)$ is measurable*

11.2 Measured spaces

A measure is a function acting on *subsets* : $\mu : S \to \mathbb{R}$ with some minimum properties.

11.2.1 Definition of a measure

Definition 777 *A function on a $\sigma-$algebra on a set E: $\mu : S \to \mathbb{R}$ is said :*
__I-subadditive__ if : $\mu(\cup_{i\in I} A_i) \le \sum_{i\in I} \mu(A_i)$ for any family $(A_i)_{i\in I}, A_i \in S$
of subsets in S
__I-additive__ if : $\mu(\cup_{i\in I} A_i) = \sum_{i\in I} \mu(A_i)$ for any family $(A_i)_{i\in I}, A_i \in S$ of disjointed subsets in S: $\forall i, j \in I : A_i \cap A_j = \varnothing$
__finitely subadditive__ if it is I-subadditive for any finite family
__finitely additive__ if it is I-additive for any finite family
__countably subadditive__ if it is I-subadditive for any countable family
__countably additive__ if it is I-additive for any countable family

Definition 778 *A __measure__ on the measurable space (E,S) is a map $\mu : S \to \overline{\mathbb{R}}_+$ which is countably additive. Then (E, S, μ) is a __measured space__.*

So a measure has the properties :
$\forall A \in S : 0 \le \mu(A) \le \infty$
$\mu(\varnothing) = 0$
For any countable *disjointed* family $(A_i)_{i\in I}$ of subsets in S : $\mu(\cup_{i\in I} A_i) = \sum_{i\in I} \mu(A_i)$ (possibly both infinite)
Notice that here a measure - without additional name - is always positive, but can take infinite value. It is necessary to introduce infinite value because the value of a measure on the whole of E is often infinite (think to the Lebesgue measure).

Definition 779 *A __Borel measure__ is a measure on a topological space with its Borel algebra.*

Definition 780 *A **signed-measure** on the measurable space (E,S) is a map $\mu : S \to \overline{\mathbb{R}}$ which is countably additive. Then (E, S, μ) is a **signed measured space**.*

So a signed measure can take negative value. Notice that a signed measure can take the values $\pm\infty$.

An outer **measure** on a set E is a map: $\lambda : 2^E \to \overline{\mathbb{R}}_+$ which is countably *subadditive,* monotone increasing and such that $\lambda(\varnothing) = 0$

So the key differences with a measure is that : there is no $\sigma-$algebra and λ is only countably subadditive (and not additive)

Finite measures

Definition 781 *A measure on E is **finite** if $\mu(E) < \infty$ so it takes only finite positive values : $\mu : S \to \mathbb{R}_+$*

A finite signed measure is a signed measure that takes only finite values : $\mu : S \to \mathbb{R}$

Definition 782 *A **locally finite measure** is a Borel measure which is finite on any compact.*

A finite measure is locally finite but the converse is not true.

Definition 783 *A measure on E is $\sigma-$**finite** if E is the countable union of subsets of finite measure. Accordingly a set is said to be $\sigma-$finite if it is the countable union of subsets of finite measure.*

Regular measure

Definition 784 *A Borel measure μ on a topological space E is*
   *   ***inner regular*** *if it is locally finite and $\mu(A) = \sup_K \mu(K)$ where K is a compact $K \sqsubseteq A$.*
   *   ***outer regular*** *if $\mu(A) = \inf_O \mu(O)$ where O is an open subset $A \sqsubseteq O$.*
   *   ***regular*** *if it is both inner and outer regular.*

Theorem 785 *(Thill p.254) An inner regular measure μ on a Hausdorff space such that $\mu(E) = 1$ is regular.*

Theorem 786 *(Neeb p.43) On a locally compact topological space, where every open subset is the countable union of compact subsets, every locally finite Borel measure is inner regular.*

11.2.2 Radon measures

Radon measures are a class of measures which have some basic useful properties and are often met in Functional Analysis.

Definition 787 *A **Radon measure** is a Borel, locally finite, regular,signed measure on a topological Hausdorff locally compact space*

So : if (E, Ω) is a topological Hausdorff locally compact space with its Borel algebra S, a Radon measure μ has the following properties :

it is locally finite : $\mu(K) < \infty$ for any compact K of E

it is regular :

$\forall X \in S : \mu(X) = \inf(\mu(Y), X \sqsubseteq Y, Y \in \Omega)$

$((\forall X \in \Omega) \vee (X \in S)) \& (\mu(X) < \infty) : \mu(X) = \sup(\mu(K), K \sqsubseteq X, K \text{ compact})$

The Lebesgue measure on $\mathbb{R}^m$ is a Radon measure.

Remark : There are other definitions : this one is the easiest to understand and use.

One useful theorem:

Theorem 788 *(Schwartz III p.452) Let (E, Ω) a topological Hausdorff locally compact space, $(O_i)_{i \in I}$ and open cover of E, $(\mu_i)_{i \in I}$ a family of Radon measures defined on each O_i. If on each non empty intersection $O_i \cap O_j$ we have $\mu_i = \mu_j$ then there is a unique measure μ defined on the whole of E such that $\mu = \mu_i$ on each O_i.*

11.2.3 Lebesgue measure

(Doob p.47)

So far measures have been reviewed through their properties. The Lebesgue measure is the basic example of a measure on the set of real numbers, and from there is used to compute integrals of functions. Notice that the Lebesgue measure is not, by far, the unique measure that can be defined on $\mathbb{R}$, but it has remarquable properties listed below.

Lebesgue measure on $\mathbb{R}$

Definition 789 *The **Lebesgue measure** on $\mathbb{R}$ denoted dx is the only complete, locally compact, translation invariant, positive Borel measure, such that $dx(]a, b]) = b - a$ for any interval in $\mathbb{R}$. It is regular and $\sigma-$finite.*

It is built as follows.

1. $\mathbb{R}$ is a metric space, thus a measurable space with its Borel algebra S

Let $F : \mathbb{R} \to \mathbb{R}$ be an increasing right continuous function, define $F(\infty) = \lim_{x \to \infty} F(x), F(-\infty) = \lim_{x \to -\infty} F(x)$

2. For any semi closed interval]a,b] define the set function :

$\lambda(]a, b]) = F(b) - F(a)$

then λ has a unique extension as a complete measure on $(\mathbb{R}, S)$ finite on compact subsets

3. Conversely if μ is a measure on $(\mathbb{R}, S)$ finite on compact subsets there is an increasing right continuous function F, defined up to a constant, such that : $\mu\left(]a, b]\right) = F(b) - F(a)$

4. If $F(x) = x$ the measure is the Lebesgue measure, also called the Lebesgue-Stieljes measure, and denoted dx. It is the usual measure on $\mathbb{R}$.

5. If μ is a probability then F is the distribution function.

Lebesgue measure on $\mathbb{R}^n$

The construct can be extended to $\mathbb{R}^n$:

1. for functions $F : \mathbb{R}^n \to \mathbb{R}$ define the operators
$D_j\left(]a, b]F\right)(x_1, x_2, ..x_{j-1}, x_{j+1}, ...x_n)$
$= F(x_1, x_2, ..x_{j-1}, b, x_{j+1}, ...x_n) - F(x_1, x_2, ..x_{j-1}, a, x_{j+1}, ...x_n)$

2. Choose F such that it is right continuous in each variable and :
$\forall a_j < b_j : \prod_{j=1}^{n} D_j\left(]a_j, b_j]F\right) \geq 0$

3. The measure of an hypercube is then defined as the difference of F between its faces.

Theorem 790 *The Lebesgue measure on $\mathbb{R}^n$ is the tensorial product $dx = dx_1 \otimes ... \otimes dx_n$ of the Lebesgue measure on each component x_k.*

So with the Lebesgue measure the measure of any subset of $\mathbb{R}^n$ which is defined as disjointed union of hypercubes can be computed. Up to a multiplicative constant the Lebesgue measure is "the volume" enclosed in an open subset of $\mathbb{R}^n$ To go further and compute the Lebesgue measure of any set on $\mathbb{R}^n$ the integral on manifolds is used.

11.2.4   Properties of measures

A measure is order preserving on subsets

Theorem 791 *(Doob p.18) A measure μ on a measurable space (E, S) is :*
i) countably subadditive:
$\mu\left(\cup_{i \in I} A_i\right) \leq \sum_{i \in I} \mu(A_i)$ for any countable family $(A_i)_{i \in I}$, $A_i \in S$ of subsets in S
ii) order preserving :
$A, B \in S, A \sqsubseteq B \Rightarrow \mu(A) \leq \mu(B)$
$\mu(\varnothing) = 0$

Extension of a finite additive function on an algebra:

Theorem 792 *Hahn-Kolmogorov (Doob p.40) There is a unique extension of a finitely-additive function $\mu_0 : S_0 \to \mathbb{R}_+$ on an algebra S_0 on a set E into a measure on $(E, \sigma(S_0))$.*

Value of a measure on a sequence of subsets

Theorem 793 *Cantelli (Doob p.26) For a sequence $(A_n)_{n\in\mathbb{N}}$ of subsets $A_n \in S$ of the measured space (E, S, μ) :*

i) $\mu\left(\liminf A_n\right) \leq \liminf \mu\left(A_n\right) \leq \limsup \mu\left(A_n\right)$

ii) if $\sum_n \mu\left(A_n\right) < \infty$ then $\mu\left(\limsup A_n\right) = 0$

iii) if μ is finite then $\limsup \mu\left(A_n\right) \leq \mu\left(\limsup A_n\right)$

Theorem 794 *(Doob p.18) For a map $\mu : S \to \mathbb{R}_+$ on a measurable space (E, S), the following conditions are equivalent :*

i) μ is a finite measure

ii) For any disjoint sequence $(A_n)_{n\in\mathbb{N}}$ in S : $\mu\left(\cup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu\left(A_n\right)$

iii) For any increasing sequence $(A_n)_{n\in\mathbb{N}}$ in S with $\lim A_n = A$: $\lim \mu\left(A_n\right) = \mu\left(A\right)$

iv) For any decreasing sequence $(A_n)_{n\in\mathbb{N}}$ in S with $\lim A_n = \varnothing$: $\lim \mu\left(A_n\right) = 0$

Tensorial product of measures

Theorem 795 *(Doob p.48) If $(E_i, S_i, \mu_i)_{i=1}^n$ are measured spaces and μ_i are $\sigma-$finite measures then there is a unique measure μ on (E,S) : $E = \prod_{i=1}^n E_i$, $S = \sigma\left(S_1 \times S_2.. \times S_n\right) = \sigma\left(\sigma\left(S_1\right) \times \sigma\left(S_2\right) \times ...\sigma\left(S_n\right)\right)$ such that : $\forall \left(A_i\right)_{i=1}^n, A_i \in S_i$: $\mu\left(\prod_{i=1}^n A_i\right) = \prod_{i=1}^n \mu_i\left(A_i\right)$*

μ is the **tensorial product of the measures** $\mu = \mu_1 \otimes \mu_2... \otimes \mu_n$ (also denoted as a product $\mu = \mu_1 \times \mu_2... \times \mu_n$)

Sequence of measures

Definition 796 *A sequence of measures or signed measures $(\mu_n)_{n\in\mathbb{N}}$ on the measurable space (E,S) converges to a limit μ if $\forall A \in S, \exists \mu\left(A\right) = \lim \mu_n\left(A\right)$.*

Theorem 797 *Vitali-Hahn-Saks (Doob p.30) The limit μ of a convergent sequence of measures $(\mu_n)_{n\in\mathbb{N}}$ on the measurable space (E,S) is a measure if each of the μ_n is finite or if the sequence is increasing.*

The limit μ of a convergent sequence of signed measures $(\mu_n)_{n\in\mathbb{N}}$ on the measurable space (E,S) is a signed measure if each of the μ_n is finite..

Pull-back, push forward of a measure

Definition 798 *Let (E_1, S_1), (E_2, S_2) be measurable spaces, $F : E_1 \to E_2$ a measurable map such that F^{-1} is measurable.*

*the **push forward** (or image) by F of the measure μ_1 on E_1 is the measure on (E_2, S_2) denoted $F_* \mu_1$ defined by : $\forall A_2 \in S_2 : F_* \mu_1 (A_2) = \mu_1 \left(F^{-1} (A_2) \right)$*

*the **pull back** by F of the measure μ_2 on E_2 is the measure on (E_1, S_1) denoted $F^* \mu_2$ defined by : $\forall A_1 \in S_1 : F^* \mu_2 (A_1) = \mu_2 (F (A_1))$*

Definition 799 *(Doob p.60) If $f_1, ... f_n : E \to F$ are measurable maps from the measured space (E, S, μ) into the measurable space (F,S') and f is the map $f : E \to F^n : f = (f_1, f_2, ..f_n)$ then $f_* \mu$ is called the **joint measure**. The i **marginal distribution** is defined as $\forall A' \in S' : \mu_i (A') = \mu \left(f_i^{-1} \left(\pi_i^{-1} (A') \right) \right)$ where $\pi_i : F^n \to F$ is the i projection.*

11.2.5 Almost eveywhere property

Definition 800 *A null set of a measured space (E, S, μ) is a set $A \in S : \mu (A) = 0$. A property which is satisfied everywhere in E but in a **null set** is said to be $\mu-$ everywhere satisfied (or almost everywhere satisfied).*

Definition 801 *The support of a Borel measure μ, denoted $Supp(\mu)$, is the complement of the union of all the null open subsets The support of a measure is a closed subset.*

Completion of a measure

It can happen that A is a null set and that $\exists B \subset A, B \notin S$ so B is not measurable.

Definition 802 *A measure is said to be **complete** if any subset of a null set is null.*

Theorem 803 *(Doob p.37) There is always a unique extension of the $\sigma-$ algebra S of a measured space such that the measure is complete (and identical for any subset of S).*

Notice that the tensorial product of complete measures is not necessarily complete

Applications to maps

Theorem 804 *(Doob p.57) If the maps $f, g : E \to F$ from the complete measured space (E, S, μ) to the measurable space (F,S') are almost eveywhere equal, then if f is measurable then g is measurable.*

Theorem 805 *Egoroff (Doob p.69) If the sequence $(f_n)_{n \in \mathbb{N}}$ of measurable maps $f_n : E \to F$ from the finite measured space (E, S, μ) to the metric space (F, d) is almost everywhere convergent in E to f, then $\forall \varepsilon > 0, \exists A_\varepsilon \in S, \mu(E \backslash A_\varepsilon) < \varepsilon$ such that $(f_n)_{n \in \mathbb{N}}$ is uniformly convergent in A_ε .*

Theorem 806 *Lusin (Doob p.69) For every measurable map $f : E \to F$ from a complete metric space (E, S, μ) endowed with a finite mesure μ to a metric space F, then $\forall \varepsilon > 0$ there is a compact A_ε , $\mu(E \backslash A_\varepsilon) < \varepsilon, A_\varepsilon$ such that f is continuous in A_ε .*

11.2.6   Decomposition of signed measures

Signed measures can be decomposed in a positive and a negative measure. Moreover they can be related to a measure (specially the Lebesgue measure) through a procedure similar to the differentiation.

Decomposition of a signed measure

Theorem 807 *Jordan decomposition (Doob p.145): If (E, S, μ) is a signed measure space,*
define : $\forall A \in S : \mu_+(A) = \sup_{B \subset A} \mu(B) ; \mu_-(A) = -\inf_{B \subset A} \mu(B)$
then :
i) μ_+, μ_- are positive measures on (E,S) such that $\mu = \mu_+ - \mu_-$
ii) μ_+ is finite if $\mu < \infty$,
iii) μ_- is finite if $\mu > -\infty$
*iv) $|\mu| = \mu_+ + \mu_-$ is a positive measure on (E,S) called the **total variation** of the measure*
v) If there are measures λ_1, λ_2 such that $\mu = \lambda_1 - \lambda_2$ then $\mu_+ \leq \lambda_1, \mu_- \leq \lambda_2$
vi) (Hahn decomposition) There are subsets E_+, E_- unique up to a null subset, such that :
$E = E_+ \cup E_- ; E_+ \cap E_- = \varnothing$
$\forall A \in S : \mu_+(A) = \mu(A \cap E_+), \mu_-(A) = \mu(A \cap E_-)$

The decomposition is not unique.

Complex measure

Theorem 808 *If μ, ν are signed measure on (E,S), then $\mu + i\nu$ is a measure valued in $\mathbb{C}$, called a **complex measure**.*
Conversely any complex measure can be uniquely decomposed as $\mu + i\nu$ where μ, ν are real signed measures.

Definition 809 *A signed or complex measure μ is said to be finite if $|\mu|$ is finite.*

Absolute continuity of a measure

Definition 810 *If λ is a positive measure on the measurable space (E,S), μ a signed measure on (E,S):*

*i) μ is **absolutely continuous** relative to λ if μ (or equivalently $|\mu|$) vanishes on null sets of λ.*

*ii) μ is **singular** relative to λ if there is a null set A for λ such that $|\mu|(A^c) = 0$*

iii) if μ is absolutely continuous (resp.singular) then μ_+, μ_- are absolutely continuous (resp.singular)

Thus with $\lambda = dx$ the Lebesgue measure, a singular measure can take non zero value for finite sets of points in $\mathbb{R}$. And an absolutely continuous measure is the product of a function and the Lebesgue measure.

Theorem 811 *(Doob p.147) A signed measure μ on the measurable space (E,S) is absolutely continuous relative to the finite measure λ on (E,S) iff :*
$\lim_{\lambda(A) \to 0} \mu(A) = 0$

Theorem 812 *Vitali-Hahn-Saks (Doob p.147) If the sequence $(\mu_n)_{n \in \mathbb{N}}$ of measures on (E,S), absolutely continuous relative to a finite measure λ, converges to μ then μ is a measure and it is also absolutely continuous relative to λ*

Theorem 813 *Lebesgue decomposition (Doob p.148) A signed measure μ on a measured space (E, S, λ) can be uniquely decomposed as : $\mu = \mu_c + \mu_s$ where μ_c is a signed measure absolutely continuous relative to λ and μ_s is a signed measure singular relative to λ*

Radon-Nikodym derivative

Theorem 814 *Radon-Nicodym (Doob p.150) For every finite signed measure μ on the finite measured space (E, S, λ), there is an integrable function $f : E \to \mathbb{R}$ uniquely defined up to null λ subsets, such that for the absolute continuous component μ_c of $\mu : \mu_c(A) = \int_A f\lambda$. For a scalar c such that $\mu_c \geq c\lambda$ (resp.$\mu_c \leq c\lambda$) then $f \geq c$ (resp.$f \leq c$) almost everywhere*

f is the **Radon-Nikodym derivative** (or density) of μ_c with respect to λ
There is a useful extension if $E = \mathbb{R}$:

Theorem 815 *(Doob p.159) Let λ, μ be locally finite measures on $\mathbb{R}, \lambda$ complete, a closed interval I containing x, then*

$\forall x \in \mathbb{R} : \varphi(x) = \lim_{I \to x} \frac{\mu(I)}{\lambda(I)}$ exists and is an integrable function on $\mathbb{R}$ almost λ everywhere finite

$\forall X \in S : \mu_c(X) = \int_X \varphi\lambda$ where μ_c is the absolutely continuous component of μ relative to λ

φ is denoted : $\varphi(x) = \frac{d\mu}{d\lambda}(x)$

11.2.7 Kolmogorov extension of a measure

These concepts are used in stochastic processes. The Kolmogorov extension can be seen as the tensorial product of an infinite number of measures.

Let (E,S,μ) a measured space and I any set. E^I is the set of maps : $\varpi : I \to E$. The purpose is to define a measure on the set E^I.

Any finite subset J of I of cardinality n can be written $J = \{j_1, j_2, ..j_n\}$

Define for $\varpi : I \to E$ the map: $\varpi_J : J \to E^n :: \varpi_J = (\varpi(j_1), \varpi(j_2), ..\varpi(j_n)) \in E^n$

For each n there is a $\sigma-$algebra : $S_n = \sigma(S^n)$ and for each $A_n \in S_n$ the condition $\varpi_J \in A_n$ defines a subset of E^I : all maps $\varpi \in E$ such that $\varpi_J \in A_n$. If, for a given J, A_n varies in S_n one gets an algebra Σ_J. The union of all these algebras is an algebra Σ_0 in E^I but usually not a $\sigma-$algebra. Each of its subsets can be expressed as the combination of Σ_J, with J finite.

However it is possible to get a measure on E^I : this is the Kolmogorov extension.

Theorem 816 *Kolmogorov (Doob p.61) If E is a complete metric space with its Borel algebra, $\lambda : \Sigma_0 \to \mathbb{R}_+$ a function countably additive on each Σ_J, then λ has an extension in a measure μ on $\sigma(\Sigma_0)$.*

Equivalently :

If for any finite subset J of n elements of I there is a finite measure μ_J on (E^n, S^n) such that :

$\forall s \in \mathfrak{S}(n), \mu_J = \mu_{s(J)}$: it is symmetric

$\forall K \subset I, card(K) = p < \infty, \forall A_j \in S : \mu_J(A_1 \times A_2.. \times A_n) \mu_K(E^p) = \mu_{J \cup K}(A_1 \times A_2.. \times A_n \times E^p)$

then there is a $\sigma-$algebra Σ, and a measure μ such that :

$\mu_J(A_1 \times A_2.. \times A_n) = \mu(A_1 \times A_2.. \times A_n)$

Thus if there are marginal measures μ_J, meeting reasonnable requirements, (E^I, Σ, μ) is a measured space.

11.3 Integral

Measures act on subsets. Integrals act on functions. Here integral are integral of *real functions* defined on a measured space. We will see later integral of r-forms on r-dimensional manifolds, which are a different breed.

11.3.1 Definition

Definition 817 *A **step function** on a measurable space (E,S) is a map : $f : E \to \mathbb{R}_+$ defined by a disjunct family $(A_i, y_i)_{i \in I}$ where $A_i \in S, y_i \in \mathbb{R}_+$: $\forall x \in E : f(x) = \sum_I y_i 1_{A_i}(x)$*

The integral of a step function on a measured space (E,S,μ) is : $\int_E f\mu = \sum_I y_i \mu(A_i)$

Definition 818 *The integral of a measurable positive function $f : E \to \overline{\mathbb{R}}_+$ on a measured space (E,S,μ) is :*

$\int_E f \mu = \sup \int_E g \mu$ *for all step functions g such that $g \le f$*

Any measurable function $f : E \to \overline{\mathbb{R}}$ can always be written as : $f = f_+ - f_-$ with $f_+, f_- : E \to \overline{\mathbb{R}}_+$ measurable such that they do not take ∞ values on the same set.

The integral of a measurable function $f : E \to \overline{\mathbb{R}}$ on a measured space (E,S,μ) is :

$\int_E f \mu = \int_E f_+ \mu - \int_E f_- \mu$

Definition 819 *A function $f : E \to \overline{\mathbb{R}}$ is **integrable** if $\left| \int_E f \mu \right| < \infty$ and $\int_E f \mu$ is the **integral** of f over E with respect to μ*

Notice that the integral can be defined for functions which take infinite values.

A function $f : E \to \overline{\mathbb{C}}$ is integrable iff its real part and its imaginary part are integrable and $\int_E f \mu = \int_E (\text{Re } f) \mu + i \int_E (\text{Im } f) \mu$

Warning ! μ is a real measure, and this is totally different from the integral of a function over a complex variable

The integral of a function on a measurable subset A of E is : $\int_A f \mu = \int_E f \times 1_A \mu$

The **Lebesgue integral** denoted $\int f dx$ is the integral with $\mu =$ the Lebesgue measure dx on $\mathbb{R}$.

Any **Riemann integrable** function is Lebesgue integrable, and the integrals are equal. But the converse is not true. A function is Rieman integrable iff it is continuous but for a set of Lebesgue null measure.

11.3.2   Properties of the integral

The spaces of integrable functions are studied in the Functional analysis part.

Theorem 820 *The set of real (resp.complex) integrable functions on a measured space (E,S,μ) is a real (resp.complex) vector space and the integral is a linear map.*

if f,g are integrable functions $f : E \to \overline{\mathbb{C}}$, a,b constant scalars then af+bg is integrable and $\int_E (af + bg) \mu = a \int_E f \mu + b \int_E g \mu$

Theorem 821 *If is an integrable function $f : E \to \overline{\mathbb{C}}$ on a measured space (E,S,μ) then : $\lambda(A) = \int_A f \mu$ is a measure on (E,S).*
If $f \ge 0$ and g is measurable then $\int_E g \lambda = \int_E g f \mu$

Theorem 822 *Fubini (Doob p.85) If $(E_1, S_1, \mu_1), (E_2, S_2, \mu_2)$ are $\sigma-$finite measured spaces, $f : E_1 \times E_2 \to \overline{\mathbb{R}}_+$ an integrable function on $(E_1 \times E_2, \sigma(S_1 \times S_2), \mu_1 \otimes \mu_2)$ then :*
i) for almost all $x_1 \in E_1$, the function $f(x_1, .) : E_2 \to \overline{\mathbb{R}}_+$ is μ_2 integrable

ii) $\forall x_1 \in E_1,$ *the function* $\int_{\{x_1\} \times E_2} f \mu_2 : E_1 \to \overline{\mathbb{R}}_+$ *is* μ_1 *integrable*

iii) $\int_{E_1 \times E_2} f \mu_1 \otimes \mu_2 = \int_{E_1} \mu_1 \left(\int_{\{x_1\} \times E_2} f \mu_2 \right) = \int_{E_2} \mu_2 \left(\int_{E_1 \times \{x_2\}} f \mu_1 \right)$

Theorem 823 *(Jensen's inequality) (Doob p.87)*
Let : $[a,b] \subset \mathbb{R}, \varphi : [a,b] \to \mathbb{R}$ *an integrable convex function, semi continuous in a,b,*
(E, S, μ) *a finite measured space, f an integrable function* $f : E \to [a,b],$
then $\varphi \left(\int_E f \mu \right) \le \int_E (\varphi \circ f) \mu$

The result holds if f,φ are not integrable but are lower bounded

Theorem 824 *If f is a function* $f : E \to \overline{\mathbb{C}}$ *on a measured space (E,S,μ) :*
i) If f$\ge$ 0 almost everywhere and $\int_E f \mu = 0$ *then f=0 almost everywhere*
ii) If f is integrable then $|f| < \infty$ *almost everywhere*
iii) if f$\ge$ 0, c $\ge$ 0 : $\int_E f \mu \ge c \mu (\{|f| \ge c\})$
iv) If f is measurable, $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ *monotone increasing,* $c \in \mathbb{R}_+$ *then*
$\int_E |f| \varphi \mu \ge \int_{|f| \ge c} \varphi (c) f \mu = \varphi (c) \mu (\{|f| \ge c\})$

Theorem 825 *(Lieb p.26) Let ν be a Borel measure on* $\overline{\mathbb{R}}_+$ *such that* $\forall t \ge 0 :$
$\phi (t) = \nu ([0,t)) < \infty,$
(E, S, μ) *a $\sigma-$finite measured space,* $f : E \to \mathbb{R}_+$ *integrable, then :*
$\int_E \phi (f(x)) \mu (x) = \int_0^\infty \mu (\{f(x) > t\}) \nu (t)$
$\forall p > 0 \in \mathbb{N} : \int_E (f(x))^p \mu (x) = p \int_0^\infty t^{p-1} \mu (\{f(x) > t\}) \nu (t)$
$f(x) = \int_0^\infty 1_{\{f > x\}} dx$

Theorem 826 *Beppo-Levi (Doob p.75) If* $(f_n)_{n \in \mathbb{N}}$ *is an increasing sequence of measurable functions* $f_n : E \to \overline{\mathbb{R}}_+$ *on a measured space (E,S,μ),which converges to f then :* $\lim_{n \to \infty} \int_E f_n \mu = \int_E f \mu$

Theorem 827 *Fatou (Doob p.82) If* $(f_n)_{n \in \mathbb{N}}$ *is a sequence of measurable functions* $f_n : E \to \overline{\mathbb{R}}_+$ *on a measured space (E,S,μ)*
and $f = \liminf f_n$ *then* $\int_E f \mu \le \liminf \left(\int_E f_n \mu \right)$
If the functions f, f_n are integrable and $\int_E f \mu = \lim_{n \to \infty} \int_E f_n \mu$
then $\lim_{n \to \infty} \int_E |f - f_n| \mu = 0$

Theorem 828 *Dominated convergence Lebesgue"s theorem (Doob p.83) If* $(f_n)_{n \in \mathbb{N}}$ *is a sequence of measurable functions* $f_n : E \to \overline{\mathbb{R}}_+$ *on a measured space (E,S,μ) if there is an integrable function g on* (E, S, μ) *such that* $\forall x \in E, \forall n :$
$|f_n (x)| \le g (x),$ *and* $f_n \to f$ *almost everywhere then :* $\lim_{n \to \infty} \int_E f_n \mu = \int_E f \mu$

11.3.3   Pull back and push forward of a Radon measure

This is the principle behind the change of variable in an integral.

Definition 829 *If μ is a Radon measure on a topological space E endowed with its Borel $\sigma-$algebra, a **Radon integral** is the integral* $\ell (\varphi) = \int \varphi \mu$ *for an integrable function :* $\varphi : E \to \overline{\mathbb{R}}.$ *ℓ is a linear functional on the functions* $C \left(E; \widetilde{\mathbb{R}} \right)$

The set of linear functional on a vector space (of functions) is a vector space which can be endowed with a norm (see Functional Analysis).

Definition 830 *Let E_1, E_2 be two sets, K a field and a map : $F : E_1 \to E_2$*

*The **pull back of a function** $\varphi_2 : E_2 \to K$ is the map : $F^* : C(E_2; K) \to C(E_1; K) :: F^* \varphi_2 = \varphi_2 \circ F$*

*The **push forward of a function** $\varphi_1 : E_1 \to K$ is the map $F_* : C(E_1; K) \to C(E_2; K) :: F_* \varphi_1 = F \circ \varphi_1$*

Theorem 831 *(Schwartz III p.535) Let $(E_1, S_1), (E_2, S_2)$ be two topological Hausdorff locally compact spaces with their Borel algebra, a continuous map $F : E_1 \to E_2$.*

i) let μ be a Radon measure in E_1, $\ell(\varphi_1) = \int_{E_1} \varphi_1 \mu$ be the Radon integral

*If F is a compact (proper) map, then there is a Radon measure on E_2, called the **push forward** of μ and denoted $F_* \mu$, such that :*

$\varphi_2 \in C(E_2; \mathbb{R})$ is $F_* \mu$ integrable iff $F^* \varphi_2$ is μ integrable and

$F_* \ell(\varphi_2) = \int_{E_2} \varphi_2 (F_* \mu) = \ell(F^* \varphi_2) = \int_{E_1} (F^* \varphi_2) \mu$

ii) let μ be a Radon measure in E_2, $\ell(\varphi_2) = \int_{E_2} \varphi_2 \mu$ be the Radon integral,

*If F is an open map, then there is a Radon measure on E_1, called the **pull back** of μ and denoted $F^* \mu$ such that :*

$\varphi_1 \in C(E_1; \mathbb{R})$ is $F^* \mu$ integrable iff $F_* \varphi_1$ is μ integrable and

$F^* \ell(\varphi_1) = \int_{E_1} \varphi_1 (F^* \mu) = \ell(F_* \varphi_1) = \int_{E_2} (F_* \varphi_2) \mu$

Moreover :

i) the maps F_*, F^* when defined, are linear on measures and functionals

ii) The support of the measures are such that : $Supp(F_* \ell) \subset F(Supp(\ell)), Supp(F^* \ell) \subset F^{-1}(Supp(\ell))$

iii) The norms of the functionals :$\|F_* \ell\| = \|\ell\| \le \infty, \|F^* \ell\| = \|\ell\| \le \infty$

iv) $F_* \mu, F^* \mu$ are positive iff μ is positive

v) If (E_3, S_3) is also a topological Hausdorff locally compact space and $G : E_2 \to E_3$, then, when defined :

$(F \circ G)_* \mu = F_* (G_* \mu)$

$(F \circ G)^* \mu = G^* (F^* \mu)$

If F is an homeomorphism then the push forward and the pull back are inverse operators :

$(F^{-1})^* \mu = F_* \mu, (F^{-1})_* \mu = F^* \mu$

Remark : the theorem still holds if E_1, E_2 are the countable union of compact subsets, F is measurable and μ is a positive finite measure. Notice that there are conditions attached to the map F.

Change of variable in a multiple integral

An application of this theorem is the change of variable in multiple integrals (in anticipation of the next part). The Lebesgue measure dx on $\mathbb{R}^n$ can be seen as the tensorial product of the measures $dx^k, k = 1...n$ which reads : $dx = dx^1 \otimes ... \otimes dx^n$ or more simply : $dx = dx^1..dx^n$ so that the integral $\int_U f dx$ of $f(x^1, ...x^n)$ over a subset U is by Fubini's theorem computed by taking

the successive integral over the variables $x^1, ...x^n$. Using the definition of the Lebesgue measure we have the following theorem.

Theorem 832 *(Schwartz IV p.71) Let U, V be two open subsets of $\mathbb{R}^n, F : U \rightarrow V$ a diffeomorphism, x coordinates in U, y coordinates in V, $y^i = F^i\left(x^1, ...x^n\right)$ then :*

$\varphi_2 \in C\left(V; \mathbb{R}\right)$ *is Lebesgue integrable iff $F^*\varphi_2$ is Lebesgue integrable and*
$\int_V \varphi_2\left(y\right) dy = \int_U \varphi_2\left(F\left(x\right)\right) \left|\det\left[F'\left(x\right)\right]\right| dx$
$\varphi_1 \in C\left(U; \mathbb{R}\right)$ *is Lebesgue integrable iff $F_*\varphi_1$ is Lebesgue integrable and*
$\int_U \varphi_1\left(x\right) dx = \int_V \varphi_1\left(F^{-1}\left(y\right)\right) \left|\det\left[F'\left(y\right)\right]^{-1}\right| dy$
So :
$F_* dx = dy = \left|\det\left[F'\left(x\right)\right]\right| dx$
$F^* dy = dx = \left|\det\left[F'\left(y\right)\right]^{-1}\right| dy$

This formula is the basis for any change of variable in a multiple integral. We use dx, dy to denote the Lebesgue measure for clarity but *there is only one measure on $\mathbb{R}^n$* which applies to different real scalar variables. For instance in $\mathbb{R}^3$ when we go from cartesian coordinates (the usual x,y,z) to spherical coordinates : $x = r\cos\theta\cos\varphi; y = r\sin\theta\cos\varphi; z = r\sin\varphi$ the new variables are real scalars (r, θ, φ) subject to the Lebesgue measure which reads $dr d\theta d\varphi$ and
$\int_U \varpi(x, y, z) dx dy dz = \int_{V(r,\theta,\varphi)} \varpi(r\cos\theta\cos\varphi, r\sin\theta\cos\varphi, r\sin\varphi) \left|r^2\cos\varphi\right| dr d\theta d\varphi$
Remark : the presence of the absolute value in the formula is due to the fact that the Lebesgue measure is positive : the measure of a set must stay positive when we use one variable or another.

11.4   Probability

Probability is a full branch of mathematics, which relies on measure theory, thus its place here.

Definition 833 *A **probability space** is a measured space (E, S, P) endowed with a measure P called a **probability** such that P(E)=1.*

So all the results above can be fully extended to a probability space, and we have many additional definitions and results. The presentation is limited to the basic concepts.

11.4.1   Definitions

Some adjustments in vocabulary are common in probability :

1. An element of a $\sigma-$algebra S is called an "**event**" : basically it represents the potential occurence of some phenomena. Notice that an event is usually not a single point in Ω, but a subset. A subset of S or a subalgebra of S can be seen as the "simultaneous" realization of events.

2. A measurable map $X : \Omega \rightarrow F$ with F usually a discrete space or a metric space endowed with its Borel algebra, is called a "**random variable**"

(or "stochastic variable"). So the events ϖ occur in Ω and the value X(ϖ) is in F.

3. two random variables X,Y are "**equal almost surely**" if $P\left(\{\varpi : X\left(\varpi\right) \neq Y\left(\varpi\right)\}\right) = 0$ so they are equal almost everywhere

4. If X is a real valued random variable :

its **distribution function** is the map : $F : \mathbb{R} \to [0,1]$ defined as :

$F\left(x\right) = P\left(\varpi \in \Omega : X\left(\varpi\right) \leq x\right)$

Its **expected value** is $E\left(X\right) = \int_{\Omega} XP$ this is its "average" value

its **moment of order r** is : $\int_{\Omega}\left(X - E(X)\right)^r P$ the moment of order 2 is the **variance**

the Jensen's inequality reads : for $1 \leq p : \left(E\left(|X|\right)\right)^p \leq E\left(|X|^p\right)$

and for X valued in [a,b], any function $\varphi : [a,b] \to \mathbb{R}$ integrable convex, semi continuous in a,b : $\varphi\left(E\left(X\right)\right) \leq E\left(\varphi \circ X\right)$

5. If $\Omega = \mathbb{R}$ then, according to Radon-Nikodym, there is a **density function** defined as the derivative relative to the Lebesgue measure :

$\rho\left(x\right) = \lim_{I \to x} \frac{P(I)}{dx(I)} = \lim_{h_1, h_2 \to 0_+} \frac{1}{h_1 + h_2}\left(F(x + h_1) - F(x - h_2)\right)$ where I is an interval containing I, $h_1, h_2 > 0$

and the absolutely continuous component of P is such that : $P_c\left(\varpi\right) = \int_{\varpi} \rho\left(x\right) dx$

11.4.2 Independant sets

Independant events

Definition 834 *The events $A_1, A_2, ...A_n \in S$ of a probability space (Ω,S,P) are **independant** if :*

$P\left(B_1 \cap B_2, .. \cap B_n\right) = P\left(B_1\right) P\left(B_2\right)..P\left(B_n\right)$ *where for any i : $B_i = A_i$ or $B_i = A_i^c$*

A family $\left(A_i\right)_{i \in I}$ of events are independant if any finite subfamily is independant

Definition 835 *Two $\sigma-$algebra S_1, S_2 are independant if any pair of subsets $\left(A_1, A_2\right) \in S_1 \times S_2$ are independant.*

If a collection of $\sigma-$algebras $\left(S_i\right)_{i=1}^n$ are independant then $\sigma\left(S_i \times S_j\right), \sigma\left(S_k \times S_l\right)$ are independant for i,j,k,l distincts

Conditional probability

Definition 836 *On a probability space (Ω,S,P), if $A \in S, P\left(A\right) \neq 0$ then $P\left(B|A\right) = \frac{P(B \cap A)}{P(A)}$ defines a new probability on (E,S) called **conditional probability** (given A).Two events are independant iff $P\left(B|A\right) = P\left(B\right)$*

Independant random variables

Definition 837 *Let (Ω,S,P) a probability space, (F,S') a measurable space, a family of random variables* $(X_i)_{i \in I}$, $X_i : E \to F$ *are **independant** if for any finite* $J \subset I$ *the* $\sigma-$*algebras* $(\sigma(X_j))_{j \in J}$ *are independant (remind that* $\sigma(X_j) = X_j^{-1}(S')$*)*

Equivalently :
$$\forall (A_j)_{j \in J}, A_j \in S', \ P\left(\cap_{j \in J} X_j^{-1}(A_j)\right) = \prod_{j \in J} P\left(X_j^{-1}(A_j)\right)$$

usually denoted : $P\left((X_j \in A_j)_{j \in J}\right) = \prod_{j \in J} P(X_j \in A)$

The 0-1 law

The basic application of the theorems on sequence of sets give the following theorem:

Theorem 838 *the 0-1 law: Let in the probability space (Ω,S,P):*

$(U_n)_{n \in \mathbb{N}}$ *an increasing sequence of* $\sigma-$*algebras of measurable subsets,*

$(V_n)_{n \in \mathbb{N}}$ *a decreasing sequence of* $\sigma-$*algebras of measurable subsets with* $V_1 \subset \sigma(\cup_{n \in \mathbb{N}} U_n)$

If, for each n, U_n, V_n *are independant, then* $\cap_{n \in \mathbb{N}} V_n$ *contains only null subsets and their complements*

Applications :

a) let the sequence of independant random variables $(X_n)_{n \in \mathbb{N}}$, $X_n \in \mathbb{R}$ take $U_n = \sigma(X_1, ...X_n), V_n = \sigma(X_{n+1}, ...)$

the series $\sum_n X_n$ converges either almost everywhere or almost nowhere

the random variables $\limsup \frac{1}{n} \left(\sum_{m=1}^n X_m\right), \liminf \frac{1}{n} \left(\sum_{m=1}^n X_m\right)$ are almost everywhere constant (possibly infinite). Thus :

Theorem 839 *On a probability space (Ω,S,P) for every sequence of independant random real variables* $(X_n)_{n \in \mathbb{N}}$*, the series* $\frac{1}{n} \left(\sum_{m=1}^n X_m\right)$ *converges almost everywhere to a constant or almost nowhere*

b) let (A_n) a sequence of Borel subsets in $\mathbb{R}$

$P(\limsup (X_n \in A_n)) = 0$ or 1 . This the probability that $X_n \in A_n$ infinitely often

$P(\liminf (X_n \in A_n)) = 0$ or 1. This is the probability that $X_n \in A_n^c$ only finitely often

11.4.3 Conditional expectation of random variables

The conditional probability is a measure acting on subsets. Similarly the conditional expectation of a random variable is the integral of a random variable using a conditional probablity.

Let (Ω,S,P) be a probability space and s a sub $\sigma-$algebra of S. Thus the subsets in s are S measurable sets. The restriction P_s of P to s is a finite measure on Ω.

Definition 840 *On a probability space (Ω,S,P), the **conditional expectation** of a random variable $X : \Omega \to F$ given a sub $\sigma-$algebra $s\subset S$ is a random variable $Y : \Omega \to F$ denoted E(X|s) meeting the two requirements:*
i) Y is s measurable and P_s integrable
ii) $\forall \varpi \in s : \int_{\varpi} Y P_s = \int_{\varpi} X P$

Thus X defined on (Ω,S,P) is replaced by Y defined on (Ω,s,P_s) with the condition that X,Y have the same expectation value on their common domain, which is s.

Y is not unique : any other function which is equal to Y almost everywhere but on P null subsets of s meets the same requirements.

With s=A it gives back the previous definition of P(B|A)=E(1_B|A)

Theorem 841 *(Doob p.183) If s is a sub $\sigma-$algebra of S on a probability space (Ω,S,P), we have the following relations for the conditional expectations of random variables $X,Z:\Omega \to F$*
i) If X=Z almost everywhere then E(X|s)=E(Z|s) almost everywhere
ii) If a,b are real constants and X,Z are real random variables : E(aX+bZ|s)=aE(X|s)+bE(Z|s)
iii) If $F = \mathbb{R} : X \leq Z \Rightarrow E(X|s) \leq E(Z|s)$ and $E(X|s) \leq E(|X||s)$
iv) if X is a constant function : $E(X|s) = X$
v) If $S'\subset S$ then $E(E(X|S')|S) = E(E(X|S)|S') = E(X|S')$

Theorem 842 *Bepo-Levi (Doob p.183) If s is a sub $\sigma-$algebra of S on a probability space (Ω,S,P) and $(X_n)_{n\in\mathbb{N}}$ an increasing sequence of positive random variables with integrable limit, then : $\lim E(X_n|s) = E(\lim X_n|s)$*

Theorem 843 *Fatou (Doob p.184) If s is a sub $\sigma-$algebra of S on a probability space (Ω,S,P) and $(X_n)_{n\in\mathbb{N}}$ is a sequence of positive integrable random variables with X=lim inf X_n integrable then :*
$E(X|s) \leq \liminf E(X_n|s)$ *almost everywhere*
$\lim E(|X - X_n||s) = 0$ *almost everywhere*

Theorem 844 *Lebesgue (Doob p.184) If s is a sub $\sigma-$algebra of S on a probability space (Ω,S,P) and $(X_n)_{n\in\mathbb{N}}$ a sequence of real random variables such that there is an integrable function g : with $\forall n, \forall x \in E : |X_n(x)| \leq g(x)$. If $X_n \to X$ almost everywhere then : $\lim E(X_n|s) = E(\lim X_n|s)$*

Theorem 845 *Jensen (Doob p.184) $[a,b] \subset \mathbb{R}, \varphi : [a,b] \to \mathbb{R}$ is an integrable convex function, semi continuous in a,b,*
X is a real random variable with range in [a,b] on a probability space $(\Omega,S,P),s$ is a sub $\sigma-$algebra of S
then $\varphi(E(X|s)) \leq E(\varphi(X)|s)$

11.4.4 Stochastic process

The problem

In a determinist process a variable X depending on time t is often known by some differential equation : $\frac{dX}{dt} = F(X, t)$ with the implication that the value of X at t depends on t and some initial value of X at t=0. But quite often in physics one meets random variables X depending on a time parameter. Thus there is no determinist rule for X(t) : even if X(0) is known the value of X(t) is still a random variable, with the complication that the probability law of X at a time t can depend on the value taken by X at a previous time t'.

Consider the simple case of coin tossing. For one shot the set of events is $\Omega = \{H, T\}$ (for "head" and "tail") with $H \cap T = \varnothing, H \cup T = E, P(H) = P(T) = 1/2$ and the variable is $X = 0, 1$. For n shots the set of events must be extended :$\Omega_n = \{HHT...T, ...\}$ and the value of the realization of the shot n is : $X_n \in \{0, 1\}$. Thus one can consider the family $(X_p)_{p=1}^n$ which is some kind of random variable, but the *set of events depend on n,* and *the probability law depends on n*, and could also depends on the occurences of the X_p if the shots are not independant.

Definition 846 *A **stochastic process** is a random variable $X = (X_t)_{t \in T}$ on a probability space (Ω, S, P) such that $\forall t \in T, X_t$ is a random variable on a probability space (Ω_t, S_t, P_t) valued in a measurable space (F,S') with $X_t^{-1}(F) = \Omega_t$*

i) T is any set ,which can be uncountable, but is well ordered so for any finite subspace J of T we can write : $J = \{j_1, j_2, .., j_n\}$

ii) so far no relation is assumed between the spaces $(\Omega_t, S_t, P_t)_{t \in T}$

iii) $X_t^{-1}(F) = E_t \Rightarrow P_t(X_t \in F) = 1$

If T is infinite, given each element, there is no obvious reason why there should be a stochastic process, and how to build E,S,P.

The measurable space (E,S)

The first step is to build a measurable space (E,S):

1. $E = \prod_{t \in T} E_t$ which always exists and is the set of all maps $\phi : T \to \cup_{t \in T} E_t$ such that $\forall t : \phi(t) \in E_t$

2. Let us consider the subsets of E of the type : $A_J = \prod_{t \in T} A_t$ where $A_t \in S_t$ and all but a finite number $t \in J$ of which are equal to E_t (they are called cylinders). For a given J and varying $(A_t)_{t \in J}$ in S_t we get a σ−algebra denoted S_J. It can be shown that the union of all these algebras for all finite J generates a σ−algebra S

3. We can do the same with F : define $A'_J = \prod_{t \in T} A'_t$ where $A'_t \in S'_t$ and all but a finite number $t \in J$ of which are equal to F. The preimage of such A'$_J$ by X is such that $X_t^{-1}(A'_t) \in S_t$ and for $t \in J : X_t^{-1}(F) = E_t$ so $X^{-1}(A'_J) = A_J \in S_J$. And the σ−algebra generated by the S'_J has a preimage in S. S is the smallest

$\sigma-$algebra for which all the $(X_t)_{t \in T}$ are simultaneouly measurable (meaning that the map X is measurable).

4. The next step, finding P, is less obvious. There are many constructs based upon relations between the (E_t, S_t, P_t), we will see some of them later. There are 2 general results.

The Kolmogov extension
This is one implementation of the extension presented above.

Theorem 847 *If all the Ω_t and F are complete metric spaces with their Borel algebras, and if for any finite subset J there are marginal probabilities P_J defined on (Ω, S_J) such that :*

$\forall s \in \mathfrak{S}(n), P_J = P_{s(J)}$ the marginal probabilities P_J do not depend on the order of J

$\forall J, K \subset I, card(J) = n, card(K) = p < \infty, \forall A_j \in S' :$
$P_J \left(X_{j_1}^{-1}(A_1) \times X_{j_2}^{-1}(A_2) .. \times X_{j_n}^{-1}(A_n) \right)$
$= P_{J \cup K} \left(X_{j_1}^{-1}(A_1) \times X_{j_2}^{-1}(A_2) .. \times X_{j_n}^{-1}(A_n) \times E^p \right)$
then there is a $\sigma-$algebra S on Ω, a probability P such that :
$P_J \left(X_{j_1}^{-1}(A_1) \times X_{j_2}^{-1}(A_2) .. \times X_{j_n}^{-1}(A_n) \right)$
$= P \left(X_{j_1}^{-1}(A_1) \times X_{j_2}^{-1}(A_2) .. \times X_{j_n}^{-1}(A_n) \right)$

These conditions are reasonable, notably in physics : if for any finite J, there is a stochastic process $(X_t)_{t \in J}$ then one can assume that the previous conditions are met and say that there is a stochastic process $(X_t)_{t \in T}$ with some probability P, usually not fully explicited, from which all the marginal probability P_J are deduced.

Conditional expectations
The second method involves conditional expectation of random variables.

Theorem 848 *Let J a finite subset of T. Consider $S_J \subset S, \Omega_J = \prod_{j \in I} \Omega_j$ and the map $X_J = (X_{j_1}, X_{j_2}, ..X_{j_n}) : \Omega_J \to F^J$*

If, for any J, there is a probability P_J on (Ω_J, S_J) and a conditional expectation $Y_J = E(X_J | S_J)$ then there is a probability on (Ω, S) such that :
$\forall \varpi \in S_J : \int_\varpi Y_J P_J = \int_\varpi X P$

This result if often presented (Tulcéa) with $T = \mathbb{N}$ and
$P_J = P(X_n = x_n | X_1 = x_1, ...X_{n-1} = x_{n-1})$ which are the transition probabilities.

11.4.5 Martingales

Martingales are classes of stochastic processes. They precise the relation between the probability spaces $(\Omega_t, S_t, P_t)_{t \in T}$

Definition 849 *Let (Ω,S) a measurable space, I an ordered set, and a map $I \to S$ where S_i is a $\sigma-$subalgebra of S such that : $S_i \sqsubseteq S_j$ whenever i<j. Then $\left(\Omega, S, (S_i)_{i \in I}, P\right)$ is a* **filtered probability space***.*

If $(X_i)_{i \in I}$ is a family of random variables $X_i : \Omega \to F$ such that each X_i is measurable in (Ω, S_j) it is said to be adapted and $\left(\Omega, S, (S_i)_{i \in I}, (X_i)_{i \in I}, P\right)$ is a filtered stochastic process.

Definition 850 *A filtered stochastic process is a* **Markov process** *if :*
$\forall i < j, A \subset F : P\left(X_j \in A | S_i\right) = P\left(X_j \in A | X_i\right)$ *almost everywhere*

So the probability at the step j depends only of the state X_i meaning the last one

Definition 851 *A filtered stochastic process is a* **martingale** *if $\forall i < j : X_i = E\left(X_j | S_i\right)$ almost everywhere*

That means that the future is totally conditionned by the past.
Then the function : $I \to F :: E(X_i)$ is constant almost everywhere
If $I = \mathbb{N}$ the condition $X_n = E\left(X_{n+1} | S_n\right)$ is sufficient
A useful application of the theory is the following :

Theorem 852 *Kolomogorov: Let (X_n) a sequence of independant real random variables on (Ω,S,P) with the same distribution law, then if X_1 is integrable :* $\lim_{n \to \infty} \left(\sum_{p=1}^{n} X_p\right) / n = E\left(X_1\right)$

.

12 BANACH SPACES

The combination of an algebraic structure and a topologic one on the same gives rise to newproperties. Topological groups are studied in the part "Lie groups". Here we study the other major algebraic structure : vector spaces, which include algebras.

A key feature of vector spaces is that all n-dimensional vector spaces are agebraically isomorphs and homeomorphs : all the topologies are equivalent and metrizable. Thus most of their properties stem from their algebraic structure. The situation is totally different for the infinite dimensional vector spaces. And the most useful of them are the complete normed vector spaces, called Banach spaces which are the spaces inhabited by many functions. Among them we have the Banach algebras and the Hilbert spaces.

On the topological and Banach vector spaces we follow Schwartz (t II).

12.1 Topological vector spaces

Topological vector spaces are the simplest of combinations : a vector space over a field K, endowed with a topological structure defined by a collection of open sets.

12.1.1 Definitions

Definition 853 *A **topological vector space** is a vector space endowed with a topology such that the operations (linear combination of vectors and scalars) are continuous.*

Theorem 854 *(Wilansky p.273, 278) A topological vector space is regular and connected*

Finite dimensional vector spaces

Theorem 855 *Every Hausdorff n-dimensional topological vector space over a field K is isomorphic (algebraically) and homeomorphic (topologically) to K^n.*

So on a finite dimensional Haussdorff topological space all the topologies are equivalent to the topology defined by a norm (see below) and are metrizable. In the following all n-dimensional vector spaces will be endowed with their unique normed topology if not stated otherwise. Conversely we have the fundamental theorem:

Theorem 856 *A Hausdorff topological vector space is finite-dimensional if and only if it is locally compact.*

And we have a less obvious result :

Theorem 857 *(Schwartz II p.97) If there is an homeomorphism between open sets of two finite dimensional vector spaces E,F on the same field, then dimE=dimF*

Vector subspace

Theorem 858 *A vector subspace F inherits a topological structure from the vector space E thus F is itself a topological vector space.*

Theorem 859 *A finite dimensional vector subspace F is always closed in a topological vector space E.*

Proof. A finite dimensional vector space is defined by a finite number of linear equations, which constitute a continuous map and F is the inverse image of 0. ■

If F is infinite dimensional it can be open or closed, or neither of the both.

Theorem 860 *If F is a vector subspace of E, then the quotient space E/F is Hausdorff iff F is closed in E. In particular E is Hausdorff iff the subset {0} is closed.*

This is the application of the general theorem on quotient topology.

Thus if E is not Hausdorff E can be replaced by the set E/F where F is the closure of {0}. For instance functions which are almost everywhere equal are taken as equal in the quotient space and the latter becomes Hausdorff.

Theorem 861 *(Wilansky p.274) The closure of a vector subspace is still a vector subspace.*

Bounded vector space

Without any metric it is still possible to define some kind of "bounded subsets". The definition is consistent with the usual one when there is a semi-norm.

Definition 862 *A subset X of a topological vector space over a field K is bounded if for any n(0) neighborhood of 0 there is $k \in K$ such that $X \subset kn(0)$*

Product of topological vector spaces

Theorem 863 *The product of topological vector spaces, endowed with its vector space structure and the product topology, is a topological vector space. It is still true with any (possibly infinite) product of topological vector spaces.*

This is the direct result of the general theorem on the product topology.

Example : the space of real functions : $f : \mathbb{R} \to \mathbb{R}$ can be seen as $\mathbb{R}^{\mathbb{R}}$ and is a topological vector space

Direct sum

The direct sum $\oplus_{i \in I} E_i$ (finite or infinite) of vector subspaces of a vector space E is algebraically isomorphic to their product $\widetilde{E} = \prod\limits_{i \in I} E_i$. Endowed with the product topology $\widetilde{E}$ is a topological vector space, and the projections π_i : $\widetilde{E} \to E_i$ are continuous. So the direct sum E is a topological vector space homeomorphic to $\widetilde{E}$.

This, obvious, result is useful because it is possible to part a vector space without any reference to a basis. A usual case if of a topological space which splits. Algebraically $E = E_1 \oplus E_2$ and it is isomorphic to $(E_1, 0) \times (0, E_2) \subset E \times E$. 0 is closed in E

Let $f : X \to E$ be a continuous map (not necessarily linear) from a topological space X to E.

$F : X \to E \times E :: F(x) = (f(x), f(x))$ is continuous, so are $\pi_i \circ F = f_i$: $X \to E_i$ and $f = f_1 + f_2$

12.1.2 Linear maps on topological vector spaces

The key point is that, in an infinite dimensional vector space, there are linear maps which are not continuous. So it is necessary to distinguish continuous linear maps, and this holds also for the dual space.

Continuous linear maps

Theorem 864 *A linear map $f \in L(E; F)$ is continuous if the vector spaces E,F are on the same field and finite dimensional.*

A multilinear map $f \in L^r(E_1, E_2, ..E_r; F)$ is continuous if the vector spaces $(E_i)_{i=1}^r$, F are on the same field and finite dimensional.

Theorem 865 *A linear map $f \in L(E; F)$ is continuous on the topological vector spaces E,F iff it is continuous at 0 in E.*

Theorem 866 *A multilinear map $f \in L^r(E_1, E_2, ..E_r; F)$ is continuous if it is continuous at $(0, .., 0)$ in $E_1 \times E_2 \times ..E_r$.*

Theorem 867 *The kernel of a linear map $f \in L(E; F)$ between topological vector space is either closed or dense in E. It is closed if f is continuous.*

Notation 868 $\mathcal{L}(E; F)$ *is the set of continuous linear map between topological vector spaces E,F on the same field*

$G\mathcal{L}(E; F)$ *is the set of continuous invertible linear map, with continuous inverse, between topological vector spaces E,F on the same field*

$\mathcal{L}^r(E_1, E_2, ...E_r; F)$ *is the set of continuous r-linear maps in $L^r(E_1, E_2, ...E_r; F)$*

Warning ! The inverse of an invertible continuous map is not necessarily continuous.

Compact maps

Compact maps (also called proper maps) are defined for any topological space, with the meaning that it maps compact sets to compact sets. However, because compact sets are quite rare on infinite dimensional vector spaces, the definition is extended as follows.

Definition 869 *(Schwartz 2 p.58) A linear map $f \in L(E;F)$ between topological vector spaces E,F is said to be* **compact** *if the closure $\overline{f(X)}$ in F of the image of a bounded subset X of E is compact in F.*

So compact maps "shrink" a set.

Theorem 870 *(Schwartz 2.p.59) A compact map is continuous.*

Theorem 871 *(Schwartz 2.p.59) A continuous linear map $f \in \mathcal{L}(E;F)$ between topological vector spaces E,F such that f(E) is finite dimensional is compact.*

Theorem 872 *(Schwartz 2.p.59) The set of compact maps is a subspace of $\mathcal{L}(E;F)$. It is a two-sided ideal of the algebra $\mathcal{L}(E;E)$*

Thus the identity map in $\mathcal{L}(E;E)$ is compact iff E is finite dimensional.

Theorem 873 *Riesz (Schwartz 2.p.66) : If $\lambda \neq 0$ is an eigen value of the compact linear endomorphism f on a topological vector space E, then the vector subspace E_λ of corresponding eigen vectors is finite dimensional.*

Dual vector space

As a consequence a linear form : $\varpi : E \rightarrow K$ is not necessarily continuous.

Definition 874 *The vector space of continuous linear form on a topological vector space E is called its* **topological dual**

Notation 875 *E' is the topological dual of a topological vector space E*

So $E^* = L(E;K)$ and $E' = \mathcal{L}(E;K)$
The topological dual E' is *included* in the algebraic dual E*, and they are identical iff E is finite dimensional.
The topological bidual $(E')'$ may be or not isomorphic to E if E is infinite dimensional.

Definition 876 *The map: $\imath : E \rightarrow (E')' :: \imath(u)(\varpi) = \varpi(u)$ between E and is topological bidual $(E')'$ is linear and injective.*
If it is also surjective then E is said to be **reflexive** *and $(E')'$ is isomorphic to E.*

The map $\imath$ is called the **evaluation map** is met quite often in this kind of problems.

Theorem 877 *The transpose of a linear continuous map : $f \in \mathcal{L}(E; F)$ is the continuous linear map : $f^t \in \mathcal{L}(F'; E') :: \forall \varpi \in F' : f'(\varpi) = \varpi \circ f$*

Proof. The transpose of a linear map $f \in L(E; F)$ is : $f^t \in L(F^*; E^*) :: \forall \varpi \in F^* : f^t(\varpi) = \varpi \circ f$

If f is continuous by restriction of F* to F' : $\forall \varpi \in F' : f'(\varpi) = \varpi \circ f$ is a continuous map ∎

Theorem 878 *Hahn-Banach (Bratelli 1 p.66) If K is a closed convex subset of a real locally convex topological Hausdorff vector space E, and $p \notin K$ then there is a continuous affine map : $f : E \to \mathbb{R}$ such that f(p)>1 and $\forall x \in K : f(x) \leq 1$*

This is one of the numerous versions of this theorem.

12.1.3   Tensor algebra

Tensor, tensor products and tensor algebras have been defined without any topology involved. All the definitions and results in the Algebra part can be fully translated by taking continuous linear maps (instead of simply linear maps).

Let be E,F vector spaces over a field K. Obviously the map $\imath : E \times F \to E \otimes F$ is continuous. So the universal property of the tensorial product can be restated as : for every topological space S and continuous bilinear map $f : E \times F \to S$ there is a unique continuous linear map : $\widehat{f} : E \otimes F \to S$ such that $f = \widehat{f} \circ \imath$

Covariant tensors must be defined in the topological dual E'. However the isomorphism between L(E;E) and $E \otimes E^*$ holds only if E is finite dimensional so, in general, $\mathcal{L}(E; E)$ is not isomorphic to $E \otimes E'$.

12.1.4   Affine topological space

Definition 879 *A **topological affine space** E is an affine space E with an underlying topological vector space $\overrightarrow{E}$ such that the map : $\to : E \times E \to \overrightarrow{E}$ is continuous.*

So the open subsets in an affine topological space E can be deduced by translation from the collection of open subsets at any given point of E.

An affine subspace is closed in E iff its underlying vector subspace is closed in $\overrightarrow{E}$. So :

Theorem 880 *A finite dimensional affine subspace is closed.*

Convexity plays an important role for topological affine spaces. In many ways convex subsets behave like compact subsets.

Definition 881 *A topological affine space (E, Ω) is **locally convex** if there is a base of the topology comprised of convex subsets.*

Such a base is a family C of open absolutely convex subsets ϖ containing a point O :

$\forall \varpi \in C, M, N \in \varpi, \lambda, \mu \in K : |\lambda| + |\mu| \leq 1 : \lambda M + \mu N \in \varpi$

and such that every neighborhood of O contains a element $k\varpi$ for some $k \in K, \varpi \in C$

A locally convex space has a family of pseudo-norms and conversely (see below).

Theorem 882 *(Berge p.262) The closure of a convex subset of a topological affine space is convex. The interior of a convex subset of a topological affine space is convex.*

Theorem 883 *Schauder (Berge p.271) If f is a continuous map f:$C \to C$ where C is a non empty compact convex subset of a locally convex affine topological space, then there is $a \in C : f(a) = a$*

Theorem 884 *An affine map f is continuous iff its underlying linear map $\overrightarrow{f}$ is continuous.*

Theorem 885 *Hahn-Banach theorem (Schwartz) : For every non empty convex subsets X,Y of a topological affine space E over $\mathbb{R}$, X open subset, such that $X \cap Y = \varnothing$, then there is a closed hyperplane H which does not meet X or Y.*

A hyperplane H in an affine space is defined by an affine scalar equation f(x)=0. If $f : E \to K$ is continuous then H is closed and $f \in E'$.

So the theorem can be restated :

Theorem 886 *For every non empty convex subsets X,Y of a topological affine space $\left(E, \overrightarrow{E}\right)$ over $\mathbb{C}$, X open subset, such that $X \cap Y = \varnothing$, there is a linear map $\overrightarrow{f} \in \overrightarrow{E'}$, $c \in \mathbb{R}$ such that for any $O \in E : \forall x \in X, y \in Y : \mathrm{Re}\, \overrightarrow{f}\left(\overrightarrow{Ox}\right) < c < \mathrm{Re}\, \overrightarrow{f}\left(\overrightarrow{Oy}\right)$*

12.2 Normed vector spaces

12.2.1 Norm on a vector space

A topological vector space can be endowed with a metric, and thus becomes a metric space. But an ordinary metric does not reflect the algebraic properties, so what is useful is a norm.

Definition 887 *A **semi-norm** on a vector space E over the field K (which is either $\mathbb{R}$ or $\mathbb{C}$) is a function :$\|\|: E \to \mathbb{R}_+$ such that :*

$\forall u, v \in E, k \in K :$
$\|u\| \geq 0;$
$\|ku\| = |k|\,\|u\|$ *where $|k|$ is either the absolute value or the module of k*
$\|u + v\| \leq \|u\| + \|v\|$

Definition 888 *A vector space endowed with a semi norm is a **semi-normed vector space***

Theorem 889 *A semi-norm is a continuous convex map.*

Definition 890 *A **norm** on a vector space E is a semi-norm such that : $\|u\| = 0 \Rightarrow u = 0$*

Definition 891 *If E is endowed with a norm $\|\|$ it is a **normed vector space** $(E, \|\|)$*

The usual norms are :
i) $\|u\| = \sqrt{g(u,u)}$ where g is a definite positive symmetric (or hermitian) form
ii) $\|u\| = \max_i |u_i|$ where u_i are the components relative to a basis
iii) $\|k\| = |k|$ is a norm on K with its vector space structure.
iv) On $\mathbb{C}^n$ we have the norms :
$\|X\|_p = \sum_{k=1}^n c_k |x_k|^p$ for p>0$\in \mathbb{N}$
$\|X\|_\infty = \sup_{k=1..n} |x_k|$
with the fixed scalars : $(c_k)_{k=1}^n, c_k > 0 \in \mathbb{R}$
The **inequalities of Hölder-Minkovski** give :
$\forall p \geq 1 : \|X + Y\|_p \leq \|X\|_p + \|Y\|_p$
and if p<∞ then $\|X + Y\|_p = \|X\|_p + \|Y\|_p \Rightarrow \exists a \in \mathbb{C} : Y = aX$

12.2.2 Topology on a semi-normed vector space

A semi-norm defines a semi-metric by : $d(u,v) = \|u - v\|$ but the converse is not true. There are vector spaces which are metrizable but not normable (see Fréchet spaces). So every result and definition for semi-metric spaces hold for semi-normed vector space.

Theorem 892 *A semi-norm (resp.norm) defines by restriction a semi-norm (resp.norm) on every vector subspace.*

Theorem 893 *On a vector space E two semi-norms $\|\|_1, \|\|_2$ are **equivalent** if they define the same topology. It is necessary and sufficient that : $\exists k, k' > 0 : \forall u \in E : \|u\|_1 \leq k \|u\|_2 \Leftrightarrow \|u\|_2 \leq k' \|u\|_1$*

Proof. The condition is necessary. If $B_1(0,r)$ is a ball centered at 0, open for the topology 1, and if the topology are equivalent then there is ball $B_2(0, r_2) \subset B_1(0, r)$ so $\|u\|_2 \leq r_2 \Rightarrow \|u\|_1 \leq r = kr_2$. And similarly for a ball $B_2(0, r)$.
The condition is sufficient. Every ball $B_1(0, r)$ contains a ball $B_2\left(0, \frac{r}{k'}\right)$ and vice versa. ∎
The theorem is still true for norms.

Theorem 894 *On a finite dimensional vector space all the norms are equivalent.*

Theorem 895 *The product $E=\prod_{i\in I} E_i$ of a finite number of semi-normed vector spaces on a field K is still a semi-normed vector space with one of the equivalent semi-norm :*

$\|\|\|_E = \max \|\|\|_{E_i}$

$\|\|\|_E = \left(\sum_{i\in I} \|\|\|_{E_i}^p\right)^{1/p}, 1 \le p < \infty$

The product of an infinite number of normed vector spaces is not a normable vector space.

Theorem 896 *(Wilansky p.268) Every first countable topological vector space is semi-metrizable*

Theorem 897 *A topological vector space is normable iff it is Hausdorff and has a convex bounded neighborhood of 0.*

Theorem 898 *(Schwartz I p.72) A subset of a finite dimensional vector space is compact iff it is bounded and closed.*

Warning ! This is false in an infinite dimensional normed vector space.

Theorem 899 *(Wilansky p.276) If a semi-normed vector space has a totally bounded neighborhood of 0 it has a dense finite dimensional vector subspace.*

Theorem 900 *(Wilansky p.271) A normed vector space is locally compact iff it is finite dimensional*

12.2.3 Linear maps

The key point is that a norm can be assigned to every continuous linear map.

Continuous linear maps

Theorem 901 *If E,F are semi-normed vector spaces on the same field, an $f\in L(E;F)$ then the following are equivalent:*
i) f is continuous
ii) $\exists k \ge 0 : \forall u \in E : \|f(u)\|_F \le k \|u\|_E$
iii) f is uniformly continuous and globally Lipschitz of order 1

So it is equivalently said that f is bounded.

Theorem 902 *Every linear map $f\in L(E;F)$ from a finite dimensional vector space E to a normed vector space F, both on the same field, is uniformly continuous and Lipschitz of order 1*

If E,F are semi-normed vector spaces on the same field f is said to be "bounded below" if : $\exists k \ge 0 : \forall u \in E : \|f(u)\|_F \ge k \|u\|_E$

Space of linear maps

Theorem 903 *The space $\mathcal{L}$(E;F) of continuous linear maps on the semi-normed vector spaces E,F on the same field is a semi-normed vector space with the semi-norm :* $\|f\|_{\mathcal{L}(E;F)} = \sup_{\|u\|\neq 0} \frac{\|f(u)\|_F}{\|u\|_E} = \sup_{\|u\|_E=1} \|f(u)\|_F$

Theorem 904 *The semi-norm $\|\|_{\mathcal{L}(E;F)}$ has the following properties :*
  i) $\forall u \in E : \|f(u)\| \leq \|f\|_{\mathcal{L}(E;F)} \|u\|_E$
  ii) If E=F $\|Id\|_E = 1$
  iii) (Schwartz I p.107) In the composition of linear continuous maps : $\|f \circ g\| \leq \|f\| \|g\|$
  iv) If $f \in \mathcal{L}(E;E)$ *then its iterated* $f^n \in \mathcal{L}(E;E)$ *and* $\|f^n\| = \|f\|^n$

Dual

Theorem 905 *The topological dual E' of the semi-normed vector spaces E is a semi-normed vector space with the semi-norm :* $\|f\|_{E'} = \sup_{\|u\|\neq 0} \frac{|f(u)|}{\|u\|_E} = \sup_{\|u\|_E=1} |f(u)|$

This semi-norm defines a topology on E' called the **strong topology**.

Theorem 906 *Banach lemna (Taylor 1 p.484): A linear form $\varpi \in F^*$ on a a vector subspace F of a semi-normed vector space E on a field K, such that on :$\forall u \in F$ $|\varpi(u)| \leq \|u\|$ can be extended in a map $\widetilde{\varpi} \in E'$ such that $\forall u \in E$:* $|\widetilde{\varpi}(u)| \leq \|u\|$

The extension is not necessarily unique. It is continuous. Similarly :

Theorem 907 *Hahn-Banach (Wilansky p.269): A linear form $\varpi \in F'$ continuous on a vector subspace F of a semi-normed vector space E on a field K can be extended in a continuous map $\widetilde{\varpi} \in E'$ such that* $\|\widetilde{\varpi}\|_{E'} = \|\varpi\|_{F'}$

Definition 908 *In a semi normed vector space E a **tangent functional** at $u \in E$ is a 1 form $\varpi \in E' : \varpi(u) = \|\varpi\| \|u\|$*

Using the Hahan-Banach theorem one can show that there are always non unique tangent functionals.

Multilinear maps

Theorem 909 *If $(E_i)_{i=1}^r$,F are semi-normed vector spaces on the same field, and f$\in L^r(E_1, E_2, ..E_r; F)$ then the following are equivalent:*
  i) f is continuous
  ii) $\exists k \geq 0 : \forall (u_i)_{i=1}^r \in E : \|f(u_1,..,u_r)\|_F \leq k \prod_{i=1}^r \|u_i\|_{E_i}$

Warning ! a multilinear map is never uniformly continuous.

Theorem 910 *If $(E_i)_{i=1}^r$,F are semi-normed vector spaces on the same field, the vector space of continuous r linear maps $f \in L^r(E_1, E_2, ..E_r; F)$ is a semi-normed vector space on the same field with the norm :*

Theorem 911 $\|f\|_{\mathcal{L}^r} = \sup_{\|u\|_i \neq 0} \frac{\|f(u_1,..,u_r)\|_F}{\|u_1\|_1...\|u_r\|_r} = \sup_{\|u_i\|_{E_i}=1} \|f(u_1,..,u_r)\|_F$

So : $\forall (u_i)_{i=1}^r \in E : \|f(u_1,..,u_r)\|_F \leq \|f\|_{\mathcal{L}^r} \prod_{i=1}^r \|u_i\|_{E_i}$

Theorem 912 *(Schwartz I p.119) If E,F are semi-normed spaces,the map : $\mathcal{L}(E.F) \times E \to F :: \varphi(f,u) = f(u)$ is bilinear continuous with norm 1*

Theorem 913 *(Schwartz I p.119) If E,F,G are normed vector spaces then the composition of maps : $\mathcal{L}(E; F) \times \mathcal{L}(F; G) \to \mathcal{L}(E; G) :: \circ(f,g) = g \circ f$ is bilinear, continuous and its norm is 1*

12.2.4 Family of semi-norms

A family of semi-metrics on a topological space can be useful because its topology can be Haussdorff (which ususally is not a semi-metric). Similarly on vector spaces :

Definition 914 *A **pseudo-normed space** is a vector space endowed with a family $(p_i)_{i \in I}$ of semi-norms such that for any finite subfamily $J : \exists k \in I : \forall j \in J : p_j \leq p_k$*

Theorem 915 *(Schwartz III p.435) A pseudo-normed space $(E,(p_i)_{i \in I})$ is a topological space with the base of open balls :*
$B(u) = \cap_{j \in J} B_j(u, r_j)$ with $B_j(u, r_j) = \{v \in E : p_j(u - v) < r_j\}$,for every finite subset J of I and familiy $(r_j)_{j \in J}$,$r_j > 0$

It works because all the balls $B_j(u, r_j)$ are convex subsets, and the open balls $B(u)$ are convex subsets.
The functions p_i must satisfy the usual conditions of semi-norms.

Theorem 916 *The topology defined above is Hausdorff iff $\forall u \neq 0 \in E, \exists i \in I : p_i(u) > 0$*

Theorem 917 *A countable family of seminorms on a vector space defines a semi-metric on E*

It is defined by :$d(x, y) = \sum_{n=0}^{\infty} \frac{1}{2^n} \frac{p_n(x-y)}{1+p_n(x-y)}$. If E is Hausdorff then this pseudo-metric is a metric.
However usualy a pseudo-normed space is not normable.

Theorem 918 *(Schwartz III p.436) A linear map between pseudo-normed spaces is continuous if it is continuous at 0. It is then uniformly continuous and Lipschitz.*

Theorem 919 *A topological vector space is locally convex iff its topology can be defined by a family of semi-norms.*

12.2.5 Weak topology

Weak topology is defined for general topological spaces. The idea is to use a collection of maps $\varphi_i : E \to F$ where F is a topological spaces to pull back a topology on E such that every φ_i is continuous.

This idea can be implemented for a topological vector space and its dual. It is commonly used when the vector space has already an initial topology, usually defined from a semi-norm. Then another topology can be defined, which is weaker than the initial topology and this is useful when the normed topology imposes too strict conditions. This is easily done by using families of semi-norms as above. For finite dimensional vector space the weak and the "strong" (usual) topologies are equivalent.

Weak-topology

Definition 920 *The **weak-topology** on a topological vector space E is the topology defined by the family of semi-norms on E:* $(p_\varpi)_{\varpi \in E'} : \forall u \in E : p_\varpi(u) = |\varpi(u)|$

It sums up to take as collection of maps the continuous (as defined by the initial topology on E) linear forms on E.

Theorem 921 *The weak topology is Hausdorff*

Proof. It is Hausdorff if E' is separating : if $\forall u \neq v \in E, \exists \varpi \in E' : \varpi(u) \neq \varpi(v)$ and this is a consequence of the Hahn-Banach theorem ∎

Theorem 922 *A sequence $(u_n)_{n \in \mathbb{N}}$ in a topological space E **converges weakly** to u if :* $\forall \varpi \in E' : \varpi(u_n) \to \varpi(u)$.

convergence (with the initial topology in E) $\Rightarrow$ weak convergence (with the weak topology in E)

So the criterium for convergence is weaker, and this is one of the main reasons for using this topology.

Theorem 923 *If E is a semi-normed vector space, then the weak-topology on E is equivalent to the topology of the semi-norm :*
$\|u\|_W = \sup_{\|\varpi\|_{E'}=1} |\varpi(u)|$

The weak norm $\|u\|_W$ and the initial norm $\|u\|$ are not equivalent if E is infinite dimensional (Wilansky p.270).

Theorem 924 *(Banach-Alaoglu): if E is a normed vector space, then the closed unit ball E is compact with respect to the weak topology iff E is reflexive.*

This the application of the same theorem for the *weak topology to the bidual.

*weak-topology

The ***weak-topology** on the topological dual E' of a topological vector space E is the topology defined by the family of semi-norms on E': $(p_u)_{u \in E} : \forall \varpi \in E' : p_u(\varpi) = |\varpi(u)|$

It sums up to take as collection of maps the evaluation maps given by vectors of E.

Theorem 925 *The *weak topology is Hausdorff*

Theorem 926 *(Wilansky p.274) With the *weak-topology E' is $\sigma-$compact, normal*

Theorem 927 *(Thill p.252) A sequence $(\varpi_n)_{n \in \mathbb{N}}$ in a the topological dual E' of a topological space E **converges weakly** to u if : $\forall u \in E : \varpi_n(u) \to \varpi(u)$.*

convergence (with the initial topology in E') $\Rightarrow$ weak convergence (with the weak topology in E')

So this is the topology of pointwise convergence (Thill p.252)

Theorem 928 *If E is a semi-normed vector space, then the weak-topology on E' is equivalent to the topology of the semi-norm :*
$\|\varpi\|_W = \sup_{\|u\|_E = 1} |\varpi(u)|$

The weak norm $\|\varpi\|_W$ and the initial norm $\|\varpi\|_{E'}$ are not equivalent if E is infinite dimensional.

Theorem 929 *Banach-Alaoglu (Wilansky p.271): If E is a semi-normed vector space, then the closed unit ball in its topological dual E' is a compact Hausdorff subset with respect to the *-weak topology.*

Remark : in both cases one can complicate the definitions by taking only a subset of E' (or E), or extend E' to the algebraic dual E*. See Bratelli (1 p.162) and Thill.

12.2.6   Fréchet space

Fréchet spaces have a somewhat complicated definition. However they are very useful, as they share many (but not all) properties of the Banach spaces which are the work-horses of analysis on vector spaces.

Definition 930 *A **Fréchet space** is a Hausdorff, complete, topological vector space, endowed with a countable family $(p_n)_{n \in \mathbb{N}}$ of semi-norms. So it is locally convex and metric.*

The metric is : $d\left(x,y\right)=\sum_{n=0}^{\infty}\frac{1}{2^{n}}\frac{p_{n}\left(x-y\right)}{1+p_{n}\left(x-y\right)}$

And because it is Hausdorff : $\forall u\neq 0\in E,\exists n\in\mathbb{N}:p_{n}\left(u\right)>0$

Theorem 931 *A closed vector subspace of a Fréchet space is a Fréchet space.*

Theorem 932 *(Taylor 1 p.482) A quotient of a Fréchet space by a closed subspace is a Fréchet space.*

Theorem 933 *The direct sum of a finite number of Fréchet spaces is a Fréchet space.*

Theorem 934 *(Taylor 1 p.481) A sequence $(u_{n})_{n\in\mathbb{N}}$ converges in a Fréchet space $\left(E,(p_{n})_{n\in\mathbb{N}}\right)$ iff $\forall m\in\mathbb{N}:p_{m}\left(u_{n}-u\right)\to_{n\to\infty}0$*

Linear functions on Fréchet spaces

Theorem 935 *(Taylor 1 p.491) For every linear map $f\in L\left(E;F\right)$ between Fréchet vector spaces :*

i) (open mapping theorem) If f is continuous and surjective then any neighborhood of 0 in E is mapped onto a neighborhood of 0 in F (f is open)

ii) If f is continuous and bijective then $f^{-1}$is continuous

iii) (closed graph theorem) if the graph of $f=\left\{\left(u,f\left(u\right)\right),u\in E\right\}$ is closed in ExF then f is continuous.

Theorem 936 *(Taylor I p.297) For any bilinear map : $B:E\times F\to\mathbb{C}$ on two complex Fréchet spaces $\left(E,(p_{n})_{n\in\mathbb{N}}\right),\left(F,(q_{n})_{n\in\mathbb{N}}\right)$ which is separately continuous on each variable, there are $C\in\mathbb{R},(k,l)\in\mathbb{N}^{2}:\forall\left(u,v\right)\in E\times F:\left|B\left(u,v\right)\right|\leq Cp_{k}\left(u\right)q_{l}\left(v\right)$*

Theorem 937 *(Zuily p.59) If a sequence $(f_{m})_{m\in\mathbb{N}}$ of continuous maps between two Fréchet spaces $(E,p_{n}),(F,q_{n})$ is such that : $\forall u\in E,\exists v\in F:f_{m}\left(u\right)_{m\to\infty}\to v$, then there is a map : $f\in\mathcal{L}\left(E;F\right)$ such that $f_{m}\left(u\right)_{m\to\infty}\to f(u)$ and for any compact K in E, any $n\in\mathbb{N}:\lim_{m\to\infty}\sup_{u\in K}q_{n}\left(f_{m}\left(u\right)-f\left(u\right)\right)=0$. If $(u_{m})_{m\in\mathbb{N}}$ is a sequence in E which converges to u then $(f_{m}\left(u_{m}\right))_{m\in\mathbb{N}}$ converges to f(u).*

This theorem is important, because it gives a simple rule for the convergence of sequence of linear maps. It holds in Banach spaces (which are Fréchet spaces).

The space $\mathcal{L}$(E;F) of continuous linear maps between Fréchet spaces E, F is usually not a Fréchet space. The topological dual of a Fréchet space is not necessarily a Fréchet space. However we have the following theorems.

Theorem 938 *Let $(E_{1},\Omega_{1}),(E_{2},\Omega_{2})$ two Fréchet spaces with their open subsets, If E_{2} is dense in E_{1}, then $E_{1}'\subset E_{2}'$*

Proof. $E_1^* \subset E_2^*$ because by restriction any linear map on E_1 is linear on E_2

take $\lambda \in E_1'$, $a \in E_2$ so $a \in E_1$

λ continuous on E_1 at a $\Rightarrow \forall \varepsilon > 0 : \exists \varpi_1 \in \Omega_1 : \forall u \in \varpi_1 : |\lambda(u) - \lambda(a)| \leq \varepsilon$

take any u in $\varpi_1, u \in \overline{E_2}, E_2$ second countable, thus first countable $\Rightarrow$ $\exists (v_n), v_n \in E_2 : v_n \to u$

So any neighborhood of u contains at least two points w, w' in E_2

So there are w$\neq$w'$\in \varpi_1 \cap E_2$

E_2 is Hausdorff $\Rightarrow \exists \varpi_2, \varpi_2' \in \Omega_2 : w \in \varpi_2, w' \in \varpi_2', \varpi_2 \cap \varpi_2' = \varnothing$

So there is $\varpi_2 \in \Omega_2 : \varpi_2 \subset \varpi_1$

and λ is continuous at a for E_2 ■

12.2.7 Affine spaces

All the previous material extends to affine spaces.

Definition 939 *An affine space $\left(E, \overrightarrow{E}\right)$ is semi-normed if its underlying vector space $\overrightarrow{E}$ is normed. The semi-norm defines uniquely a semi-metric : $d(A, B) = \left\|\overrightarrow{AB}\right\|$*

Theorem 940 *The closure and the interior of a convex subset of a semi-normed affine space are convex.*

Theorem 941 *Every ball B(A,r) of a semi-normed affine space is convex.*

Theorem 942 *A map $f : E \to F$ valued in,an affine normed space F is **bounded** if for a point $O \in F : \sup_{x \in E} \|f(x) - O\|_{\overrightarrow{F}} < \infty$. This property does not depend on the choice of O.*

Theorem 943 *(Schwartz I p.173) A hyperplane of a normed affine space E is either closed or dense in E. It is closed if it is defined by a continuous affine map.*

12.3 Banach spaces

For many applications a complete topological space is required, thanks to the fixed point theorem. So for vector spaces there are Fréchet spaces and Banach spaces. The latter is the structure of choice, whenever it is available, because it is easy to use and brings several useful tools such as series, analytic functions and one parameter group of linear maps. Moreover all classic calculations on series, usually done with scalars, can readily be adapted to Banach vector spaces.

Banach spaces are named after the Polish mathematician Stefan Banach who introduced them in 1920–1922 along with Hans Hahn and Eduard Helly

12.3.1 Banach Spaces

Definitions

Definition 944 *A **Banach vector space** is a complete normed vector space over a topologically complete field K*

usually K= $\mathbb{R}$ or $\mathbb{C}$

Definition 945 *A **Banach affine space** is a complete normed affine space over a topologically complete field K*

Usually a "Banach space" is a Banach vector space.

Any finite dimensional vector space is complete. So it is a Banach space when it is endowed with any norm.

A normed vector space can be completed. If the completion procedure is applied to a normed vector space, the result is a Banach space containing the original space as a dense subspace, and if it is applied to an inner product space, the result is a Hilbert space containing the original space as a dense subspace. So for all practical purposes the completed space can replace the initial one.

Subspaces

The basic applications of general theorems gives:

Theorem 946 *A closed vector subspace of a Banach vector space is a Banach vector space*

Theorem 947 *Any finite dimensional vector subspace of a Banach vector space is a Banach vector space*

Theorem 948 *If F is a closed vector subspace of the Banach space E then E/F is still a Banach vector space*

It can be given (Taylor I p.473) the norm :$\|u\|_{E/F} = \lim_{v \in F, v \to 0} \|u - v\|_E$

Series on a Banach vector space

Series must be defined on sets endowed with an addition, so many important results are on Banach spaces. Of course they hold for series defined on $\mathbb{R}$ or $\mathbb{C}$. First we define three criteria for convergence.

1. Absolutely convergence

Definition 949 *A series $\sum_{n \in \mathbb{N}} u_n$ on a semi-normed vector space E is **absolutely convergent** if the series $\sum_{n \in \mathbb{N}} \|u_n\|$ converges.*

Theorem 950 *(Schwartz I p.123) If the series* $\sum_{n\in\mathbb{N}} u_n$ *on a Banach E is* **absolutely convergent** *then :*

i) $\sum_{n\in\mathbb{N}} u_n$ *converges in E*

ii) $\left\|\sum_{n\in\mathbb{N}} u_n\right\| \leq \sum_{n\in\mathbb{N}} \|u_n\|$

iii) If $\varphi : \mathbb{N} \to \mathbb{N}$ *is any bijection, the series* $\sum_{n\in\mathbb{N}} u_{\varphi(n)}$ *is also absolutely convergent and* $\lim \sum_{n\in\mathbb{N}} u_{\varphi(n)} = \lim \sum_{n\in\mathbb{N}} u_n$

2. Commutative convergence:

Definition 951 *A series* $\sum_{i\in I} u_i$ *on a topological vector space E, where I is a countable set, is* **commutatively convergent** *if there is* $u \in E$ *such that :for every bijective map* φ *on I :* $\lim \sum_n u_{\varphi(j_n)} = u$

Then on a Banach : absolute convergence $\Rightarrow$ commutative convergence
Conversely :

Theorem 952 *(Neeb p.21) A series on a Banach commutatively convergent is absolutely convergent.*

Commutative convergence enables to define quantities such as $\sum_{i\in I} u_i$ for any set.

3. Summable family:

Definition 953 *(Neeb p.21) A family* $(u_i)_{i\in I}$ *of vectors on a semi-normed vector space E is said to be* **summable** *with sum u if :*

$\forall \varepsilon > 0, \exists J \subset I, card(J) < \infty : \forall K \subset J : \left\|\left(\sum_{i\in K} u_i\right) - x\right\| < \varepsilon$ *then one writes :* $u = \sum_{i\in I} u_i$.

Theorem 954 *(Neeb p.25) If a family* $(u_i)_{i\in I}$ *of vectors in the Banach E is summable, then only countably many* u_i *are non zero*

So for a countable set I, E Banach
summability $\Leftrightarrow$ commutative convergence$\Leftrightarrow$absolute convergence $\Rightarrow$ convergence in the usual meaning, but the converse is not true.

4. Image of a series by a continuous linear map:

Theorem 955 *(Schwartz I p.128) For every continuous map* $L \in \mathcal{L}(E; F)$ *between normed vector spaces : if the series* $\sum_{n\in\mathbb{N}} u_n$ *on E is convergent then the series0* $\sum_{n\in\mathbb{N}} L(u_n)$ *on F is convergent and* $\sum_{n\in\mathbb{N}} L(u_n) = L\left(\sum_{n\in\mathbb{N}} u_n\right)$.

If E,F are Banach, then the theorem holds for absolutely convergent (resp. commutatively convergent) and :

$\sum_{n\in\mathbb{N}} \|L(u_n)\| \leq \|L\| \sum_{n\in\mathbb{N}} \|u_n\|$

5. Image of 2 series by a continuous bilinear map:

Theorem 956 *(Schwartz I p.129) For every continuous bilinear map $B \in \mathcal{L}^2(E,F;G)$ between the Banach spaces E,F,G, if the series $\sum_{i \in I} u_i$ on E, $\sum_{j \in J} v_j$ on F, for I,J countable sets, are both absolutely convergent, then the series on $G : \sum_{(i,j) \in I} B(u_i, v_j)$ is absolutely convergent and $\sum_{(i,j) \in I} B(u_i, v_j) = B\left(\sum_{i \in I} u_i, \sum_{j \in J} v_j\right)$*

Theorem 957 *Abel criterium (Schwartz I p.134) For every continuous bilinear map $B \in \mathcal{L}^2(E,F;G)$ between the Banach spaces E,F,G on the field K, if :*

the sequence $(u_n)_{n \in \mathbb{N}}$ on E converges to 0 and is such that the series $\sum_{p=0}^{\infty} \|u_{p+1} - u_p\|$ converges,

the sequence $(v_n)_{n \in \mathbb{N}}$ on F is such that $\exists k \in K : \forall m, n : \left\|\sum_{p=m}^{n} v_p\right\| \leq k$,

Theorem 958 *then the series: $\sum_{n \in \mathbb{N}} B(u_n, v_n)$ converges to S, and*

$$\|S\| \leq \|B\| \left(\sum_{p=0}^{\infty} \|u_{p+1} - u_p\|\right) \left(\left\|\sum_{p=0}^{\infty} v_p\right\|\right)$$

$$\left\|\sum_{p>n} B(u_p, v_p)\right\| \leq \|B\| \left(\sum_{p=n+1}^{\infty} \|u_{p+1} - u_p\|\right) \left(\sup_{p>n} \left\|\sum_{m=p}^{\infty} v_m\right\|\right)$$

The last theorem covers all the most common criteria for convergence of series.

12.3.2   Continuous linear maps

It is common to say "operator" for a "continuous linear map" on a Banach vector space.

Properties of continuous linear maps on Banach spaces

Theorem 959 *For every linear map $f \in L(E;F)$ between Banach vector spaces :*

i) open mapping theorem (Taylor 1 p.490): If f is continuous and surjective then any neighborhood of 0 in E is mapped onto a neighborhood of 0 in F (f is open)

ii) closed graph theorem (Taylor 1 p.491): if the graph of $f = \{(u, f(u)), u \in E\}$ is closed in ExF then f is continuous.

iii) (Wilansky p.276) if f is continuous and injective then it is a homeomorphism

iv) (Taylor 1 p.490) If f is continuous and bijective then f^{-1} is continuous

v) (Schwartz I p.131) If f,g are continuous, f invertible and $\|g\| < \|f^{-1}\|^{-1}$ then f+g is invertible and $\left\|(f+g)^{-1}\right\| \leq \frac{1}{\|f^{-1}\|^{-1} - \|g\|}$

Theorem 960 *(Rudin) For every linear map $f \in L(E;F)$ between Banach vector spaces and sequence $(u_n)_{n \in \mathbb{N}}$ in E:*

i) If f is continuous then for every sequence $(u_n)_{n \in \mathbb{N}}$ in $E : u_n \to u \Rightarrow f(u_n) \to f(u)$

ii) Conversely if for every sequence $(u_n)_{n \in \mathbb{N}}$ in E which converges to 0: $f(u_n) \to v$ then v=0 and f is continuous.

Theorem 961 *(Wilansky p.273) If $(\varpi_n)_{n \in \mathbb{N}}$ is a sequence in the topological dual E' of a Banach space such that $\forall u \in E$ the set $\{\varpi_n(u), n \in \mathbb{N}\}$ is bounded, then the set $\{\|\varpi_n\|, n \in \mathbb{N}\}$ is bounded*

Theorem 962 *(Schwartz I p.109) If $f \in \mathcal{L}(E_0; F)$ is a continuous linear map from a dense subspace E_0 of a normed vector space to a Banach vector space F, then there is a unique continuous map $\widetilde{f} : E \to F$ which extends f, $\widetilde{f} \in \mathcal{L}(E; F)$ and $\left\| \widetilde{f} \right\| = \|f\|$*

If F is a vector subspace, the annihiliator $F^\intercal$ of F is the set : $\{\varpi \in E' : \forall u \in F : \varpi(u) = 0\}$

Theorem 963 *Closed range theorem (Taylor 1 p.491): For every linear map $f \in L(E; F)$ between Banach vector spaces : $\ker f^t = f(E)^\intercal$. Moreover if f(E) is closed in F then $f^t(F')$ is closed in E' and $f^t(F') = (\ker f)^\intercal$*

Properties of the set of linear continuous maps

Theorem 964 *(Schwartz I p.115) The set of continuous linear maps $\mathcal{L}(E;F)$ between a normed vector space and a Banach vector space F on the same field is a Banach vector space*

Theorem 965 *(Schwartz I p.117) The set of continuous multilinear maps $\mathcal{L}^r(E_1, E_2, ..E_r; F)$ between normed vector spaces $(E_i)_{i=1}^r$ and a Banach vector space F on the same field is a Banach vector space*

Theorem 966 *if E,F are Banach : $\mathcal{L}(E;F)$ is Banach*

Theorem 967 *The topological dual E' of a Banach vector space is a Banach vector space*

A Banach vector space may be not reflexive : the bidual (E')' is not necessarily isomorphic to E.

Theorem 968 *(Schwartz II p.81) The sets of invertible continuous linear maps $G\mathcal{L}(E;F), G\mathcal{L}(F;E)$ between the Banach vector spaces E,F are open subsets in $\mathcal{L}(E;F), \mathcal{L}(F;E)$, thus they are normed vector spaces but not complete. The map $\Im : G\mathcal{L}(E;F) \to G\mathcal{L}(F;E) :: \Im(f) = f^{-1}$ is an homeomorphism (bijective, continuous as its inverse).*

Then we have : $\left\| f \circ f^{-1} \right\| = \|Id\| = 1 \leq \|f\| \left\| f^{-1} \right\| \leq \|\Im\|^2 \|f\| \left\| f^{-1} \right\| \Rightarrow \|\Im\| \geq 1, \left\| f^{-1} \right\| \geq 1/\|f\|$

Theorem 969 *The set $G\mathcal{L}(E;E)$ of invertible endomorphisms on a Banach vector space is a topological group with compose operation and the metric associated to the norm, open subset in $\mathcal{L}(E; E)$.*

Notice that an "invertible map f in $G\mathcal{L}$(E;F)" means that f^{-1} must also be a continuous map, and for this it is sufficient that f is continuous and bijective .

Theorem 970 *(Neeb p.141) If X is a compact topological space, endowed with a Radon measure μ, E,F are Banach vector spaces, then:*

i) for every continuous map : $f \in C_0\left(X; E\right)$ there is a unique vector U in E such that :

$\forall \lambda \in E' : \lambda\left(U\right) = \int_X \lambda\left(f\left(x\right)\right)\mu\left(x\right)$ *and we write :* $U = \int_X f\left(x\right)\mu\left(x\right)$

ii) for every continuous map : $L \in \mathcal{L}\left(E; F\right) : L\left(\int_X f\left(x\right)\mu\left(x\right)\right) = \int_X \left(L \circ f\left(x\right)\right)\mu\left(x\right)$

Spectrum of a map

A scalar λ if an eigen value for the endomorphism f$\in\mathcal{L}(E; E)$ if there is a vector u such that f(u)=λu, so $f - \lambda I$ cannot be inversible. On infinite dimensional topological vector space the definition is enlarged as follows.

Definition 971 *For every linear continuous endomorphism f on a topological vector space E on a field K,*

*i) the **spectrum** $Sp\left(f\right)$ of f is the subset of the scalars $\lambda \in K$ such that $(f - \lambda Id_E)$ has no inverse in $\mathcal{L}(E; E)$.*

*ii) the **resolvent set** $\rho\left(f\right)$ of f is the complement of the spectrum*

*iii) the map: $R : K \to L\left(E; E\right) :: R\left(\lambda\right) = \left(\lambda Id - f\right)^{-1}$ is called the **resolvent** of f.*

If λ is an eigen value of f, it belongs to the spectrum, but the converse is no true. If $f \in G\mathcal{L}\left(E; E\right)$ then $0 \notin Sp(f)$.

This definition can be extended to any algebra, and more properties are seen in the next section (Normed algebras). However t the spectrum has some specificities on Banach vector spaces.

Theorem 972 *The spectrum of a continuous endomorphism f on a complex Banach vector space E is a non empty compact subset of $\mathbb{C}$ bounded by $\|f\|$*

Proof. It is a consequence of general theorems on Banach algebras : $\mathcal{L}$(E;E) is a Banach algebra, so the spectrum is a non empty compact, and is bounded by the spectral radius, which is $\leq \|f\|$ $\blacksquare$

Theorem 973 *(Schwartz 2 p.69) The set of eigen values of a compact endomorphism on a Banach space is either finite, or countable in a sequence convergent to 0 (which is or not an eigen value).*

Theorem 974 *(Taylor 1 p.493) If f is a continuous endomorphism on a complex Banach space:*

$|\lambda| > \|f\| \Rightarrow \lambda \in \rho\left(f\right)$. In particular if $\|f\| < 1$ then Id-f is invertible and $\sum_{n=0}^{\infty} f^n = \left(Id - f\right)^{-1}$

If $\lambda_0 \in \rho\left(f\right)$ then : $R\left(\lambda\right) = R\left(\lambda_0\right)\sum_{n=0}^{\infty} R\left(\lambda_0\right)^n \left(\lambda - \lambda_0\right)^n$

If $\lambda_1, \lambda_2 \in \rho\left(f\right)$ then : $R\left(\lambda_1\right) - R\left(\lambda_2\right) = \left(\lambda_1 - \lambda_2\right)R\left(\lambda_1\right) \circ R\left(\lambda_2\right)$

Compact maps

Theorem 975 *(Schwartz 2 p.60) If f is a continuous compact map* $f \in \mathcal{L}(E;F)$ *between a reflexive Banach vector space E and a topological vector space f, then the closure in F* $\overline{f(B(0,1))}$ *of the image by f of the unit ball B(0,1) in E is compact in F.*

Theorem 976 *(Taylor 1 p.496) The transpose of a compact map is compact.*

Theorem 977 *(Schwartz 2 p.63) If* $(f_n)_{n \in \mathbb{N}}$ *is a sequence of linear continuous maps of finite rank between Banach vector spaces, which converges to f, then f is a compact map.*

Theorem 978 *(Taylor 1 p.495)The set of compact maps between Banach vector spaces is a closed vector subspace of the space of continuous linear maps* $\mathcal{L}(E;F)$.

Theorem 979 *(Taylor 1 p.499) The spectrum Sp(f) of a compact endomorphism* $f \in \mathcal{L}(E;E)$ *on a complex Banach space has only 0 as point of accumulation, and all* $\lambda \neq 0 \in Sp(f)$ *are eigen values of f.*

Fredholm operators

Fredholm operators are "proxy" for isomorphisms. Their main feature is the index.

Definition 980 *(Taylor p.508) A continuous linear map* $f \in \mathcal{L}(E;F)$ *between Banach vector spaces E,F is said to be a* **Fredholm operator** *if* $\ker f$ *and* $F/f(E)$ *are finite dimensional. Equivalentely if there exists* $g \in \mathcal{L}(F;E)$ *such that :* $Id_E - g \circ f$ *and* $Id_F - f \circ g$ *are continuous and compact. The* **index** *of f is :* $Index(f) = \dim \ker f - \dim F/f(E) = \dim \ker f - \dim \ker f^t$

Theorem 981 *(Taylor p.508) The set Fred(E;F) of Fredholm operators is an open vector subspace of* $\mathcal{L}(E;F)$. *The map : Index:Fred(E;F)* $\to \mathbb{Z}$ *is constant on each connected component of Fred(E;F).*

Theorem 982 *(Taylor p.508) The compose of two Fredholm operators is Fredholm :* *If* $f \in Fred(E;F), g \in Fred(F;G)$ *then* $g \circ f \in Fred(E;G)$ *and Index(gf)=Index(f) + Index(g). If f is Fredholm and g compact then f+g is Fredholm and Index(f+g)=Index(f)*

Theorem 983 *(Taylor p.508) The transpose* f^t *of a Fredholm operator f is Fredholm and* $Index(f^t) = -Index(f)$

12.3.3 Analytic maps on Banach spaces

With the vector space structure of $\mathcal{L}(E;E)$ one can define any linear combination of maps. But in a Banach space one can go further and define "functions" of an endomorphism.

Exponential of a linear map

Theorem 984 *The **exponential** of a continuous linear endomorphism $f \in \mathcal{L}(E;E)$ on a Banach space E is the continuous linear map : $\exp f = \sum_{n=0}^{\infty} \frac{1}{n!} f^n$ where f^n is the n iterated of f*

Proof. $\forall u \in F$, the series $\sum_{n=0}^{\infty} \frac{1}{n!} f^n(u)$ converges absolutely :

$\sum_{n=0}^{N} \frac{1}{n!} \|f^n(u)\| \leq \sum_{n=0}^{N} \frac{1}{n!} \|f^n\| \|u\| = \sum_{n=0}^{N} \frac{1}{n!} \|f\|^n \|u\| \leq (\exp \|f\|) \|u\|$

we have an inceasing bounded sequence on $\mathbb{R}$ which converges.

and $\left\| \sum_{n=0}^{\infty} \frac{1}{n!} f^n(u) \right\| \leq (\exp \|f\|) \|u\|$ so exp is continuous with $\|\exp f\| \leq \exp \|f\|$ ∎

A simple computation as above brings (Neeb p.170):

If $f \circ g = g \circ f \Rightarrow \exp(f+g) = (\exp f) \circ (\exp g)$

$\exp(-f) = (\exp f)^{-1}$

$\exp(f^t) = (\exp f)^t$

$g \in GL(E;E) : \exp \left(g^{-1} \circ f \circ g \right) = g^{-1} \circ (\exp f) \circ g$

If E,F are finite dimensional : $\det(\exp f) = \exp(Trace(f))$

If E is finite dimensional the inverse log of exp is defined as :

$(\log f)(u) = \int_{-\infty}^{0} [(s-f)^{-1} - (s-1)^{-1}](u)\, ds$ if f has no eigen value ≤ 0

Then : $\log(g \circ f \circ g^{-1}) = g \circ (\log f) \circ g^{-1}$

$\log(f^{-1}) = -\log f$

Holomorphic groups

The exponential can be generalized.

Theorem 985 *If f is a continuous linear endomorphism $f \in \mathcal{L}(E;E)$ on a complex Banach space E then the map : $\exp zf = \sum_{n=0}^{\infty} \frac{z^n}{n!} f^n \in \mathcal{L}(E;E)$ and defines the holomorphic group : $U : \mathbb{C} \to \mathcal{L}(E;E) :: U(z) = \exp zf$ with $U(z_2) \circ U(z_1) = U(z_1 + z_2), U(0) = Id$*

U is holomorphic on $\mathbb{C}$ and $\frac{d}{dz}(\exp zf)|_{z=t_0} = f \circ \exp z_0 f$

Proof. i) The previous demonstration can be generalized in a complex Banach space for $\sum_{n=0}^{\infty} \frac{z^n}{n!} f^n$

Then, for any continuous endomorphism f we have a map : $\exp zf = \sum_{n=0}^{\infty} \frac{z^n}{n!} f^n \in \mathcal{L}(E;E)$

ii) $\exp zf(u) = \exp f(zu)$, $z_1 f, z_2 f$ commutes so : $\exp(z_1 f) \circ \exp(z_2 f) = \exp(z_1 + z_2)f$

iii) $\frac{1}{z}(U(z) - I) - f = \sum_{n=1}^{\infty} \frac{z^{n-1}}{n!} f^n - f = f \circ \left(\sum_{n=1}^{\infty} \frac{z^{n-1}}{n!} f^{n-1} - Id \right) = f \circ (\exp zf - Id)$

$\left\| \frac{1}{z}(U(z) - I) - f \right\| \leq \|f\| \|(\exp zf - Id)\|$

$\lim_{z \to 0} \left\| \frac{1}{z}(U(z) - I) - f \right\| \leq \lim_{z \to 0} \|f\| \|(\exp zf - Id)\| = 0$

Thus U is holomorphic at z=0 with $\frac{dU}{dz}|_{z=0} = f$

iv) $\frac{1}{h}(U(z+h) - U(z)) - f \circ U(z) = \frac{1}{h}(U(h) - I) \circ U(z) - f \circ U(z) = \left(\frac{1}{h}(U(h) - I) - f \right) \circ U(z)$

$\left\| \frac{1}{h}(U(z+h) - U(z)) - f \circ U(z) \right\| \leq \left\| \left(\frac{1}{h}(U(h) - I) - f \right) \right\| \|U(z)\|$

$\lim_{h\to 0}\left\|\frac{1}{h}\left(U\left(z+h\right)-U\left(z\right)\right)-f\circ U\left(z\right)\right\|\leq\lim_{h\to 0}\left\|\left(\frac{1}{h}\left(U\left(h\right)-I\right)-f\right)\right\|\left\|U\left(z\right)\right\|=$ 0

So U is holomorphic on $\mathbb{C}$ and $\frac{d}{dz}\left(\exp tf\right)|_{z=t_0}=f\circ\exp z_0 f$ ∎

For every endomorphism $f\in\mathcal{L}\left(E;E\right)$ on a complex or real Banach space E then the map : $\exp tf:\mathbb{R}\to\mathcal{L}\left(E;E\right)$ defines a one parameter group and U(t)=exptf is smooth and $\frac{d}{dt}\left(\exp tf\right)|_{t=t_0}=f\circ\exp t_0 f$

Map defined through a holomorphic map

The previous procedure can be generalized. This is an appication of the spectral theory (see the dedicated section).

Theorem 986 *(Taylor 1 p.492) Let $\varphi:\Omega\to\mathbb{C}$ be a holomorphic map on a bounded region, with smooth border, of $\mathbb{C}$*

and $f\in\mathcal{L}\left(E;E\right)$ a continuous endomorphism on the complex Banach E.

i) If Ω contains the spectrum of f, the following map is a continuous endomorphism on E:

$\Phi\left(\varphi\right)\left(f\right)=\frac{1}{2i\pi}\int_{\partial\Omega}\varphi\left(\lambda\right)\left(\lambda I-f\right)^{-1}d\lambda\in\mathcal{L}\left(E;E\right)$

ii) If $\varphi\left(\lambda\right)=1$ then $\Phi\left(\varphi\right)\left(f\right)=Id$

iii) If $\varphi\left(\lambda\right)=\lambda$ then $\Phi\left(\varphi\right)\left(f\right)=f$

iv) If φ_1,φ_2 are both holomorphic on Ω, then : $\Phi\left(\varphi_1\right)\left(f\right)\circ\Phi\left(\varphi_2\right)\left(f\right)=\Phi\left(\varphi_1\times\varphi_2\right)\left(f\right)$

12.3.4 One parameter group

The main purpose is the study of the differential equation $\frac{dU}{dt}=SU\left(t\right)$ where $U\left(t\right),S\in\mathcal{L}\left(E;E\right).$ S is the infinitesimal generator of U. If S is continuous then the general solution is $U\left(t\right)=\exp tS$ but as it is not often the case we have to distinguish norm and weak topologies. On this topic we follow Bratelli (I p.161). See also Spectral theory on the same topic for unitary groups on Hilbert spaces.

Definition

Definition 987 *A **one parameter group** of operators on a Banach vector space E is a map : $U:\mathbb{R}\to\mathcal{L}(E;E)$ such that :*

$U\left(0\right)=Id,U\left(s+t\right)=U\left(s\right)\circ U\left(t\right)$

the family U(t) has the structure of an abelian group, isomorphic to $\mathbb{R}$.

Definition 988 *A **one parameter semi-group** of operators on a Banach vector space E is a map : $U:\mathbb{R}_+\to\mathcal{L}(E;E)$ such that :*

$U\left(0\right)=Id,U\left(s+t\right)=U\left(s\right)\circ U\left(t\right)$

the family U(t) has the structure of a monoid (or semi-group)

So we denote T=$\mathbb{R}$ or $\mathbb{R}_+$

Notice that U(t) (the value at t) *must be continuous.* The continuity conditions below do not involve U(t) but the map $U:T\to\mathcal{L}\left(E;E\right).$

Norm topology

Definition 989 *(Bratelli p.161) A one parameter (semi) group U of continuous operators on E is said to be **uniformly continuous** if one of the equivalent conditions is met:*

i) $\lim_{t \to 0} \|U(t) - Id\| = 0$

ii) $\exists S \in \mathcal{L}(E; E) : \lim_{t \to 0} \left\| \frac{1}{t}(U(t) - I) - S \right\| = 0$

iii) $\exists S \in \mathcal{L}(E; E) : U(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} S^n = \exp tS$

*S is the **infinitesimal generator** of U and one writes $\frac{dU}{dt} = SU(t)$*

A uniformly continuous one parameter semi group U can be extended to T=$\mathbb{R}$ such that $\|U(t)\| \leq \exp(|t| \|S\|)$

If these conditions are met the problem is solved. And conversely a one parameter (semi) group of continuous operators is uniformly continuous iff its generator is continuous.

Weak topology

Definition 990 *(Bratelli p.164) A one parameter (semi) group U of continuous operators on the banach vector space E on the field K is said to be **weakly continuous** if $\forall \varpi \in E'$ the map $\phi_\varpi : T \times E \to K :: \phi_\varpi(t, u) = \varpi(U(t)u)$ is such that :*

$\forall t \in T : \phi_\varpi(t, .) : E \to K$ *is continuous*

$\forall u \in E : \phi_\varpi(., u) : T \to K$ *is continuous*

So one can say that U is continuous in the weak topology on E.

Similarly a one parameter group U on E' : $U : \mathbb{R} \to \mathcal{L}(E'; E')$ is continuous in the *weak topology if $\forall u \in E$ the map $\phi_u : T \times E' \to K :: \phi_u(t, \varpi) = U(t)(\varpi)(u)$ is such that :

$\forall t \in T : \phi_u(t, .) : E' \to K$ is continuous

$\forall \varpi \in E' : \phi_u(., \varpi) : T \to K$ is continuous

Theorem 991 *(Bratelli p.164-165) If a one parameter (semi) group U of operators on E is weakly continuous then :*

i) $\forall u \in E : \psi_u : T \to E : \psi_u(t) = U(t)u$ *is continuous in the norm of E*

ii) $\exists M \geq 1, \exists \beta \geq \inf_{t>0} \frac{1}{t} \ln \|U(t)\| : \|U(t)\| \leq M \exp \beta t$

iii) for any complex borelian measure μ on T such that $\int_T e^{\beta t} |\mu(t)| < \infty$ the map :

$U_\mu : E \to E :: U_\mu(u) = \int_T U(t)(u) \mu(t)$ *belongs to $\mathcal{L}(E; E)$*

The main difference with the uniformly continuous case is that the infinitesimal generator does not need to be defined over the whole of E.

Theorem 992 *(Bratelli p.165-166) A map $S \in L(D(S); E)$ with domain $D(S) \subset E$ is the infinitesimal generator of the weakly continuous one parameter (semi) group U on a Banach E if :*

$\forall u \in D(S), \exists v \in E : \forall \varpi \in E' : \varpi(v) = \lim_{t \to 0} \frac{1}{t} \varpi((U(t) - Id) u)$

then :

i) $\forall u \in E : T \to E :: U(t)u$ is continuous in the norm of E

ii) D(S) is dense in E in the weak topology

iii) $\forall u \in D(S) : S \circ U(t) u = U(t) \circ Su$

iii) if $\operatorname{Re}\lambda > \beta$ then the range of $(\lambda Id - S)^{-1} = E$ and $\forall u \in D(S) :$
$\|(\lambda Id - S) u\| \geq M^{-1} (\operatorname{Re}\lambda - \beta) \|u\|$

iv) the resolvent $(\lambda Id - S)^{-1}$ is given by the Laplace transform : $\forall \lambda : \operatorname{Re}\lambda > \beta, \forall u \in E : (\lambda Id - S)^{-1} u = \int_0^\infty e^{-\lambda t} U(t)u\, dt$

Notice that $\frac{d}{dt}U(t)u = Su$ only if u$\in D(S)$. The parameters β, M refer to the previous theorem.

The following theorem gives a characterization of the linear endomorphisms S defined on a subset D(S) of a Banach space which can be an infintesimal generator.

Theorem 993 *Hille-Yoshida (Bratelli p.171): Let $S \in L(D(S); E), D(S) \sqsubseteq E$, E Banach vector space, then the following conditions are equivalent :*

i) S is the infinitesimal generator of a weakly continuous semi group U in E and U(t) is a contraction

ii) D(S) is dense in E and S closed in D(S) (in the weak topology) and

$\forall u \in D(S), \forall \alpha \geq 0 : \|(Id - \alpha S) u\| \geq \|u\|$ and for some $\alpha > 0$: the range of $(Id - \alpha S)^{-1} = E$

If so then U is defined by : $\forall u \in D(S) : U(t)u = \lim_{\varepsilon \to 0} \exp\left(tS(Id - \varepsilon S)^{-1}\right) u = \lim_{n \to \infty} \left(I - \frac{1}{n}tS\right)^{-n} u$ where the exponential is defined by power series expansion. The limits exist in compacts in the weak topology uniformly for t, and if u is in the norm closure of D(S) the limits exist in norm.

12.4   Normed algebras

Algebras are vector spaces with an internal operation. Their main algebraic properties are seen in the Algebra part. To add a topology the most natural way is to add a norm and one has a normed algebra and, if it is complete, a Banach algebra. Several options are common : assumptions about the norm and the internal operation on one hand, the addition of an involution (copied from the adjoint of matrices) on the other hand, and both lead to distinguish several classes of normed algebras, notably C*-algebras.

Normed algebras are met frequently in mathematics : square matrices with the Schmidt norm, space of linear endomorphisms, spaces of functions. Their interest in physics come from quantum mechanics : a system is represented as a set of observables, which are linear operators on a Hilbert space, and states, which are functionals on the observables. So the axiomatisation of quantum mechanics has lead to give the center place to C*-algebras (see Bratelli for more on the subject).

In this section we review the fundamental properties of normed algebras, their representation is seen in the Spectral theory section. We use essentially the comprehensive study of M.Thill. We strive to address as many subjects as possible, while staying simple and practical. Much more can be found in M.Thill's study. Bratelli gives an in depth review of the dynamical aspects, more centered on the C*algebras and one parameter groups.

12.4.1 Algebraic structure

This is a reminder of definitions from the Algebra part.

1. Algebra:

An algebra A is a vector space on a field K (it will be $\mathbb{C}$, if K=$\mathbb{R}$ the adjustments are obvious) endowed with an internal operation (denoted as multiplication XY with inverse X^{-1}) which is associative, distributive over addition and compatible with scalar multiplication. *We assume that it is unital*, with unity element denoted I.

An algebra is commutative if XY=YX for every element.

2. Commutant:

The commutant, denoted S', of a subset S of an algebra A is the set of all elements in A which commute with all the elements of S for the internal operation. This is a subalgebra, containing I. The second commutant, denoted S", is the commutant of S'.

3. Projection and reflexion:

An element X of A is a projection if $XX = X$, a reflexion if $X = X^{-1}$, nilpotent if $X \cdot X = 0$

4. Star algebra:

Inspired from the "adjoint" operation on matrix. A *algebra is endowed with an involution such that : $(X + Y)^* = X^* + Y^*; (X \cdot Y)^* = Y^* \cdot X^*; (\lambda X)^* = \overline{\lambda} X^*; (X^*)^* = X$

Then the adjoint of an element X is X*

An element X of a *-algebra is : normal if XX*=X*X, self-adjoint (or hermitian) if X=X*, anti self-adjoint (or antihermitian) if X=-X*, unitary if XX*=X*X=I

The subset of self-adjoint elements in A is a real vector space, real form of the vector space A.

12.4.2 Topological structures

An algebra has the structure of a vector space, so we distinguish in the obvious way : topological algebra, normed algebra, Banach algebra. Further we can distinguish an algebra and a *algebra. For the sake of simplicity we will make use only of :

- normed algebra, normed *-algebra
- Banach algebra, Banach *-algebra, C*-algebra

Topological algebra

Definition 994 *A **topological algebra** is a topological vector space such that the internal operation is continuous.*

Normed algebras

Definition 995 *A **normed algebra** is a normed vector space endowed with the structure of a topological algebra with the topology induced by the norm $\|\|$, with the additional properties that : $\|XY\| \leq \|X\| \|Y\|, \|I\| = 1$.*

Notice that each element in A must have a finite norm.

There is always an equivalent norm such that $\|I\| = 1$

A normed algebra is a rich structure, so much so that if we go further we fall in known territories :

Theorem 996 *Gel'fand-Mazur (Thill p.40) A normed algebra which is also a division ring (each element has an inverse) is isomorphic to $\mathbb{C}$*

Definition 997 *A normed *-algebra is a normed algebra and a *algebra such that the involution is continuous. We will require also that : $\forall X \in A : \|X^*\| = \|X\|$ and $\|X\|^2 = \|X^*X\|$*

It implies $\|I\| = 1$

(so a normed *algebra is a pre C*-algebra in Thill's nomenclature)

Theorem 998 *(Thill p.120) In a normed *-algebra, if the involution * is continuous, then the map $X \to X^*X$ is continuous in 0*

Theorem 999 *(Thill p.120) If the sequence $(X_n)_{n \in \mathbb{N}}$ in a normed *-algebra converges to 0, then the sequence $(X_n^*)_{n \in \mathbb{N}}$ is bounded*

Banach algebra

Definition 1000 *A **Banach algebra** is a normed algebra which is complete with the norm topology.*

It is always possible to complete a normed algebra to make it a Banach algebra.

Theorem 1001 *(Thill p.12) A Banach algebra is isomorphic and homeomorphic to the space of continuous endomorphisms on a Banach space.*

Take A as vector space and the maps : $\rho : A \to \mathcal{L}(A; A) :: \rho(X)Y = XY$ this is the left regular representation of A on itself.

Definition 1002 *A **Banach *-algebra** is a Banach algebra which is endowed with a continuous involution such that $\|XY\| \leq \|X\| \|Y\|, \|I\| = 1$.*

Definition 1003 *A **C*-algebra** is a Banach *-algebra with a continuous involution * such that $\|X^*\| = \|X\|$ and $\|X\|^2 = \|X^*X\|$*

The results for series seen in Banach vector space still hold, but the internal product opens additional possibilities. The main theorem is the following:

Theorem 1004 *Mertens (Thill p.53): If the series in a Banach algebra, $\sum_{n\in\mathbb{N}} X_n$ is absolutely convergent, $\sum_{n\in\mathbb{N}} Y_n$ is convergent, then the series (called the Cauchy product) $\sum_{n\in\mathbb{N}} Z_n = \sum_{n\in\mathbb{N}} \left(\sum_{k=0}^{n} X_k Y_{n-k}\right)$ converges and $\sum_{n\in\mathbb{N}} Z_n = \left(\sum_{n\in\mathbb{N}} X_n\right)\left(\sum_{n\in\mathbb{N}} Y_n\right)$*

12.4.3 Examples

1. Banach vector space

Theorem 1005 *On a Banach vector space the set $\mathcal{L}(E; E)$ is a Banach algebra with composition of maps as internal product. If E is a Hilbert space $\mathcal{L}(E; E)$ is a C*-algebra*

2. Spaces of functions:
(see the Functional analysis part for more)
Are commutative C*-algebra with pointwise multiplication and the norm : $\|f\| = \max |f|$
   i) The set $C_b(E; \mathbb{C})$ of bounded functions
   ii) if E Hausdorff, the set $C_{0b}(E; \mathbb{C})$ of bounded continuous functions
   iii) if E Hausdorff, locally compact, the set $C_{0v}(E; \mathbb{C})$ of continuous functions vanishing at infinity.
   If is E Hausdorff, locally compact, the set $C_{0c}(E; \mathbb{C})$ of continuous functions with compact support with the norm : $\|f\| = \max |f|$ is a normed *-algebra which is dense in $C_{0\nu}(E; \mathbb{C})$

3. Matrices:

Theorem 1006 *The set $\mathbb{C}(r)$ of square complex matrices is a finite dimensional C*-algebra with the norm $\|M\| = \frac{1}{r} Tr(MM^*)$*

12.4.4 Morphisms

The morphisms are maps between sets endowed with the same structures and preserving this structure.

Definition 1007 *An **algebra morphism** between the topological algebras A,B is a continuous linear map $f \in \mathcal{L}(A; B)$ such that $f(XY) = f(X) \cdot f(Y), f(I_A) = I_B$*

Definition 1008 *A ***-algebra morphism** between the topological *-algebras A,B is an algebra morphism f such that $f(X)^* = f(X^*)$*

As usual a morphism which is bijective and whose inverse map is also a morphism is called an isomorphism.

When the algebras are normed, a map which preserves the norm is an isometry. It is necessarily continuous.

A *-algebra isomorphism between C*-algebras is necessarily an isometry, and will be called a C*-algebra isomorphism.

Theorem 1009 *(Thill p.119,120) A map $f : A \to E$ from a normed *-algebra A to a normed vector space F is $\sigma-$contractive if $\|f(X)\| \leq r_\sigma(X)$. Then it is continuous.*

Theorem 1010 *(Thill p.48) A map $f \in L(A; B)$ between a Banach *-algebra A, and a normed *-algebra B, such that : $f(XY) = f(X) \cdot f(Y), f(I) = I$ and $f(X)^* = f(X^*)$ is continuous, and a *-algebra morphism*

Theorem 1011 *(Thill p.46) A *morphism f from a C*-algebra A to a normed *-algebra B :*
   i) is contractive ($\|f\| \leq 1$)
   ii) $f(A)$ is a C-algebra*
   iii) $A/\ker f$ is a C-algebra*
   iv) if f is injective, it is an isometry
   v) f factors in a C-algebra isomorphism $A/\ker f \to f(A)$*

12.4.5   Spectrum

The spectrum of an element of an algebra is an extension of the eigen values of an endomorphism. This is the key tool in the study of normed algebras.

Invertible elements
   "Invertible" will always mean "invertible for the internal operation".

Theorem 1012 *The set G(A) of invertible elements of a topological algebra is a topological group*

Theorem 1013 *(Thill p.38, 49) In a Banach algebra A, the set G(A) of invertible elements is an open subset and the map $X \to X^{-1}$ is continuous.*

If the sequence $(X_n)_{n \in \mathbb{N}}$ in G(A) converges to X, then the sequence $\left(X_n^{-1}\right)_{n \in \mathbb{N}}$ converges to X^{-1} iff it is bounded.

The border $\partial G(A)$ is the set of elements X such that there are sequences $(Y_n)_{n \in \mathbb{N}}, (Z_n)_{n \in \mathbb{N}}$ in A such that :
   $\|Y_n\| = 1, \|Z_n\| = 1, XY_n \to 0, Z_n X \to 0$

Spectrum

Definition 1014 *For every element X of an algebra A on a field K:*

*i) the **spectrum** $Sp(X)$ of X is the subset of the scalars $\lambda \in K$ such that $(f - \lambda Id_E)$ has no inverse in A.*

*ii) the **resolvent set** $\rho(f)$ of X is the complement of the spectrum*

*iii) the map: $R : K \to A :: R(\lambda) = (\lambda Id - X)^{-1}$ is called the **resolvent** of X.*

As we have assumed that K=$\mathbb{C}$ the spectrum is in $\mathbb{C}$.

Warning ! the spectrum is *relative to an algebra* A, and the inverse must be *in the algebra* :

i) is A is a normed algebra then we must have $\left\| (X - \lambda I)^{-1} \right\| < \infty$

ii) When one considers the spectrum in a subalgebra and when necessary we will denote $\mathrm{Sp}_A(X)$.

Spectral radius

The interest of the spectral radius is that, in a Banach algebra : $\max(|\lambda|, \lambda \in Sp(X)) = r_\lambda(X)$ (Spectral radius formula)

Definition 1015 *The **spectral radius** of an element X of a normed algebra is the real scalar:*
$$r_\lambda(X) = \inf \|X^n\|^{1/n} = \lim_{n \to \infty} \|X^n\|^{1/n}$$

Theorem 1016 *(Thill p.35, 40, 41)*
$$r_\lambda(X) \leq \|X\|$$
$$k \geq 1 : r_\lambda(X^k) = (r_\lambda(X))^k$$
$$r_\lambda(XY) = r_\lambda(YX)$$
If XY=YX :
$$r_\lambda(XY) \leq r_\lambda(X) r_\lambda(Y) ; r_\lambda(X + Y) \leq r_\lambda(X) + r_\lambda(Y) ; r_\lambda(X - Y) \leq |r_\lambda(X) - r_\lambda(Y)|$$

Theorem 1017 *(Thill p.36) For every element X of a Banach algebra the series $f(z) = \sum_{n=0}^\infty z^n X^n$ converges absolutely for $|z| < 1/r_\lambda(X)$ and it converges nowhere for $|z| > 1/r_\lambda(X)$. The radius of convergence is $1/r_\lambda(X)$*

Theorem 1018 *(Thill p.60) For $\mu > r_\lambda(X)$, the **Cayley transform** of X : $C_\mu(X) = (X - \mu i I)(X + \mu i I)^{-1}$ of every self adjoint element of a Banach *-algebra is unitary*

Structure of the spectrum

Theorem 1019 *(Thill p.40) In a normed algebra the spectrum is never empty.*

Theorem 1020 *(Thill p.39, 98) In a Banach algebra :*

- the spectrum is a non empty compact in $\mathbb{C}$, bounded by $r_\lambda(X) \leq \|X\|$: $\max(|\lambda|, \lambda \in Sp(X)) = r_\lambda(X)$

- the spectrum of a reflexion X is Sp(X)=(-1,+1)

Theorem 1021 *(Thill p.34) In a *-algebra :*
$Sp(X^*) = \overline{Sp(X)}$
for every normal element $X : r_\lambda(X) = \|X\|$

Theorem 1022 *(Thill p.41) In a Banach *-algebra: $r_\lambda(X) = r_\lambda(X^*), Sp(X^*) = \overline{Sp(X)}$*

Theorem 1023 *(Thill p.60) In a C*-algebra the spectrum of an unitary element is contained in the unit circle*

Theorem 1024 *(Thill p.33) For every element : $(Sp(XY)) \backslash 0 = (Sp(YX)) \backslash 0$*

Theorem 1025 *(Thill p.73) In a Banach algebra if XY=YX then : $Sp(XY) \subset Sp(X) Sp(Y) ; Sp(X+Y) \subset \{Sp(X) + Sp(Y)\}$*

Theorem 1026 *(Thill p.32, 50, 51) For every normed algebra, B subalgebra of A, $X \in B$*
$Sp_A(X) \sqsubseteq Sp_B(X)$
(Silov) if B is complete or has no interior : $\partial Sp_B(X) \subset \partial Sp_A(X)$

Theorem 1027 *(Thill p.32, 48) If $f : A \to B$ is an algebra morphism then :*
$Sp_B(f(X)) \subset Sp_A(X)$
$r_\lambda(f(X)) \leq r_\lambda(X)$

Theorem 1028 *(Rational Spectral Mapping theorem) (Thill p.31) For every element X in an algebra A, the rational map :*
$Q : A \to A :: Q(X) = \prod_k (X - \alpha_k I) \prod_l (X - \beta_l I)^{-1}$ *where all $\alpha_k \neq \beta_k$,*
$\beta_k \notin Sp(X)$
is such that : $Sp(Q(X)) = Q(Sp(X))$

Ptàk function

Definition 1029 *On a normed *-algebra A the Ptàk function is : $r_\sigma : A \to \mathbb{R}_+ :: r_\sigma(X) = \sqrt{r_\lambda(X^*X)}$*

Theorem 1030 *(Thill p.43, 44,120) The Ptàk function has the following properties :*
$r_\sigma(X) \leq \sqrt{\|X^*X\|}$
$r_\sigma(X^*) = r_\sigma(X)$
$r_\sigma(X^*X) = r_\sigma(X)^2$
If X is hermitian : $r_\lambda(X) = r_\sigma(X)$
If X is normal : $r_\sigma(X) = \|X\|$ and in a Banach *-algebra: $r_\lambda(X) \geq r_\sigma(X)$
the map r_σ is continuous at 0 and bounded in a neighborhood of 0

Hermitian algebra

For any eement : Sp(X*)=$\overline{Sp(X)}$ so for a self-adjoint X : Sp(X)=$\overline{Sp(X)}$ but it does not imply that each element of the spectrum is real.

Definition 1031 *A ***-**algebra is said to be **hermitian** if all its self-adjoints elements have a real spectrum*

Theorem 1032 *(Thill p.57) A closed *-algebra of a hermitian algebra is hermitian. A C*-algebra is hermitian.*

Theorem 1033 *(Thill p.56, 88) For a Banach *-algebra A the following conditions are equivalent :*
i) A is hermitian
ii) $\forall X \in A : X = X^* : i \notin Sp(X)$
iii) $\forall X \in A : r_\lambda(X) \leq r_\sigma(X)$
iv) $\forall X \in A : XX^* = X^*X \Rightarrow r_\lambda(X) = r_\sigma(X)$
v) $\forall X \in A : XX^* = X^*X \Rightarrow r_\lambda(X) \leq \|X^*X\|^{1/2}$
vi) $\forall X \in A :$ unitary $\Rightarrow Sp(X)$ is contained in the unit circle
vii) Shirali-Ford: $\forall X \in A : X^*X \geq 0$

12.4.6   Order on a *-algebra

If self adjoint elements have a real spectrum we can define a partial ordering on the self-adjoint elements of an algebra endowed with an involution.

Positive elements

Definition 1034 *On a *-algebra the set of positive elements denoted A^+ is the set of self-adjoint elements with positive spectrum*

$A^+ = \{X \geq 0\} = \{X \in A : X = X^*, Sp(X) \subset [0, \infty[\}$
A^+ is a cone in A

Theorem 1035 *(Thill p.85) If $f : A \rightarrow B$ is a *-morphism : $X \in A^+ \Rightarrow f(X) \in B^+$*

Square root

We say that Y is a square root for X if Y^2=X. There are no solution or usually at least two solutions (depending of A). In some conditions it is possible to distinguish one of the solution (as the square root of a real scalar) and it is denoted $X^{1/2}$.

Theorem 1036 *(Thill p.55) In a Banach algebra every element X such that $Sp(X) \subset]0, \infty[$ has a unique square root such that $Sp(X^{1/2}) \subset]0, \infty[$.*

Theorem 1037 *(Thill p.62) In a Banach *-algebra every invertible positive element X has a unique positive square root which is also invertible.*

Theorem 1038 *(Thill p.100,101) In a C*-algebra every positive element X has a unique positive square root. Conversely if there is Y such that X=Y^2 or X=Y*Y then X is positive.*

Theorem 1039 *(Thill p.51) The square root $X^{1/2}$ of X, when it exists, belongs to the closed subalgebra generated by X. If X=X* then $\left(X^{1/2}\right) = \left(X^{1/2}\right)^*$*

C*-algebra

A C*-algebra is hermitian, so all self-adjoint elements have a real spectrum and their set is well ordered by :

$X \geq Y \Leftrightarrow X - Y \geq 0 \Leftrightarrow X - Y$ have a spectrum in $\mathbb{R}_+$

Theorem 1040 *(Thill p.88) A^+ is a convex and closed cone*

For every X : $X^*X \geq 0$

Theorem 1041 *(Thill p.100,102) In a C*algebra the absolute value of every element X is $|X| = (X^*X)^{1/2}$ .It lies in the closed *-subalgebra generated by X. And we have : $\||X|\| = \|X\| , |X| \leq |Y| \Rightarrow \|X\| \leq \|Y\|$*

If f is an *-homomorphism between C*-algebras : $f\left(|X|\right) = |f\left(X\right)|$

Theorem 1042 *(Thill p.102) In a C*algebra, for every self-adjoint element X we have :*

$-\|X\| I \leq X \leq \|X\| I, -|X| \leq X \leq +|X| , 0 \leq X \leq Y \Rightarrow \|X\| \leq \|Y\|$

Theorem 1043 *(Thill p.100,103) In a C*-algebra every self-adjoint element has a unique decomposition : $X = X_+ - X_-$ such that $X_+, X_- \geq 0, X_+X_- = X_-X_+ = 0$*

It is given by : $X_+ = \frac{1}{2}\left(|X| + X\right), X_- = \frac{1}{2}\left(|X| - X\right)$

Theorem 1044 *(Thill p.95) In a C*-algebra every invertible element X has a unique polar decomposition : X=UP with P=$|X|$, $UU^* = I$*

12.4.7 Linear functionals

Linear functionals play a specific role. They can be used to build representations of the algebra on itself. In Quantum Mechanics they define the "mixed states".

Definitions

Definition 1045 *A **linear functional** on a topological algebra A is an element of its algebraic dual A'*

Definition 1046 *In a *-algebra A a linear functional φ is :*

i) **hermitian** *if $\forall X \in A : \varphi(X^*) = \overline{\varphi(X)}$*

ii) **positive** *if $\forall X \in A : \varphi(X^*X) \geq 0$*

The **variation** *of a positive linear functional is :*

$v(\varphi) = \inf_{X \in A} \left\{ \gamma : |\varphi(X)|^2 \leq \gamma \varphi(X^*X) \right\}.$

If it is finite then $|\varphi(X)|^2 \leq v(\varphi)\varphi(X^*X)$

iii) **weakly continuous** *if for every self-adjoint element X the map $Y \in A \to \varphi(Y^*XY)$ is continuous*

iv) a **quasi-state** *if it is positive, weakly continuous, and $v(\varphi) \leq 1$. The set of states will be denoted QS(A)*

iv) a **state** *if it is a quasi-state and $v(\varphi) = 1$. The set of states will be denoted S(A).*

v) a **pure state** *if it is an extreme point of S(A). The set of pure states is denoted PS(A).*

Theorem 1047 *(Thill p.139,140) QS(A) is the closed convex hull of $PS(A)\cup 0$, and a compact Hausdorff space in the *weak topology.*

Definition 1048 *In a *-algebra a positive linear functional φ_2 is* **subordinate** *to a positive linear fonctional φ_1 if $\forall \lambda \geq 0 : \lambda\varphi_2 - \varphi_1$ is a positive linear fonctional. A positive linear functional φ is* **indecomposable** *if any other positive linear functional subordinate to φ is a multiple of φ*

Theorems on linear functionals

Theorem 1049 *(Thill p.142, 144, 145) The variation of a positive linear functional φ on a normed *-algebra is finite and given by $v(\varphi) = \varphi(I)$. A positive linear functional on a Banach *-algebra is continuous and on a C*-algebra $v(\varphi) = \|\varphi\|$.*

Theorem 1050 *(Thill p.139) A quasi-state on a normed *-algebra is $\sigma-$contractive and hermitian*

Theorem 1051 *(Thill p.141,151) A state φ on a normed *-algebra is continuous and $\forall X \in A_+ : \varphi(X) \geq 0, \forall X \in A : \sqrt{\psi(X^*X)} \leq r_\sigma(X)$.*

Theorem 1052 *(Thill p.145) On a C*-algebra a state is a continuous linear functional such that $\|\varphi\| = \varphi(I) = 1$. Thenit is hermitian and $v(\varphi) = \|\varphi\|$*

Theorem 1053 *(Thill p.139) On a normed *-algebra, a state is pure iff it is indecomposable*

Theorem 1054 *(Thill p.158,173) On a Banach *-algebra A a state (resp a pure state) on a closed *-subalgebra can be extended to a state (resp. a pure state) if A is hermitian*

Theorem 1055 *(Thill p.146) If E is a locally compact Hausdorff topological space, for every state φ in $C_\nu(E;\mathbb{C})$ there is a unique inner regular Borel probability measure P on E such that : $\forall f \in C_\nu(E;\mathbb{C}) : \varphi(f) = \int_E fP$*

Theorem 1056 *If φ is a positive linear functional on a *-algebra A, then $\langle X, Y \rangle = \varphi(Y^*X)$ defines a sesquilinear form on A, called a Hilbert form.*

Multiplicative linear functionals

Definition 1057 *A **multiplicative linear functional** on a topological algebra is an element of the algebraic dual A' : $\varphi \in L(A;\mathbb{C})$ such that $\varphi(XY) = \varphi(X)\varphi(Y)$ and $\varphi \neq 0$*

$$\Rightarrow \varphi(I) = 1$$

Notation 1058 $\Delta(A)$ *is the set of multiplicative linear functionals on an algebra A.*

It is also sometimes denoted $\widehat{A}$.

Definition 1059 *For X fixed in an algebra A, the **Gel'fand transform** of X is the map : $\widehat{X} : \Delta(A) \to \mathbb{C} :: \widehat{X}(\varphi) = \varphi(X)$ and the map $\widehat{\ } : A \to C(\Delta(A);\mathbb{C})$ is the **Gel'fand transformation.***

The Gel'fand transformation is a morphism of algebras.

Using the Gel'fand transformation $\Delta(A) \subset A'$ can be endowed with the *weak topology, called Gel'fand topology. With this topology $\overline{\Delta(A)}$ is compact Hausdorff and $\overline{\Delta(A)} \sqsubseteq \Delta(A) \cup 0$

Theorem 1060 *(Thill p.68) For every topological algebra A, and $X \in A$: $\widehat{X}(\Delta(A)) \subset Sp(X)$*

Theorem 1061 *(Thill p.67, 68, 75) In a Banach algebra A:*
i) a multiplicative linear functional is continuous with norm $\|\varphi\| \leq 1$
ii) the Gel'fand transformation is a contractive morphism in $C_{0v}(\Delta(A);\mathbb{C})$
iii) $\Delta(A)$ is compact Hausdorff in the Gel'fand topology

Theorem 1062 *(Thill p.70, 71) In a commutative Banach algebra A:*
i) for every element $X \in A$: $\widehat{X}(\Delta(A)) = Sp(X)$
ii) (Wiener) An element X of A is not invertible iff $\exists\varphi \in \Delta(A) : \widehat{X}(\varphi) = 0$

Theorem 1063 *(Thill p.72) The set of multiplicative linear functional is not empty : $\Delta(A) \neq \varnothing$*

Theorem 1064 *Gel'fand - Naimark (Thill p.77) The Gel'fand transformation is a C*-algebra isomorphism between A and $C_{0v}(\Delta(A);\mathbb{C})$, the set of continuous, vanishing at infinity, functions on $\Delta(A)$.*

Theorem 1065 *(Thill p.79) For any Hausdorff, locally compact topological space, $\Delta(C_{0v}(E;\mathbb{C}))$ is homeomorphic to E.*

The homeomorphism is : $\delta : E \to \Delta(C_{0v}(E;\mathbb{C})) :: \delta_x(f) = f(x)$

12.5 Hilbert Spaces

12.5.1 Hilbert spaces

Definition

Definition 1066 *A complex **Hilbert space** is a complex Banach vector space whose norm is induced by a positive definite hermitian form. A real Hilbert space is a real Banach vector space whose norm is induced by a positive definite symmetric form*

As a real hermitian form is a symmetric form we will consider only complex Hilbert space, all results can be easily adjusted to the real case.

The hermitian form g will be considered as *antilinear in the first variable,* so :

$g(x,y) = \overline{g(y,x)}$

$g(x, ay + bz) = a g(x,y) + b g(x,z)$

$g(ax + by, z) = \bar{a} g(x,z) + \bar{b} g(y,z)$

$g(x, \bar{x}) \geq 0$

$g(x,x) = 0 \Rightarrow x = 0$

g is continuous. It induces a norm on H : $\|x\| = \sqrt{g(x,x)}$

Definition 1067 *A **pre-Hilbert** space is a complex normed vector space whose norm is induced by a positive definite hermitian form*

A normed space can always be "completed" to become a complete space.

Theorem 1068 *(Schwartz 2 p.9) If E is a separable complex vector space endowed with a definite positive sesquilinear form g, then its completion is a Hilbert space with a sesquilinear form which is the extension of g.*

But it is not always possible to deduce a sesquilinear form from a norm.

Let E be a vector space on the field K with a semi-norm $\|\|$. This semi norm is induced by :

- a sequilinear form iff K=$\mathbb{C}$ and $g(x,y) = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\left(\|x+iy\|^2 - \|x-iy\|^2\right)\right)$

is a sequilinear form (not necessarily definite positive).

- a symmetric bilinear forms iff K=$\mathbb{R}$ and $g(x,y) = \frac{1}{2}\left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right)$

is a symmetric bilinear form (not necessarily definite positive).

And the form g is necessarily unique for a given norm.

Similarly not any norm can lead to a Hilbert space : in $\mathbb{R}^n$ it is possible only with the euclidian norm.

Theorem 1069 *(Schwartz 2 p.21) Every closed vector subspace of a Hilbert space is a Hilbert space*

Warning! a vector subspace is not necessarily closed if H is infinite dimensional

Projection

Theorem 1070 *(Neeb p.227) For any vectors x,y in a Hilbert space H, the map : $P_{xy} : H \to H :: P_{xy}(u) = g(y,u)x$ is a continuous operator with the properties :*
$$P_{xy} = P_{yx}^*$$
$$\forall X, Y \in \mathcal{L}(H;H) : P_{Xx,Yy} = XP_{xy}Y^*$$

Theorem 1071 *(Schwartz 2 p.11) For every closed convex non empty subset F of a Hilbert space (H,g):*

i) $\forall u \in H, \forall v \in F : \operatorname{Re} g(u-v, u-v) \le 0$

ii) for any u∈ H there is a unique v∈ F such that : $\|u-v\| = \min_{w \in F} \|u-w\|$

*iii) the map $\pi_F : H \to F :: \pi_F(u) = v$,called the **projection** on F, is continuous.*

Theorem 1072 *(Schwartz 2 p.13) For every closed convex family $(F_n)_{n \in \mathbb{N}}$ subsets of a Hilbert space (H,g), such that their intersection F is non empty, and every vector $u \in H$, the sequence $(v_n)_{n \in \mathbb{N}}$ of the projections of u on each F_n converges to the projection v of u on F and $\|u - v_n\| \to \|u - v\|$*

Theorem 1073 *(Schwartz 2 p.15) For every closed convex family $(F_n)_{n \in \mathbb{N}}$ subsets of a Hilbert space (H,g), with union F, and every vector $u \in H$, the sequence $(v_n)_{n \in \mathbb{N}}$ of the projections of u on each F_n converges to the projection v of u on the closure of F and $\|u - v_n\| \to \|u - v\|$.*

Theorem 1074 *(Schwartz 2 p.18) For every closed vector subspace F of a Hilbert space (H,g) there is a unique projection $\pi_F : H \to F \in \mathcal{L}(F;H)$. If $F \ne \{0\}$ then $\|\pi_F\| = 1$*

If $u \in F : \pi_F(u) = u$

Orthogonal complement
2 vectors u,v are orthogonal if g(u,v)=0

Definition 1075 *The **orthogonal complement** $F^{\perp}$ of a vector subspace F of a Hilbert space H is the set of all vectors which are orthogonal to vectors of F.*

Theorem 1076 *(Schwartz 2 p.17) The orthogonal complement $F^{\perp}$ of a vector subspace F of a Hilbert space H is a closed vector subspace of H, which is also a Hilbert space and we have : $H = F \oplus F^{\perp}, F^{\perp\perp} = \overline{F}$*

Theorem 1077 *(Schwartz 2 p.19) For every finite family $(F_i)_{i \in I}$ of closed vector subspaces of a Hilbert space H :$(\cup_i F_i)^{\perp} = \cap_i F_i^{\perp}; (\cap_i F_i)^{\perp} = \overline{(\cup_i F_i^{\perp})}$*

Theorem 1078 *(Schwartz 2 p.16) A vector subspace F of a Hilbert space H is dense in H iff its orthogonal complement is 0*

If S is a *subset* of H then the orthogonal complement of S is the orthogonal complement of the linear span of S (intersection of all the vector subspaces containing S). It is a closed vector subspace, which is also a Hilbert space.

Quotient space

Theorem 1079 *(Schwartz 2 p.21) For every closed vector subspace F of a Hilbert space the quotient space H/F is a Hilbert space and the projection $\pi_F : F^\perp \to H/F$ is a Hilbert space isomorphism.*

Hilbert sum of Hilbert spaces

Theorem 1080 *(Neeb p.23, Schwartz 2 p.34) The **Hilbert sum,** denoted $H = \oplus_{i \in I} H_i$ of a family $(H_i, g_i)_{i \in I}$ of Hilbert spaces is the subset of families $(u_i)_{i \in I}$, $u_i \in H_i$ such that : $\sum_{i \in I} g_i (u_i, u_i) < \infty$. For every family $(u_i)_{i \in I} \in H$, $\sum_{i \in I} u_i$ is summable and H has the structure of a Hilbert space with the scalar product : $g(u, v) = \sum_{i \in I} g_i (u_i, v_i)$. The vector subspace generated by the H_i is dense in H.*

The sums are understood as :
$(u_i)_{i \in I} \in H \Leftrightarrow \forall J \subset I, card(J) < \infty : \sum_{i \in J} g_i (u_i, u_i) < \infty$
and :
$\exists u : \forall \varepsilon > 0, \forall J \subset I : card(J) < \infty, \sqrt{\sum_{i \notin J} \|u_i\|_{H_i}^2} < \varepsilon, \forall K : J \subset K \subset I : \left\| u - \sum_{i \in K} u_i \right\| < \varepsilon$
which implies that for any family of H only *countably* many u_i are non zero. So this is significantly different from the usual case.

The vector subspace generated by the H_i comprises any family $(u_i)_{i \in I}$ such that only *finitely* many u_i are non zero.

Definition 1081 *For a complete field K ($=\mathbb{R}, \mathbb{C}$) and any set I, the set $\ell^2(I)$ is the set of families $(x_i)_{i \in I}$ over K such that :*

$\left(\sup_{J \subset I} \sum_{i \in J} |x_i|^2 \right) < \infty$ *for any countable subset J of I. $\ell^2(I)$ is a Hilbert space with the sesquilinear form : $\langle x, y \rangle = \sum_{i \in I} \overline{x}_i y_i$*

Theorem 1082 *(Schwartz 2 p.37) $\ell^2(I), \ell^2(I')$ are isomorphic iff I and I' have the same cardinality.*

12.5.2 Hilbertian basis

Definition 1083 *A family $(e_i)_{i \in I}$ of vectors of a Hilbert space (H,g) is **orthornal** is $\forall i, j \in I : g(e_i, e_j) = \delta_{ij}$*

Theorem 1084 *(Schwartz 2 p.42) For any orthonormal family the map : $\ell^2(I) \to H :: y = \sum_i x_i e_i$ is an isomorphism of vector space from $\ell^2(I)$ to the closure $\overline{L}$ of the linear span L of $(e_i)_{i \in I}$ and*

Perceval inequality : $\forall x \in H : \sum_{i \in I} |g(e_i, x)|^2 \leq \|x\|^2$
Perceval equality   : $\forall x \in \overline{L} : \sum_{i \in I} |g(e_i, x)|^2 = \|x\|^2, \sum_{i \in I} g(e_i, x) e_i = x$

Definition 1085 A **Hilbertian basis** of H is an orthonormal family $(e_i)_{i \in I}$ such that the linear span of the family is dense in H. Equivalently if the only vector orthogonal to the family is 0.

$\forall x \in H : \sum_{i \in I} g(e_i, x) e_i = x, \sum_{i \in I} |g(e_i, x)|^2 = \|x\|^2$

$\forall x, y \in H : \sum_{i \in I} \overline{g(e_i, x)} g(e_i, y) = g(x, y)$

$\forall (x_i)_{i \in I} \in \ell^2(I)$ (which means $\left(\sup_{J \subset I} \sum_{i \in J} |x_i|^2 \right) < \infty$ for any countable subset J of I) then : $\sum_{i \in I} x_i e_i = x \in H$ and $(x_i)_{i \in I}$ is the unique family such that $\sum_{i \in I} y_i e_i = x$.

The quantities $g(e_i, x)$ are the Fourier coefficients.

Conversely a family $(e_i)_{i \in I}$ of vectors of H is a Hilbert basis iff :

$\forall x \in H : \sum_{i \in I} |g(e_i, x)|^2 = \|x\|^2$

Warning ! As vector space, a Hilbert space has bases, for which only a finite number of components are non zero. In a Hilbert basis there can be countably non zero components. So the two kinds of bases are not equivalent if H is infinite dimensional.

Theorem 1086 *(Schwartz 2 p.44) A Hilbert space has always a Hilbertian basis. All the Hilbertian bases of a Hilbert space have the same cardinality.*

Theorem 1087 *A Hilbert space is **separable** iff it has a Hilbert basis which is at most countable.*

Theorem 1088 *(Lang p.37) For every non empty closed disjoint subsets X,Y of a separable Hilbert space H there is a smooth function $f : H \to [0,1]$ such that f(x)=0 on X and f(x)=1 on Y.*

Ehrardt-Schmidt procedure :

It is the extension of the Graham Schmidt procecure to Hilbert spaces. Let $(u_n)_{n=1}^N$ be independant vectors in a Hilbert space H. Define :

$v_1 = u_1 / \|u_1\|$

$w_2 = u_2 - g(u_2, v_1) v_1$ and $v_2 = w_2 / \|w_2\|$

$w_p = u_p - \sum_{q=1}^{p-1} g(u_p, v_q) v_q$ and $v_p = w_p / \|w_p\|$

then the vectors $(u_n)_{n=1}^N$ are orthonormal.

Conjugate:

The conjugate of a vector can be defined if we have a real structure on the complex Hilbert space H, meaning an anti-linear map $:\sigma : H \to H$ such that $\sigma^2 = Id_H$. Then the conjugate of u is $\sigma(u)$.

The simplest way to define a real structure is by choosing a Hilbertian basis which is stated as real, then the conjugate $\overline{u}$ of $u = \sum_{i \in I} x_i e_i \in H$ is $\overline{u} = \sum_{i \in I} \overline{x_i} e_i$.

So we must keep in mind that conjugation is always with respect to some map, and practically to some Hermitian basis.

12.5.3   Operators

Linear endomorphisms on a Hilbert space are commonly called **operators** (in physics notably).

Theorem 1089 *The set of continuous linear maps $\mathcal{L}(H;H')$ between Hilbert spaces on the field K is a Banach vector space on the field K.*

Theorem 1090 *(Schwartz 2 p.20) Any continuous linear map $f \in \mathcal{L}(F;G)$ from the subspace F of a separable pre-Hilbert space E, to a complete topological vector space G can be extended to a continuous linear map $\widetilde{f} \in \mathcal{L}(E;G)$. If G is a normed space then $\left\|\widetilde{f}\right\|_{\mathcal{L}(E;G)} = \|f\|_{\mathcal{L}(F;G)}$*

The conjugate $\overline{f}$ (with respect to a real structure on H) of a linear endomorphism over a Hilbert space H is defined as : $\overline{f} : H \to H :: \overline{f}(u) = \overline{f(\overline{u})}$

Dual

One of the most important feature of Hilbert spaces is that there is an anti-isomorphism with the dual.

Theorem 1091 *(Riesz) Let (H,g) be a complex Hilbert space with hermitian form g, H' its topological dual. There is a continuous anti-isomorphism $\tau :$ $H' \to H$ such that :*
$\forall \lambda \in H', \forall u \in H : g(\tau(\lambda), u) = \lambda(u)$
(H',g*) is a Hilbert space with the hermitian form : $g^*(\lambda, \mu) = g(\tau(\mu), \tau(\lambda))$ and $\|\tau(\mu)\|_H = \|\mu\|_{H'}, \left\|\tau^{-1}(u)\right\|_{H'} = \|u\|_H$

Theorem 1092 *(Schwartz 2 p.27) A Hilbert space is reflexive : $(H')' = H$*

So :
for any $\varpi \in H'$ there is a unique $\tau(\varpi) \in H$ such that : $\forall u \in H :$ $g(\tau(\varpi), u) = \varpi(u)$ and conversely for any $u \in H$ there is a unique $\tau^{-1}(u) \in H'$ such that : $\forall v \in H : g(u, v) = \tau^{-1}(u)(v)$
$\tau(z\varpi) = \overline{z}\varpi, \tau^{-1}(zu) = \overline{z}u$
These relations are usually written in physics with the **bra-ket notation** :
a vector $u \in H$ is written $|u>$ (ket)
a form $\varpi \in H'$ is written $< \varpi|$ (bra)
the inner product of two vectors u,v is written $\langle u|v \rangle$
the action of the form ϖ on a vector u is written : $< \varpi||u >$ so $< \varpi|$ can be identified with $\tau(\varpi) \in H$ such that :
$\langle \tau(\varpi)|u \rangle =< \varpi||u >$
As a consequence :

Theorem 1093 *For every continuous sesquilinear map $B : H \times H \to \mathbb{C}$ in the Hilbert space H, there is a unique continuous endomorphism $A \in \mathcal{L}(H;H)$ such that $B(u, v) = g(Au, v)$*

Proof. Keep u fixed in H. The map : $B_u : H \to K :: B_u(v) = B(u, v)$ is continuous linear, so $\exists \lambda_u \in H' : B(u, v) = \lambda_u(v)$
Define : $A : H \to H :: A(u) = \tau(\lambda_u) \in H : B(u, v) = g(Au, v)$ ∎

Adjoint of a linear map

Theorem 1094 *(Schwartz 2 p.44) For every continuous linear maps f in $\mathcal{L}(H;H')$ between the Hilbert spaces $(H,g),(H',g')$ on the field K there is a map f^* in $\mathcal{L}(H';H)$ called the **adjoint** of f such that :*

$\forall u \in H, v \in H' : g(u, f^*v) = g'(fu, v)$

The map $*:\mathcal{L}(H;H')\rightarrow\mathcal{L}(H';H)$ is antilinear, bijective, continuous, isometric and $f^{**} = f, (f \circ g)^* = g^* \circ f^*$

If f is invertible, then f^* is invertible and $\left(f^{-1}\right)^* = (f^*)^{-1}$

There is a relation between transpose and adjoint : $f^*(v) = \overline{f^t(\overline{v})}$

$f \in \mathcal{L}(H;H') : f(H)^{\perp} = f^{*-1}(H), f^{-1}(0)^{\perp} = \overline{f^*(H')}$

Theorem 1095 *(Schwartz 2 p.47) f is injective iff $f^*(H)$ is dense in H', f(H) is dense in H' iff f^* is injective*

Compact operators

A continuous linear map $f \in \mathcal{L}(E;F)$ between Banach spaces E,F is compact if the the closure $\overline{f(X)}$ of the image of a bounded subset X of E is compact in F.

Theorem 1096 *(Schwartz 2 p.63) A continuous linear map $f \in \mathcal{L}(E;H)$ between a Banach space E and a Hilbert space H is compact iff it is the limit of a sequence $(f_n)_{n\in\mathbb{N}}$ of finite rank continuous maps in $\mathcal{L}(E;H)$*

Theorem 1097 *(Schwartz 2 p.64) The adjoint of a compact map between Hilbert spaces is compact.*

Hilbert sum of endomorphisms

Theorem 1098 *(Thill p.124) For a family of Hilbert space $(H_i)_{i\in I}$, a family of operators $(X_i)_{i\in I} : X_i \in \mathcal{L}(H_i; H_i)$, if $\sup_{i\in I} \|X_i\|_{H_i} < \infty$ there is a continuous operator on $\oplus_{i\in I}H_i$ with norm : $\|\oplus_{i\in I}H_i\| = \sup_{i\in I} \|X_i\|_{H_i}$, called the **Hilbert sum of the operators**, defined by : $(\oplus_{i\in I}X_i)(\oplus_{i\in I}u_i) = \oplus_{i\in I}X_i(u_i)$*

Topologies on $\mathcal{L}$(H;H)

On the space $\mathcal{L}$(H;H) of continuous endomorphisms of a Hilbert space H, we have topologies :

i) Strong operator topology, induced by the semi-norms : $u \in H : p_u(X) = \|Xu\|$

ii) Weak operator topology, weak topology induced by the functionals : $\mathcal{L}$(H;H)$\rightarrow \mathbb{C} :: u, v \in H : p_{u,v}(X) = g(u, Xv)$

iii) $\sigma-$strong topology, induced by the semi-norms : $p_U(X) = \sqrt{\sum_{n\in\mathbb{N}} \|Xu_n\|^2}, U = (u_n)_{n\in\mathbb{N}} : \sum_{n\in\mathbb{N}} \|u_n\|^2 < \infty$

iv) $\sigma-$weak topology, weak topology induced by the functionals :

$\mathcal{L}$(H;H)$\rightarrow \mathbb{C} :: p_{UV}(X) = \sum_{n \in \mathbb{N}} g(u_n, X v_n), U, V : \sum_{n \in \mathbb{N}} \|u_n\|^2 < \infty, \sum_{n \in \mathbb{N}} \|v_n\|^2 <$
∞

Weak operator topology < Strong operator topology < Norm topology

$\sigma-$weak toplogy < $\sigma-$strong toplogy < Norm topology

Weak operator topology < $\sigma-$weak toplogy

Strong operator topology < $\sigma-$strong toplogy

The $\sigma-$weak topology is the *weak topology induced by the trace class operators.

12.5.4 C*-algebra of continuous endomorphisms

With the map * which associates to each endomorphism its adjoint, the space $\mathcal{L}$(H;H) of endomorphisms on a Hilbert space over a field K is a C*-algebra over K.

So all the previous results can be fully implemented, with some simplifications and extensions.

General properties

All these results are the applications of theorems about C*-algebras.

For every endomorphism $f \in \mathcal{L}(H; H)$ on a Hilbert space on the field K :

$f^* \circ f$ is hermitian, and positive

$\overline{\exp f} = \exp \overline{f}$

$\exp f^* = (\exp f)^*$

The absolute value of f is : $|f| = (f^* f)^{1/2}$ and $\||f|\| = \|f\|, |f| \leq |g| \Rightarrow \|f\| \leq \|g\|$

The set of **unitary** endomorphism f in a Hilbert space : $f \in \mathcal{L}(H; H) :$ $f^* = f^{-1}$ is a closed subgroup of G$\mathcal{L}$(H;H).

Warning ! we must have both : f inversible and $f^* = f^{-1}$. $f^* \circ f = Id$ is not sufficient.

The set of invertible operators is an open subset and the map $f \rightarrow f^{-1}$ is continuous.

Every invertible element f has a unique polar decomposition : f=UP with P=$|f|, UU^* = I$

Theorem 1099 *Trotter Formula (Neeb p.172) If f,g are continuous operators in a Hilbert space over the field K, then :* $\forall k \in K : e^{k(f+g)} = \lim_{n \to \infty} (e^{k\frac{f}{n}} e^{k\frac{g}{n}})^n$

Theorem 1100 *(Schwartz 2 p.50) If K=$\mathbb{C}$:* $\frac{1}{2} \|f\| \leq \sup_{\|u\| \leq 1} |g(u, fu)| \leq \|f\|$

Hermitian maps

Definition 1101 $f \in \mathcal{L}(H; H)$ *is self adjoint (or hermitian) if f=f*, then* $\forall u, v \in H, : g(u, fv) = g(fu, v)$

$$\|f\| = \sup\nolimits_{\|u\| \le 1} |g\left(u, fu\right)|$$

Definition 1102 *A **symmetric map** on a Hilbert space (H,g) is a linear map $f \in L\left(D\left(f\right); H\right)$, where D(f) is a vector subspace of H, such that $\forall u, v \in D\left(f\right),:$ $g\left(u, fv\right) = g(fu, v)$*

Theorem 1103 *(Hellinger–Toeplitz theorem) A symmetric map $f \in L(H; H)$ on a Hilbert space H is continuous and self adjoint.*

The key condition is here that f is defined over the whole of H.

Theorem 1104 *(Thill p.104) For a continuous endomorphism f on a Hilbert space H the following conditions are equivalent :*
  i) f is hermitian positive : $f \ge 0$
  ii) $\forall u \in H : \langle u, fu \rangle \ge 0$

Spectrum

Theorem 1105 *The spectrum of an endomorphism f on a Hilbert space H is a non empty compact in $\mathbb{C}$, bounded by $r_\lambda\left(f\right) \le \|f\|$*
  $Sp(f^*) = \overline{Sp(f)}$
  If f is self-adjoint then its eigen values λ are real and $-\|f\| \le \lambda \le \|f\|$
  The spectrum of an unitary element is contained in the unit circle

Theorem 1106 *Riesz (Schwartz 2 p.68) The set of eigen values of a compact normal endomorphism f on a Hilbert space H on the field K is either finite, or countable in a sequence convergent to 0 (which is or not an eigen value). It is contained in a disc of radius $\|f\|$. If λ is eigen value for f, then $\overline{\lambda}$ is eigen value for f*. If K=$\mathbb{C}$, or if K=$\mathbb{R}$ and f is symmetric, then at least one eigen value equal to $\|f\|$. For each eigen value λ , except possibly for 0, the eigen space H_λ is finite dimensional. The eigen spaces are orthonormal for distinct eigen values. H is the direct Hilbert sum of the H_λ thus f can be written $u = \sum_\lambda u_\lambda \to fu = \sum_\lambda \lambda u_\lambda$ and f* : $f^*u = \sum_\lambda \overline{\lambda} u_\lambda$*
  Conversely if $(H_\lambda)_{\lambda \in \Lambda}$ is a family of closed, finite dimensional, orthogonal vector subspaces, with direct Hilbert sum H, then the operator $u = \sum_\lambda u_\lambda \to fu = \sum_\lambda \lambda u_\lambda$ is normal and compact

Hilbert-Schmidt operator
  This is the way to extend the definition of trace operator to Hilbert spaces.

Theorem 1107 *(Neeb p.228) For every endomorphism $f \in \mathcal{L}(H;H)$ of a Hilbert space H, and Hilbert basis $(e_i)_{i \in I}$ of H, the quantity $\|f\|_{HS} = \sqrt{\sum_{i \in I} g\left(fe_i, fe_i\right)}$ does not depend of the choice of the basis. If $\|f\|_{HS} < \infty$ then f is said to be a **Hilbert-Schmidt operator**.*

Notation 1108 *HS(H) is the set of Hilbert-Schmidt operators on the Hilbert space H.*

Theorem 1109 *(Neeb p.229) Hilbert Schmidt operators are compact*

Theorem 1110 *(Neeb p.228) For every Hilbert-Schmidt operators f,h$\in HS(H)$ on a Hilbert space H:*
$\|f\| \leq \|f\|_{HS} = \|f^*\|_{HS}$
$\langle f, h \rangle = \sum_{i \in I} g(e_i, f^* \circ h(e_i))$ *does not depend of the basis, converges and gives to HS(H) a structure of a Hilbert space such that* $\|f\|_{HS} = \sqrt{\langle f, f \rangle}$
$\langle f, h \rangle = \langle h^*, f^* \rangle$
If $f_1 \in \mathcal{L}(H; H), f_2, f_3 \in HS(H)$ *then :* $f_1 \circ f_2, f_1 \circ f_3 \in HS(H), \|f_1 \circ f_2\|_{HS} \leq \|f_1\| \|f_2\|_{HS}, \langle f_1 \circ f_2, f_3 \rangle = \langle f_2, f_1^* f_3 \rangle$

Trace

Definition 1111 *(Neeb p.230) A Hilbert-Schmidt endomorphism X on a Hilbert space H is **trace class** if*
$\|X\|_T = \sup \{|\langle X, Y \rangle|, Y \in HS(H), \|Y\| \leq 1\} < \infty$

Notation 1112 *T(H) is the set of trace class operators on th Hilbert space H*

Theorem 1113 *(Neeb p.231)* $\|X\|_T$ *is a norm on T(H) and T(H)$\sqsubseteq HS(H)$ is a Banach vector space with* $\|X\|_T$

Theorem 1114 *(Neeb p.230) The trace class operator X on a Hilbert space H has the following properties:*
$\|X\|_{HS} \leq \|X\|_T = \|X^*\|_T$
If $X \in \mathcal{L}(H; H), Y \in T(H)$ *then :* $XY \in T(H), \|XY\|_T \leq \|X\| \|Y\|_T$
If $X, Y \in HS(H)$ *then* $XY \in T(H)$

Theorem 1115 *(Taylor 1 p.502) A continuous endomorphism X on a Hilbert space is trace class iff it is compact and the set of eigen values of* $(X^*X)^{1/2}$ *is summable.*

Theorem 1116 *(Neeb p.231) For any trace class operator X on a Hilbert space H and any Hilbertian basis* $(e_i)_{i \in I}$ *of H, the sum* $\sum_{i \in I} g(e_i, Xe_i)$ *converges absolutely and :* $\sum_{i \in I} g(e_i, Xe_i) = Tr(X)$ *is the trace of X. It has the following properties:*
i) $|Tr(X)| \leq \|X\|_T$
ii) Tr(X) does not depend on the choice of a basis, and is a linear continuous functional on T(H)
iii) For $X, Y \in HS(H) : Tr(XY) = Tr(YX), \langle X, Y \rangle = Tr(XY^*)$
iv) For $X \in T(H)$ *the map :* $\mathcal{L}(H; H) \rightarrow \mathbb{C} :: Tr(YX)$ *is continuous, and Tr(XY)=Tr(YX).*
v) $\forall X \in \mathcal{L}(H; H) : \|X\|_T \leq \sum_{i,j \in I} |g(e_i, Xe_j)|$
vi) The space of continuous, finite rank, endomorphims on H is dense in T(H)

For H finite dimensional the trace coincides with the usual operator.

265

Irreducible operators

Definition 1117 *A continuous linear endomorphism on a Hilbert space H is **irreducible** if the only invariant closed subspaces are 0 and H. A set of operators is invariant if each of its operators is invariant.*

Theorem 1118 *(Lang p.521) For an irreducible set S of continuous linear endomorphism on a Hilbert space H. If f is a self-adjoint endomorphism commuting with all elements of S, then f=kId for some scalar k.*

Theorem 1119 *(Lang p.521) For an irreducible set S of continuous linear endomorphism on a Hilbert space H. If f is a normal endomorphism commuting,as its adjoint f*, with all elements of S, then f=kId for some scalar k.*

Ergodic theorem
    In mechanics a system is ergodic if the set of all its invariant states (in the configuration space) has either a null measure or is equal to the whole of the configuration space. Then it can be proven the the system converges to a state which does not depend on the initial state and is equal to the averadge of possible states. As the dynamic of such systems is usually represented as one parameter group of operators on Hilbert spaces, the topic has received a great attention.

Theorem 1120 *Alaoglu-Birkhoff (Bratelli 1 p.378) Let $\mathfrak{U}$ be a set of linear continuous endomorphisms on a Hilbert space H, such that : $\forall U \in \mathfrak{U} : \|U\| \leq 1, \forall U_1, U_2 \in \mathfrak{U} : U_1 \circ U_2 \in \mathfrak{U}$ and V the subspace of vectors invariant by all U: $V = \{u \in H, \forall U \in \mathfrak{U} : Uu = u\}$.*
    Then the orthogonal projection $\pi_V : H \to V$ belongs to the closure of the convex hull of $\mathfrak{U}$.

Theorem 1121 *For every unitary operator U on a Hilbert space $H : \forall u \in H : \lim_{n\to\infty} \frac{1}{n+1} \sum_{p=0}^{n} U^p u = Pu$ where P is the orthogonal projection on the subspace V of invariant vectors $u \in V : Uu = u$*

Proof. Take $\mathfrak{U} =$ the algebra generated by U in $\mathcal{L}(H; H)$ ∎

12.5.5   Unbounded operators

In physics it is necessary to work with linear maps which are not bounded, so not continuous, on the whole of the Hilbert space. The most common kinds of unbounded operators are operators defined on a dense subset and closed operators.

General definitions
    An unbounded operator is a linear map $X \in L(D(X); H)$ where D(X) is a vector subspace of H.

Definition 1122 *The **extension** of a linear map Y∈ L(D(Y); H), where H is a Hilbert space and D(X) a vector subspace of H is a linear map $X \in L(D(X); H)$ where $D(Y) \subset D(X)$ and X=Y on D(Y)*

It is usally denoted $Y \subset X$

Definition 1123 *The spectrum of a linear map $X \in L(D(X); H)$, where H is a Hilbert space and D(X) a vector subspace of H is the set of scalar $\lambda \in \mathbb{C}$ such that $\lambda I - X$ is injective and surjective on D(X) and has a bounded left-inverse*

X is said to be regular out of its spectrum

Definition 1124 *The adjoint of a linear map X∈ L(D(X); H), where H is a Hilbert space and D(X) a vector subspace of H is a map X*∈ $L(D(X^*); H)$ such that : $\forall u \in D(X), v \in D(X^*) : g(Xu, v) = g(u, X^*v)$*

The adjoint does not necessarily exist or be unique.

Definition 1125 *X is self-adjoint if X=X*, it is normal if XX*=X*X*

Theorem 1126 *(von Neumann) X*X and XX* are self-adjoint*

Definition 1127 *A **symmetric map** on a Hilbert space (H,g) is a linear map X∈ $L(D(X); H)$, where D(X) is a vector subspace of H, such that $\forall u, v \in D(X), : g(u, Xv) = g(Xu, v)$*

If X is symmetric, then $X \subset X^*$ and X can be extended on D(X*) but the extension is not necessarily unique.

Definition 1128 *A symmetric operator which has a unique extension which is self adjoint is said to be **essentially self-adjoint**.*

Theorem 1129 *(Hellinger–Toeplitz theorem) (Taylor 1 p.512) A symmetric map f∈ $L(H; H)$ on a Hilbert space H is continuous and self adjoint.*

The key condition is here that X is defined over the whole of H.

Definition 1130 *Two linear operators $X \in \mathcal{L}(D(X); H), Y \in \mathcal{L}(D(Y); H)$ on the Hilbert space H commute if :*
i) D(X) is invariant by Y : $YD(X) \subset D(X)$
ii) $YX \subset XY$

The set of maps in $\mathcal{L}(H;H)$ commuting with X is still called the commutant of X and denoted X'

Densely defined linear maps

Definition 1131 *A densely defined operator is a linear map X defined on a dense subspace D(X) of a Hilbert space*

Theorem 1132 *(Thill p.238, 242) A densely defined operator X has an adjoint X* which is a closed map.*
If X is self-adjoint then it is closed, X* is symmetric and has no symmetric extension.

Theorem 1133 *(Thill p.238, 242) If X,Y are densely defined operator then :*
i) $X \subset Y \Rightarrow Y^* \subset X^*$
ii) if XY is continuous on a dense domain then Y*X* is continuous on a dense domain and $Y^*X^* \subset (XY)^*$

Theorem 1134 *(Thill p.240,241) The spectrum of a self-adjoint,densely defined operator is a closed, locally compact subset of $\mathbb{R}$.*

Theorem 1135 *(Thill p.240, 246) The Cayley transform $Y = (X - iI)(X + iI)^{-1}$ of the densely defined operator X is an unitary operator and 1 is not an eigen value. If $\lambda \in Sp(X)$ then $(\lambda - i)(\lambda + i)^{-1} \in Sp(Y)$. Furthermore the commutants are such that Y'=X'. If X is self-adjoint then : $X = i(I + Y)(1 - Y)^{-1}$. Two self adjoint densely defined operators commute iff their Cayley transform commutes.*

If X is closed and densely defined, then X*X is self adjoint and I+X*X has a bounded inverse.

Closed linear maps

Definition 1136 *A linear map $X \in L(D(X); H)$, where H is a Hilbert space and D(X) a vector subspace of H is **closed** if its graph is closed in HxH.*

Definition 1137 *A linear map $X \in L(D(X); H)$ is **closable** if X has a closed extension denoted $\widetilde{X}$. Not all operators are closable.*

Theorem 1138 *A densely defined operator X is closable iff X* is densely defined. In this case $\widetilde{X}$=X** and $\left(\widetilde{X}\right)^* = X^*$*

Theorem 1139 *A linear map $X \in L(D(X); H)$ where D(X) is a vector subspace of an Hilbert space H, is closed if for every sequence $(u_n), u_n \in D(X)$ which converges in H to u, such that $Xu_n \to v \in H$ then : $u \in D(X)$ and v=Xu*

Theorem 1140 *(closed graph theorem) (Taylor 1 p.511) Any closed linear operator defined on the whole space H is bounded thus continuous.*

Theorem 1141 *The kernel of a closed linear map $X \in L(D(X); H)$ is a closed subspace of H*

Theorem 1142 *If the map X is closed and injective, then its inverse X^{-1} is also closed;*

Theorem 1143 *If the map X is closed then $X - \lambda I$ is closed where λ is a scalar and I is the identity function;*

Theorem 1144 *An operator X is closed and densely defined if and only if $X^{**} = X$*

12.5.6   Von Neumann algebra

Definition

Definition 1145 *A **von Neumann algebra** W denoted W^*-algebra is a *-subalgebra of $\mathcal{L}(H;H)$ for a Hilbert space H, such that $W=W$"*

Theorem 1146 *For every Hilbert space, $\mathcal{L}(H;H)$, its commutant $\mathcal{L}(H;H)$', $\mathbb{C}I$ are W^*-algebras.*

Theorem 1147 *(Thill p.203) A C^*-subalgebra A of $\mathcal{L}(H;H)$ is a W^*-algebra iff A"=A*

Theorem 1148 *(Thill p.204) If W is a von Neumann algebra then W' is a von Neumann algebra*

Theorem 1149 *Sakai (Bratelli 1 p.76) A C^*-algebra is isomorphic to a von Neumann algebra iff it is the dual of a Banach space.*

Properties

Theorem 1150 *(Thill p.206) For Hilbert space H and any subset S of $\mathcal{L}(H;H)$ the smallest W^*-algebra which contains S is $W(S)=(S \cup S^*)$". If $\forall X, Y \in S$: $X^* \in S, XY = YX$ then $W(S)$ is commutative.*

Theorem 1151 *von Neumann density theorem (Bratelli 1 p.74) If B is a *-subalgebra of $\mathcal{L}(H;H)$ for a Hilbert space H, such that the orthogonal projection on the closure of the linear span $Span\{Xu, X \in B, u \in H\}$ is H, then B is dense in B"*

Theorem 1152 *(Bratelli 1 p.76) A state φ of a von Neumann algebra W in $\mathcal{L}(H;H)$ is normal iff there is positive, trace class operator ρ in $\mathcal{L}(H;H)$ such that : $Tr(\rho) = 1, \forall X \in W : \varphi(X) = Tr(\rho X)$.*

ρ is called a density operator.

Theorem 1153 *(Neeb p.152) Every von Neumann algebra A is equal to the bi-commutant P" of the set P of projections belonging to A : $P = \left\{ p \in A : p = p^2 = p^* \right\}$*

Theorem 1154 *(Thill p.207) A von Neuman algebra is the closure of the linear span of its projections.*

12.5.7 Reproducing Kernel

All vector spaces on the same field, of the same dimension, and endowed with a definite positive form g are isometric. So they are characterized by g. We have something similar for infinite dimensional Hilbert spaces of functions over a topological space E. In a Hilbert basis the scalar product $g(e_i, e_i)$ can be in some way linked to the values of $g(e_i(x), e_i(y))$ for x,y in E. With a reproducing kernel it is then possible to build other Hilbert spaces of functions over E.

Definitions

Definition 1155 *For any set E and field K=$\mathbb{R}, \mathbb{C}$, a function $N : E \times E \to K$ is a **definite positive kernel** of E if :*
 i) it is definite positive : for any finite set $(x_1, ..., x_n)$ the matrix $[N(x_i, x_j)]_{n \times n} \subset K(n)$ is semi definite positive : $[X]^* [N(x_i, x_j)] [X] \geq 0$ with $[X] = [x_i]_{n \times 1}$.
 *ii) it is either **symmetric** (if $\underline{K=\mathbb{R}}$) : $N(x, y)^* = N(y, x) = N(x, y)$, or **hermitian** (if K=$\mathbb{C}$) : $N(x, y)^* = \overline{N(y, x)} = N(x, y)$*

 Then $|N(x, y)|^2 \leq |N(x, x)| |N(y, y)|$
 A Hilbert space defines a reproducing kernel:
 Let (H,g) be Hilbert space (H,g), on a field K, of functions $f : E \to K$ on a topological space E. If the evaluation maps : $x \in E : \widehat{x} : H \to K :: \widehat{x}(f) = f(x)$ are continuous, then $\widehat{x} \in H'$, and there is $N_x \in H$ such that :
 $\forall x \in E, f \in H : \exists N_x \in H : g(N_x, f) = \widehat{x}(f) = f(x)$
 The corresponding function : $N : E \times E \to K :: N(x, y) = N_y(x)$ is called the **reproducing kernel** of H.
 Conversely reproducing kernel defines a Hilbert space:

Theorem 1156 *(Neeb p.55) If $N : E \times E \to K$ is a positive definite kernel of E, then :*
 i) $H_0 = Span\{N(x, .), x \in E\}$ carries a unique positive definite hermitian form g such that :
 $\forall x, y \in E : g(N_x, N_y) = N(x, y)$
 ii) the completion H of H_0 with injection : $\imath : H_0 \to H$ carries a Hilbert space structure H consistent with this scalar product, and whose reproducing kernel is N.
 iii) this Hilbert space is unique

Theorem 1157 *(Neeb p.55) A function $N : E \times E \to K$ is positive definite iff it is the reducing kernel of some Hilbert space $H \subset C(E; K)$*

5. Examples (Neeb p.59)

i) Let (H,g) be a Hilbert space, E any set, $f \in C(E;H)$ then : $N(x,y) = g(f(x), f(y))$ is a positive definite kernel

ii) If P is a positive definite kernel of E, $f \in C(E;H)$, then $Q(x,y) = \overline{f(x)}P(x,y)f(y)$ is a positive definite kernel of E

iii) If P is a positive definite kernel of E, $f \in C(F;E)$, then $Q(x,y) = P(f(x), f(y))$ is a positive definite kernel of F

iv) Let (H,g) be a Hilbert space, take N(x,y)=g(x,y) then $H_N = H'$

v) Fock space : let H be a complex Hilbert space. Then $N : H \times H \to \mathbb{C} ::$ $N(u,v) = \exp g(u,v)$ is a positive definite kernel of H. The corresponding Hilbert space is the symmetric Fock space $\mathfrak{F}(H)$.

Properties

Theorem 1158 *(Neeb p.55) If $N : E \times E \to K$ is the reproducing kernel of the Hilbert space H, then :*

i) N is definite positive

ii) $H_0 = Span\{N(x,.), x \in E\}$ is dense in H

iii) For any orthonormal basis $(e_i)_{i \in I}$ of H : $N(x,y) = \sum_{i \in I} g(e_i(x), e_i(y))$

(remember that the vectors of H are functions)

Theorem 1159 *(Neeb p.57) The set N(E) of positive definite kernels of a topological space E is a convex cone in K^E which is closed under pointwise convergence and pointwise multiplication :*

$\forall P, Q \in N(E), \lambda \in \mathbb{R}_+ : P + Q \in N(E), \lambda P \in N(E), (PQ)(x,y) = P(x,y)Q(x,y) \in N(E)$

If $K = \mathbb{C} : P \in N(E) \Rightarrow \operatorname{Im} P \in N(E), |P| \in N(E)$

Theorem 1160 *(Neeb p.57) Let (T, S, μ) a measured space, $(P_t)_{t \in T}$ a family of positive definite kernels of E, such that $\forall x, y \in E$ the maps : $t \to P_t(x,y)$ are measurable and the maps : $t \to P_t(x,x)$ are integrable, then : $P(x,y) = \int_T P_t(x,y) \mu(t)$ is a positive definite kernel of E.*

Theorem 1161 *(Neeb p.59) If the series : $f(z) = \sum_{n=0}^{\infty} a_n z^n$ over K is convergent for $|z| < r$, if P is a positive definite kernel of E and $\forall x, y \in E$: $|P(x,y)| < r$ then : $f(P)(x,y) = \sum_{n=0}^{\infty} a_n P(x,y)^n$ is a positive definite kernel of E.*

Theorem 1162 *(Neeb p.60) For any positive definite kernel P of E, there are : a Hilbert space H, a map : $f : E \to H$ such that f(E) spans a dense subset of H. Then $Q(x,y) = g(f(x), f(y))$ is the corresponding reproducing kernel. The set (E,H,f) is called a triple realization of P. For any other triple (E,H',f') there is a unique isometry : $\varphi : H \to H'$ such that $f' = \varphi \circ f$*

Tensor product of Hilbert spaces

The definition of the tensorial product of two vector spaces on the same field extends to Hilbert spaces.

Theorem 1163 *(Neeb p.87) If $(e_i)_{i \in I}$ is a Hilbert basis of H and $(f_j)_{j \in J}$ is a Hilbert basis of F then $\sum_{(i,j) \in I \times J} e_i \otimes f_j$ is a Hilbert basis of $H \otimes F$*

The scalar product is defined as : $\langle u_1 \otimes v_1, u_2 \otimes v_2 \rangle = g_H(u_1, u_2) \, g_F(v_1, v_2)$

The reproducing Kernel is : $N_{H \otimes F}(u_1 \otimes v_1, u_2 \otimes v_2) = g_H(u_1, u_2) \, g_F(v_1, v_2)$

13 SPECTRAL THEORY

The C*-algebras have been modelled on the set of continuous linear maps on a Hilbert space, so it is natural to look for representations of C*algebras on Hilbert spaces. In quantum physics most of the work is done through "representations of observables" : one starts with a set of observables, with an algebra structure, and look for a representation on a Hilbert space.

In many ways this topic looks like the representation of Lie groups. One of the most useful outcome of this endeavour is the spectral theory which enables to resume the action of an operator as an integral with measures which are projections on eigen spaces.

On this subject we follow mainly Thill. See also Bratelli.

13.1 Representation of an algebra

13.1.1 General Properties

Representations

Definition 1164 *A linear representation of an algebra* $(A, \cdot)$ *over the field K is a pair (H,ρ) of a vector space H over the field K and the algebra morphism* $\rho : A \to L(H; H) :$

$\forall X, Y \in A, k, k' \in K :$
$\rho(kX + k'Y) = k\rho(X) + k'\rho(Y)$
$\rho(X \cdot Y) = \rho(X) \circ \rho(Y)$
$\rho(I) = Id \Rightarrow$ if $X \in G(A) : \rho(X)^{-1} = \rho(X^{-1})$

Definition 1165 *A linear representation of a *-algebra* $(A, \cdot)$ *over the field K is a linear representation (H,ρ) of A, such that H is endowed with an involution and :* $\forall X \in A : \rho(X^*) = \rho(X)^*$

In the following we will consider representation (H, ρ) **of a Banach *-algebra A on a Hilbert space (H,g).**

Definition 1166 *A Hilbertian representation of a Banach *-algebra A is a linear representation (H,ρ) of A, where H is a Hilbert space, and ρ is a continuous *-morphism* $\rho : A \to \mathcal{L}(H; H)$.

So: $\forall u \in H, X \in A : g(\rho(X)u, v) = g(u, \rho(X^*)v)$ with the adjoint X* of X.

The adjoint map $\rho(X)^*$ is well defined if $\rho(X)$ is continuous on H or at least on a dense subset of H

$\rho \in \mathcal{L}(A; \mathcal{L}(H; H))$ and we have the norm : $\|\rho\| = \sup_{\|X\|_A=1} \|\rho(X)\|_{\mathcal{L}(H;H)} < \infty$

Properties of a representation

1.Usual definitions of representation theory for any linear representation (H, ρ) of A:

i) the representation is faithful if ρ is injective

ii) a vector subspace F of H is invariant if $\forall u \in F, \forall X \in A : \rho(X) u \in F$

iii) (H, ρ) is irreducible if there is no other invariant vector space than 0,H.

iv) If $(H_k, \rho_k)_{k \in I}$ is a family of Hilbertian representations of A, then and $\forall X \in A, \|\rho_k(X)\| < \infty$ the Hilbert sum of representations $(\oplus_i H_i, \oplus_i \rho_i)$ is defined with : $(\oplus_i \rho_i)(X)(\oplus_i u_i) = \oplus_i (\rho_i(X) u_i)$ and norm $\|\oplus_i \rho_i\| = \sup_{i \in I} \|\rho_i\|$

v) An operator $f \in \mathcal{L}(H_1; H_2)$ is an interwiner between two representations $(H_k, \rho_k)_{k=1,2}$ if :

$\forall X \in A : f \circ \rho_1(X) = \rho_2(X) \circ f$

vi) Two representations are equivalent if there is an interwiner which an isomorphism

vii) A representation (H, ρ) is contractive if $\|\rho\| \leq 1$

viii) A representation (H, ρ) of the algebra A is isometric if $\forall X \in A : \|\rho(X)\|_{\mathcal{L}(H;H)} = \|X\|_A$

2. Special definitions :

Definition 1167 *The **commutant** ρ' of the linear representation (H, ρ) of a algebra A is the set $\{\pi \in \mathcal{L}(H; H) : \forall X \in A : \pi \circ \rho(X) = \rho(X) \circ \pi\}$*

Definition 1168 *A vector $u \in H$ is **cyclic** for the linear representation (H, ρ) of a algebra A if the set $\{\rho(X) u, X \in A\}$ is dense in H. (H, ρ) is said cyclic if there is a cyclic vector u_c and is denoted (H, ρ, u_c)*

Definition 1169 *Two linear representations $(H_1, \rho_1), (H_2, \rho_2)$ of the algebra A are **spatially equivalent** if there is a unitary interwiner $U : U \circ \rho_1(X) = \rho_2(X) U$*

General theorems

Theorem 1170 *(Thill p.125) If the vector subspace $F \subset H$ is invariant in the linear representation (H, ρ) of A, then the orthogonal complement $F^\perp$ is also invariant and (F, ρ) is a subrepresentation*

(Thill p.125 A closed vector subspace $F \subset H$ is invariant in the linear representation (H, ρ) of A iff $\forall X \in A : \pi_F \circ \rho(X) = \rho(X) \circ \pi_F$ where $\pi_F : H \to F$ the projection on F

Theorem 1171 *If (H, ρ) is a linear representation of A, then for every unitary map $U \in \mathcal{L}(H; H)$, $(H, U\rho U^*)$ is an equivalent representation.*

Theorem 1172 *(Thill p.122) Every linear representation of a Banach *-algebra with isometric involution on a pre Hilbert space is contractive*

Theorem 1173 *(Thill p.122) Every linear representation of a C*algebra on a pre Hilbert space is contractive*

Theorem 1174 *(Thill p.122) Every faithful linear representation of a C*algebra on a Hilbert space is isometric*

Theorem 1175 *If (H, ρ) is a linear representation of a *-algebra then the commutant ρ' is a W*-algebra.*

Theorem 1176 *(Thill p.123) For every linear representation (H, ρ) of a C*algebra A: $A/\ker \rho, \rho(A)$ are C*-algebras and the representation factors to : $A/\ker \rho \to \rho(A)$*

Theorem 1177 *(Thill p.127) For every linear representation (H, ρ) of a Banach *-algebra, and any non null vector $u \in H$, the closure of the linear span of $F = \{\rho(X) u, X \in A\}$ is invariant and (F, ρ, u) is cyclic*

Theorem 1178 *(Thill p.129) If $(H_1, \rho_1, u_1), (H_2, \rho_2, u_2)$ are two cyclic linear representations of Banach *-algebra A and if $\forall X \in A : g_1(\rho_1(X) u_1, u_1) = g_2(\rho_2(X) u_2, u_2)$ then the representations are equivalent and there is a unitary operator U: $U \circ \rho_1(X) \circ U^* = \rho_2(X)$*

Theorem 1179 *(Thill p.136) For every linear representation (H, ρ) of a C*algebra A and vector u in H such that: $\|u\| = 1$,the map : $\varphi : A \to \mathbb{C} :: \varphi(X) = g(\rho(X) u, u)$ is a state*

13.1.2 Representation GNS

A Lie group can be represented on its Lie algebra through the adjoint representation. Similarly an algebra has a linear representation on itself. Roughly $\rho(X)$ is the translation operator $\rho(X) Y = XY$. A Hilbert space structure on A is required..

Theorem 1180 *(Thill p.139, 141) For any linear positive functional φ, a Banach *-algebra has a Hilbertian representation, called **GNS** (for Gel'fand, Naimark, Segal) and denoted $(H_\varphi, \rho_\varphi)$, which is continuous and contractive.*

The construct is the following:

i) Any linear positive functional φ on A define the sesquilinear form : $\langle X, Y \rangle = \varphi(Y^* X)$ called a Hilbert form

ii) It can be null for non null X,Y. Let J=$\{X \in A : \forall Y \in A : \langle X, Y \rangle = 0\}$. It is a left ideal of A and we can pass to the quotient A/J: Define the equivalence relation : $X \sim Y \Leftrightarrow X - Y \in J$. A class x in A/J is comprised of elements of the kind : X + J

iii) Define on A/J the sesquilinear form : $\langle x, y \rangle_{A/J} = \langle X, Y \rangle_A$. So A/J becomes a pre Hilbert space which can be completed to get a Hilbert space H_φ.

iv) For each x in A/J define the operator on A/J : T(x)y=xy .If T is bounded it can be extended to the Hilbert space H_φ and we get a representation of A.

v) There is a vector $u_\varphi \in H_\varphi$ such that : $\forall X \in A : \varphi(X) = \langle x, u_\varphi \rangle$, $v(\varphi) = \langle u_\varphi, u_\varphi \rangle$. u_φ can be taken as the class of equivalence of I.

If φ is a state then the representation is cyclic with cyclic vector u_φ such that $\varphi(X) = \langle T(X) u_\varphi, u_\varphi \rangle$, $v(\varphi) = \langle u_\varphi, u_\varphi \rangle = 1$

Conversely:

Theorem 1181 *(Thill p.140) If (H, ρ, u_v) is a cyclic linear representation representation of the Banach *-algebra A, then each cyclic vector u_c of norm 1 defines a state $\varphi(X) = g(\rho(X) u_c, u_c)$ such that the associated representation $(H_\varphi, \rho_\varphi, u_\varphi)$ is equivalent to (H, ρ) and $\rho_\varphi = U \circ \rho \circ U^*$ for an unitary operator. The cyclic vectors are related by $U : u_\varphi = U u_c$*

So each cyclic representation of A on a Hilbert space can be labelled by the equivalent GNS representation, meaning labelled by a state. Up to equivalence the GNS representation $(H_\varphi, \rho_\varphi)$ associated to a state φ is defined by the condition :

$\varphi(X) = \langle \rho_\varphi(X) u_\varphi, u_\varphi \rangle$

Any other cyclic representation (H, ρ, u_c) such that : $\varphi(X) = \langle \rho(X) u_c, u_c \rangle$ is equivalent to $(H_\varphi, \rho_\varphi)$

13.1.3 Universal representation

The universal representation is similar to the sum of finite dimensional representations of a compact Lie group : it contains all the classes of equivalent representations. As any representation is sum of cyclic representations, and that any cyclic representation is equivalent to a GNS representation, we get all the representations with the sum of GNS representations.

Theorem 1182 *(Thill p.152) The **universal representation** of the Banach *-algebra A is the sum :$\left(\oplus_{\varphi \in S(A)} H_\varphi ; \oplus_{\varphi \in S(A)} \rho_\varphi \right) = (H_u, \rho_u)$ where $(H_\varphi, \rho_\varphi)$ is the GNS representation $(H_\varphi, \rho_\varphi)$ associated to the state φ and S(A) is the set of states on A. It is a $\sigma-$contractive Hilbertian representation and $\|\rho_u(X)\| \leq p(X)$ where p is the semi-norm : $p(X) = \sup_{\varphi \in S(A)} (\varphi(X^*X))^{1/2}$.*

This semi-norm is well defined as : $\forall X \in A, \varphi \in S(A) : \varphi(X) \leq p(X) \leq r_\sigma(X) \leq \|X\|$ and is required to sum the GNS representations.

1. The subset rad(A) of A such that p(X)=0 is a two-sided ideal, * stable and closed, called the **radical**.

2. The quotient set A/rad(A) with the norm p(X) is a pre C*-algebra whose completion is a C*-algebra denoted C*(A) called the envelopping C*algebra of A. The map : $j : A \to C^*(A)$ is a *-algebra morphism, continuous and j(C*(A)) is dense in C*(A).

To a representation (H, ρ_*) of C*(A) one associates a unique representation (H, ρ) of A by : $\rho = \rho_* \circ j$.

3. A is said **semi-simple** if rad(A)=0. Then A with the norm p is a pre-C*-algebra whose completion if C*(A).

If A has a faithful representation then A is semi-simple.

Theorem 1183 *(Gelfand-Naïmark) (Thill p.159) if A is a C*-algebra :* $\|\rho_u(X)\| = p(X) = r_\sigma(X)$

The universal representation is a C* isomorphism between A and the set $\mathcal{L}(H; H)$ *of a Hilbert space, thus C*(A) can be assimilated to A*

5. If A is commutative and the set of its multiplicative linear functionals $\Delta(A) \neq \varnothing$, then C*(A) is isomorphic as a C*-algebra to the set $C_{0v}(\Delta(A); \mathbb{C})$ of continuous functions vanishing at infinity.

13.1.4   Irreducible representations

Theorem 1184 *(Thill p.169) For every Hilbertian representation* (H, ρ) *of a Banach *-algebra the following are equivalent :*
 i) (H, ρ) *is irreducible*
 ii) any non null vector is cyclic
 iii) the commutant ρ' *of* ρ *is the set* $zI, z \in \mathbb{C}$

Theorem 1185 *(Thill p.166) If the Hilbertian representation* (H, ρ) *of a Banach *-algebra A is irreducible then, for any vectors u,v of H such that* $\forall X \in A : g(\rho(X)u, u) = g(\rho(X)v, v) : \exists z \in \mathbb{C}, |z| = 1 : v = zu$

Theorem 1186 *(Thill p.171) For every Hilbertian representation* (H, ρ) *of a Banach *-algebra the following are equivalent :*
 i) φ *is a pure state*
 ii) φ *is indecomposable*
 iii) $(H_\varphi, \rho_\varphi)$ *is irreducble*

Thus the pure states label the irreducible representations of A up to equivalence

Theorem 1187 *(Thill p.166) A Hilbertian representation of a commutative algebra is irreducible iff it is unidimensional*

13.2   Spectral theory

Spectral theory is a general method to replace a linear map on an infinite dimensional vector space by an integral. It is based on the following idea. Let $X \in L(E; E)$ be a diagonalizable operator on a finite dimensional vector space. On each of its eigen space E_λ it acts by $u \rightarrow \lambda u$ thus X can be written as : $X = \sum_\lambda \lambda \pi_\lambda$ where π_λ is the projection on E_λ (which can be uniquely defined if we have a bilinear symmetric form). If E is infinite dimensional then we can hope to replace $\sum$ by an integral. For an operator on a Hilbert space the same idea involves the spectrum of X and an integral. The interest lies in the fact that many properties of X can be studied through the spectrum, meaning a set of complex numbers. Several steps are necessary to address the subject.

13.2.1   Spectral measure

Definition 1188 *A **spectral measure** defined on a mesurable space (E,S) and acting on a Hilbert space (H,g) is a map $P : S \to \mathcal{L}(H; H)$ such that:*

    i) $\forall \varpi \in S : P(\varpi) = P(\varpi)^* = P(\varpi)^2 : P(\varpi)$ *is a projection*

    ii) P(E)=I

    iii) $\forall u \in H$ *the map :* $\varpi \to g(P(\varpi)u, u) = \|P(\varpi)u\|^2 \in \mathbb{R}_+$ *is a finite measure on (E,S).*

    Thus if g(u,u)=1 $\|P(\varpi)u\|^2$ is a probability

    For u,v in H we define a bounded complex measure by :

    $\langle Pu, v \rangle (\varpi) = \frac{1}{4} \sum_{k=1}^{4} i^k g \left(P(\varpi) \left(u + i^k v \right), \left(u + i^k v \right) \right) \Rightarrow \langle Pu, v \rangle (\varpi) = \langle P(\varpi)v, u \rangle$

    The support of P is the complement in E of the largest open subset on which P=0

Theorem 1189 *(Thill p.184, 191) A spectral measure P has the following properties :*

    i) P is finitely additive : for any finite disjointed family $(\varpi_i)_{i \in I}, \varpi_i \in S, \forall i \neq j : \varpi_i \cap \varpi_j = \varnothing : P(\cup_i \varpi_i) = \sum_i P(\varpi_i)$

    ii) $\forall \varpi_1, \varpi_2 \in S : \varpi_1 \cap \varpi_2 = \varnothing : P(\varpi_1) \circ P(\varpi_2) = 0$

    iii) $\forall \varpi_1, \varpi_2 \in S : P(\varpi_1) \circ P(\varpi_2) = P(\varpi_1 \cap \varpi_2)$

    iv) $\forall \varpi_1, \varpi_2 \in S : P(\varpi_1) \circ P(\varpi_2) = P(\varpi_2) \circ P(\varpi_1)$

    v) If the sequence $(\varpi_n)_{n \in \mathbb{N}}$ *in S is disjointed or increasing then* $\forall u \in H :$ $P(\cup_{n \in \mathbb{N}} \varpi_n) u = \sum_{n \in \mathbb{N}} P(\varpi_n) u$

    vi) $\overline{Span\,(P(\varpi))_{\varpi \in S}}$ *is a commutative C*-subalgebra of $\mathcal{L}(H,H)$*

    Warning ! P is not a measure on (E,S), $P(\varpi) \in \mathcal{L}(H; H)$

    A property is said to hold P almost everywhere in E if $\forall u \in H$ if holds almost everywhere in E for $g(P(\varpi)u, u)$

    Image of a spectral measure : let (F,S') another Borel measurable space, and $\varphi : E \to F$ a measurable map, then P defines a spectral measure on (F,S') by : $\varphi^* P(\varpi') = P(\varphi^{-1}(\varpi'))$

Examples

  (Neeb p.145)

    1. Let (E,S,μ) be a measured space. Then the set $L^2(E, S, \mu, \mathbb{C})$ is a Hilbert space. The map :

    $\varpi \in S : P(\varpi)\varphi = \chi_\varpi \varphi$ where χ_ϖ is the characteristic function of ϖ , is a spectral measure on $L^2(E, S, \mu, \mathbb{C})$

    2. Let $H = \oplus_{i \in I} H_i$ be a Hilbert sum, define $P(J)$ as the orthogonal projection on the closure : $\overline{(\oplus_{i \in J} H_i)}$. This is a spectral measure

    3. If we have a family $(P_i)_{i \in I}$ of spectral measures on some space (E,S), each valued in $\mathcal{L}(H_i; H_i)$, then :

    $P(\varpi)u = \sum_{i \in I} P_i(\varpi) u_i$ is a spectral measure on H=$\oplus_{i \in I} H_i$.

13.2.2 Spectral integral

For a measured space (E, S, μ), a bounded function $f \in C_b(E; \mathbb{C})$, a Hilbert space H and a map $P : S \to \mathcal{L}(H; H)$, for each $\varpi \in S : f(\varpi) P(\varpi) \in \mathcal{L}(H; H)$ so we can consider an integral $\int_E f(\varpi) P(\varpi) \mu$ which will be some linear map on H. The definition of the integral of a function with a real valued measure is given in the Measure section. Here we have to extend the concept to a measure valued in $\mathcal{L}(H; H)$, proceeding along a similar line.

Definition

Theorem 1190 *If P is a spectral measure on the space (E,S), acting on the Hilbert space (H,g), a complex valued measurable bounded function on E is **P-integrable** if there is $X \in \mathcal{L}(H; H)$ such that :*

$\forall u, v \in H : g(Xu, v) = \int_E f(\varpi) g(P(\varpi) u, v)$
*If so X is unique and called the **spectral integral** of f : $X = \int_P f P$*

The contruct is the following (Thill p.185).

1. A step function is given by a finite set I, a partition $(\varpi_i)_{i \in I}$ of E such that $\varpi_i \in S$, and a family of complex scalars $(\alpha_i)_{i \in I} \in \ell^2(I)$ by : $f = \sum_{i \in I} \alpha_i 1_{\varpi_i}$, where 1_{ϖ_i} is the characteristic function of ϖ_i

The set $C_b(E; \mathbb{C})$ of complex valued measurable bounded functions in E, endowed with the norm: $\|f\| = \sup |f|$ is a commutative C*-algebra with the involution : $f^* = \overline{f}$.

The set $C_S(E; \mathbb{C})$ of complex valued step functions on (E,S) is a C*-subalgebra of $C_b(E; \mathbb{C})$

2. For $h \in C_S(E; \mathbb{C})$ define the integral $\rho_S(h) = \int_E h(\varpi) P(\varpi) = \sum_{i \in I} \alpha_i h(\varpi_i) P(\varpi_i) \in \mathcal{L}(H; H)$

H with the map $\rho_S : C_S(E; \mathbb{C}) \to \mathcal{L}(H; H)$ defines a representation (H, ρ) of $C_S(E; \mathbb{C})$

We have : $\forall u \in H : g\left(\left(\int_E h(\varpi) P(\varpi)\right) u, u\right) = \int_E h(\varpi) g(P(\varpi) u, u)$

3. We say that $f \in C_b(E; \mathbb{C})$ is P integrable (in norm) if there is $X \in \mathcal{L}(H; H)$

$\forall h \in C_S(E; \mathbb{C}) : \left\| X - \int_E h(\varpi) P(\varpi) \right\|_{\mathcal{L}(H; H)} \leq \|f - h\|_{C_b(E; \mathbb{C})}$

We say that $f \in C_b(E; \mathbb{C})$ is P integrable (weakly) if there is $Y \in \mathcal{L}(H; H)$

such that : $\forall u \in H : g(Yu, u) = \int_E f(\varpi) g(P(\varpi) u, u)$

4. f P integrable (in norm) $\Rightarrow$ f P integrable (weakly) and there is a unique $X = Y = \rho_b(f) = \int_E f P \in \mathcal{L}(H; H)$

5. conversely f P integrable (weakly) $\Rightarrow$ f P integrable (in norm)

Remark : the norm on a C*-algebra of functions is necessarily equivalent to : $\|f\| = \sup_{x \in E} |f(x)|$ (see Functional analysis). So the theorem holds for any C*-algebra of functions on E.

Properties of the spectral integral

Theorem 1191 *(Thill p.188) For every P integrable function f:*

i) $\left\| \left(\int_E fP \right) u \right\|_H = \sqrt{\int_E |f|^2 g\left(P\left(\varpi\right)u,u\right)}$

ii) $\int_E fP = 0 \Leftrightarrow f = 0$ *P almost everywhere*

iii) $\int_E fP \geq 0 \Leftrightarrow f \geq 0$ *P almost everywhere*

Notice that the two last results are inusual.

Theorem 1192 *(Thill p.188) For a spectral measure P on the space (E,S), acting on the hilbert space (H,g), H and the map :* $\rho_b : C_b\left(E;\mathbb{C}\right) \to \mathcal{L}\left(H;H\right) :: \rho_b\left(f\right) = \int_E fP$ *is a representation of the C*-algebra* $C_b\left(E;\mathbb{C}\right)$*.* $\rho_b\left(C_b\left(E;\mathbb{C}\right)\right) = \overline{Span\left(P\left(\varpi\right)\right)_{\varpi \in S}}$ *is the C*-subalgebra of* $\mathcal{L}(H,H)$ *generated by P and the commutants :* $\rho' = Span\left(P\left(\varpi\right)\right)'_{\varpi \in S}$*.*

Every projection in $\rho_b\left(C_b\left(E;\mathbb{C}\right)\right)$ *is of the form :* $P(s)$ *for some* $s \in S$*.*

Theorem 1193 *Monotone convergence theorem (Thill p.190) If P is a spectral measure P on the space (E,S), acting on the hilbert space (H,g),* $\left(f_n\right)_{n \in \mathbb{N}}$ *an increasing bounded sequence of real valued mesurable functions on E, bounded P almost everywhere, then* $f = \lim f_n \in C_b\left(E;\mathbb{R}\right)$ *and* $\int fP = \lim \int f_n P$ *,* $\int fP$ *is self adjoint and* $\forall u \in H : g\left(\left(\int_E fP\right)u,u\right) = \lim \int_E f_n\left(\varpi\right)g\left(P\left(\varpi\right)u,u\right)$

Theorem 1194 *Dominated convergence theorem (Thill p.190) If P is a spectral measure P on the space (E,S), acting on the hilbert space (H,g),* $\left(f_n\right)_{n \in \mathbb{N}}$ *a norm bounded sequence of functions in* $C_b\left(E;\mathbb{C}\right)$ *which converges pointwise to f,then :* $\forall u \in H : \left(\int fP\right)u = \lim\left(\int f_n P\right)u$

Theorem 1195 *Image of a spectral measure (Thill p.192) : If P is a spectral measure P on the space (E,S), acting on the hilbert space (H,g),(F,S') another Borel measurable space, and* $\varphi : E \to F$ *a mesurable map then :* $\forall h \in C_b\left(F;\mathbb{C}\right) :$ $\int_F h\varphi^* P = \int_E \left(h \circ \varphi\right)P$

13.2.3 Spectral resolution

The purpose is now, conversely, starting from an operator X, find f and a spectral measure P such that X=$\int_E f\left(\varpi\right)P\left(\varpi\right)$

Exitence

Definition 1196 *A **resolution of identity** is a spectral measure on a measurable Hausdorff space (E,S) acting on a Hilbert space (H,g) such that for any* $u \in H$*,* $g(u,u) = 1 : g\left(P\left(\varpi\right)u,u\right)$ *is inner regular.*

Theorem 1197 *(Thill p.197) For any continuous normal operator X on a Hilbert space H there is a unique resolution of identity :* $P : Sp(X) \to \mathcal{L}(H;H)$ *called the **spectral resolution** of X such that :* $X = \int_{Sp(X)} zP$ *where Sp(X) is the spectrum of X*

X *normal* : X*X=XX*

so the function f is here the identity map : $Id : Sp(X) \to Sp(X)$

We have a sometimes more convenient formulation of this theorem

Theorem 1198 *(Taylor 2 p.72) Let X be a self adjoint operator on a separable Hilbert space H, then there is a Borel measure μ on $\mathbb{R}$, a unitary map $W:L^2(\mathbb{R}, \mu, \mathbb{C}) \to H$, a real valued function $a \in L^2(\mathbb{R}, \mu, \mathbb{R})$ such that :*
$\forall \varphi \in L^2(\mathbb{R}, \mu, \mathbb{C}) : W^{-1}XW\varphi(x) = a(x)\varphi(x)$

Theorem 1199 *(Taylor 2 p.79) If $A_k, k = 1..n$ are commuting, self adjoint continuous operators on a Hilbert space H, there are a measured space (E,μ), a unitary map : $W:L^2(E, \mu, \mathbb{C}) \to H$, functions $a_k \in L^\infty(E, \mu, \mathbb{R})$ such that :*
$\forall f \in L^2(E, \mu, \mathbb{C}) : W^{-1}A_kW(f)(x) = a_k(x)f(x)$

Commutative algebras

For any algebra (see multiplicative linear functionals in Normed algebras) :

$\Delta(A) \in \mathcal{L}(A; \mathbb{C})$ is the set of multiplicative linear functionals on A

$\widehat{X} : \Delta(A) \to \mathbb{C} :: \widehat{X}(\varphi) = \varphi(X)$ is the Gel'fand transform of X

Theorem 1200 *Representation of a commutative *-algebra (Thill p.201) For every Hilbertian representation (H, ρ) of a commutative *-algebra A, there is a unique resolution of identity P sur $Sp(\rho)$ acting on H such that : $\forall X \in A$:* $\rho(X) = \int_{Sp(X)} \widehat{X}|_{Sp(X)}P$ *and Sup(P)=Sp(ρ)*

Theorem 1201 *Representation of a commutative Banach *-algebra (Neeb p.152) For any Banach commutative *-algebra A :*

i) If P is a spectral measure on $\Delta(A)$ then $\rho(X) = P\left(\widehat{X}\right)$ defines a spectral measure on A

ii) If (H,ρ) is a non degenerate Hilbertian representation of A, then there is a unique spectral measure P on $\Delta(A)$ such that $\rho(X) = P\left(\widehat{X}\right)$

Theorem 1202 *(Thill p.194) For every Hilbert space H, commutative C*-subalgebra A of $\mathcal{L}(H;H)$, there is a unique resolution of identity $P : \Delta(A) \to \mathcal{L}(H;H)$ such that : $\forall X \in A : X = \int_{\Delta(A)} \widehat{X}P$*

Properties of the spectral resolution

Theorem 1203 *If P is the spectral resolution of X :*
i) Support of P = all of Sp(X)
ii) Commutants : X'=Span $(P(z))'_{z \in Sp(X)}$

Theorem 1204 *Eigen-values (Thill p.198) If P is the spectral resolution of the continuous normal operator on a Hilbert space H, $\lambda \in Sp(X)$ is an eigen value of X iff $P(\{\lambda\}) \neq 0$. Then the range of $P(\lambda)$ is the eigen space relative to λ*

So the eigne values of X are the isolated points of its spectrum.

13.2.4 Extension to unbounded operators

(see Hilbert spaces for definitions)

Spectral integral

Theorem 1205 *(Thill p.233) If P is a spectral measure on the space (E,S), acting on the Hilbert space (H,g), for each complex valued measurable function f on (E,S) there is a linear map $X=\int fP$ called the **spectral integral** of f, defined on a subspace D(X) of H such that : $\forall u \in D(X) : g(Xu,u) = \int_E f(\varpi) g(P(\varpi)u,u)$ and $D(\int fP) = \left\{ u \in H : \int_E |g(u,fu)P|^2 < \infty \right\}$ is dense in H*

Comments:

1) the conditions on f are very weak : almost any function is integrable

2) the difference with the previous spectral integral is that $\int fP$ is neither necessarily defined over the whole of H, nor continuous

The consruct is the following (Thill p.233)

i) For each complex valued measurable function f on (E,S) $D(f)=\left\{ u \in H : \int_E |g(u,fu)P|^2 < \infty \right\}$ is dense in H

ii) one says that f is weakly integrable if :$\exists X \in L(D(X);H) : D(X)=D(f)$ and $\forall u \in H : g(Xu,u) = \int_E f(\varpi) g(P(\varpi)u,u)$

one says that f is pointwise integrable if :$\exists X \in L(D(X);H) : D(X)=D(f)$ and

$\forall h \in C_b(E;\mathbb{C}), \forall u \in H : \left\| \left(X - \int_E hP \right) u \right\|^2 = \sqrt{\int_E \|f(\varpi) - h(\varpi)\|^2 g(P(\varpi)u,u)}$

iii) f is weakly integrable $\Rightarrow$ f is pointwise integrable and X is unique.

For any complex valued measurable function f on (E,S) there exists a unique $X=\Psi_P(f)$ such that $X = \int_E fP$ pointwise

f is pointwise integrable $\Rightarrow$ f is weakly integrable

Properties of the spectral integral

Theorem 1206 *(Thill p.236, 237, 240) If P is a spectral measure on the space (E,S), acting on the Hilbert space (H,g), and f,f_1, f_2 are complex valued measurable functions on (E,S) :*

i) $\forall u \in D(f) : \left\| \left(\int_E f(\varpi) P(\varpi) \right) u \right\|_H = \sqrt{\int_E |f|^2 g(P(\varpi)u,u)}$

ii) $D(|f_1| + |f_2|) = D\left(\int_E f_1 P + \int_E f_2 P \right)$

$D\left(\left(\int_E f_1 P \right) \circ \left(\int_E (f_2) P \right) \right) = D(f_1 \circ f_2) \cap D(f_2)$

which reads with the meaning of extension of operators (see Hilbert spaces)

$\int_E f_1 P + \int_E f_2 P \subset \int_E (f_1 + f_2) P$

$\left(\int_E f_1 P \right) \circ \left(\int_E (f_2) P \right) \subset \int_E (f_1 f_2) P$

iii) $\left(\int_E fP \right)^* = \int_E \overline{f} P$ *so if f is a measurable real valued function on E then* $\int_E fP$ *is self-adjoint*

$\int_E fP$ is a closed map

$$\left(\int_E fP\right)^* \circ \left(\int_E fP\right) = \left(\int_E fP\right) \circ \left(\int_E fP\right)^* = \int_E |f|^2 P$$

Theorem 1207 *Image of a spectral measure (Thill p.236) : If P is a spectral measure on the space (E,S), acting on the Hilbert space (H,g), (F,S') a Borel measurable space, and $\varphi : E \to F$ a mesurable map then for any complex valued measurable functions on (F,S') : $\int_F f\varphi^* P = \int_E (f \circ \varphi) P$*

Spectral resolution

It is the converse of the previous result.

Theorem 1208 *(Spectral theorem for unbounded operators) (Thill p.243) For every densely defined, linear, self-adjoint operator X in the Hilbert space H, there is a unique resolution of identity $P : Sp(X) \to L(H;H)$ called the **spectral resolution** of X, such that : $X = \int_{Sp(X)} \lambda P$ where Sp(X) is the spectrum of X.*

(the function f is real valued and equal to the identity)
We have a sometimes more convenient formulation of this theorem

Theorem 1209 *(Taylor 2 p.79) Let X be a self adjoint operator, defined on a dense subset D(X) of a separable Hilbert space H, then there is a measured space (E,μ), a unitary map $W:L^2(E,\mu,\mathbb{C}) \to H$, a real valued function $a \in L^2(E,\mu,\mathbb{R})$ such that :*

$\forall \varphi \in L^2(E,\mu,\mathbb{C}) : W^{-1}XW\varphi(x) = a(x)\varphi(x)$
$W\varphi \in D(X)$ *iff* $\varphi \in L^2(E,\mu,\mathbb{C})$

If f is a bounded measurable function on E, then : $W^{-1}f(X)W\varphi(x) = f(a(x))\varphi(x)$ defines a bounded operator f(X) on $L^2(E,\mu,\mathbb{C})$

With $f(x) = e^{ia(x)}$ we get the get the strongly continuous one parameter group $e^{iXt} = U(t)$ with generator iX.

Theorem 1210 *(Thill p.243) The spectral resolution has the following properties:*

i) Support of P = all of Sp(X)
ii) Commutants : $X'=Span\,(P(\lambda))'_{\lambda \in Sp(X)}$

Theorem 1211 *(Thill p.246) If P is the spectral resolution of a densely self adjoint operator on the Hilbert space H, $f : Sp(X) \to \mathbb{C}$ a Borel measurable function, then $\int_E fP$ is well defined on $D\left(\int_E fP\right)$ and denoted f(X)*

13.2.5   Application to one parameter unitary groups

One parameters groups are seen in the Banach Spaces subsection. Here we address some frequently used results, notably in quantum physics.

Theorem 1212 *(Thill p.247) A map : $U : \mathbb{R} \to \mathcal{L}(H;H)$ such that :*

i) U(t) is unitary

ii) U(t+s)=U(t)U(s)=U(s)U(t)

defines a one parameter unitary group on a Hilbert space H.

If $\forall u \in H$ the map : $\mathbb{R} \to H :: U(t)u$ is continuous then U is differentiable, and there is an infinitesimal generator $S \in L(D(S), H)$ such that : $\forall u \in D(S)$: $-\frac{1}{i}\frac{d}{dt}U(t)|_{t=0}u = Su$ which reads $U(t) = \exp(itS)$

We have a sometime more convenient formulation of this theorem :

Theorem 1213 *(Taylor 2 p.76) Let H be a Hilbert space and U a map $U : \mathbb{R} \to \mathcal{L}(H;H)$ which defines an uniformly continuous one parameter group, having a cyclic vector v, then there exists a positive Borel measure μ on $\mathbb{R}$ and a unitary map : $W : L^2(\mathbb{R}, \mu, \mathbb{C}) \to H$ such that : $\forall \varphi \in L^2(\mathbb{R}, \mu, \mathbb{C}) : W^{-1}U(t)W\varphi(x) = e^{itx}\varphi(x)$*

The measure $\mu = \widehat{\zeta}(t)\,dt$ where $\zeta(t) = \langle v, U(t)v \rangle$ is a tempered distribution.

Conversely :

Theorem 1214 *(Thill p.247) For every self adjoint operator S defined on a dense domain D(X) of a Hilbert space H, the map : $U : \mathbb{R} \to \mathcal{L}(H;H) :: U(t) = \exp(-itS) = \int_{Sp(S)} (-it\lambda) P(\lambda)$ defines a one parameter unitary group on H with infinitesimal generator S. U is differentiable and $-\frac{1}{i}\frac{d}{ds}U(s)|_{s=t}u = SU(t)u$*

So U is the solution to the problem : $-\frac{1}{i}\frac{d}{ds}U(s)|_{s=t} = SU(t)$ with the initial value solution U(0)=S

Remark : U(t) is the Fourier transform of S

Part IV

PART 4 : DIFFERENTIAL GEOMETRY

Differential geometry is the extension of elementary geometry and deals with manifolds. Nowodays it is customary to address many issues of differential geometry with the fiber bundle formalism. However a more traditional approach is sometimes useful, and enables to start working with the main concepts without the hurdle of getting acquainted with a new theory. So we will deal with fiber bundles later, after the review of Lie groups.

Many concepts and theorems about manifolds can be seen as extensions from the study of derivatives in affine normed spaces. So we will start with a comprehensive review of derivatives in this context.

14 DERIVATIVE

In this section we will address the general theory of derivative of a map (non necessarily linear) between affine normed spaces. It leads to some classic results about extremum and implicit functions. We will also introduce holomorphic functions.

We will follow mainly Schwartz (t I).

14.1 Differentiable maps

14.1.1 Definitions

In elementary analysis the derivative of a function f(x) in a point a is introduced as $f'(x)|_{x=0} = \lim_{h \to 0} \frac{1}{h} (f(a+h) - f(a))$. This idea can be generalized once we have normed linear spaces. As the derivative is taken at a point, the right structure is an affine space (of course a vector space is an affine space and the results can be fully implemented in this case) .

Differentiable at a point

Definition 1215 *A map $f : \Omega \to F$ defined on an open subset Ω of the normed affine space $\left(E, \overrightarrow{E} \right)$ and valued in the normed affine space $\left(F, \overrightarrow{F} \right)$, both on the same field K, is **differentiable** at $a \in \Omega$ if there is a linear, continuous map $L \in \mathcal{L} \left(\overrightarrow{E}; \overrightarrow{F} \right)$ such that :*

$\exists r > 0, \forall \overrightarrow{h} \in \overrightarrow{E}, \left\| \overrightarrow{h} \right\|_E < r : a + \overrightarrow{h} \in \Omega : f(a + \overrightarrow{h}) - f(a) = L \overrightarrow{h} + \varepsilon(h) \left\| \overrightarrow{h} \right\|_F$

where $\varepsilon(h) \in \overrightarrow{F}$ is such that $\lim_{h \to 0} \varepsilon(h) = 0$
*L is called the **derivative** of f in a.*

Speaking plainly : f can be approximated by an affine map in the neighborhood of a: $f(a + \overrightarrow{h}) \simeq f(a) + L \overrightarrow{h}$

Theorem 1216 *If the derivative exists, it is unique and f is continuous in a.*

This derivative is often called Fréchet's derivative. If we take E=F=$\mathbb{R}$ we get back the usual definition of a derivative.

Notice that $f(a + \overrightarrow{h}) - f(a) \in \overrightarrow{F}$ and that no assumption is made about the dimension of E,F or the field K, but E and F must be on the same field (because a linear map must be between vector spaces on the same field). This remark will be important when K=$\mathbb{C}$.

Remark :the domain Ω must be open. If Ω is a closed subset and a$\in \partial \Omega$ then we must have $a + \overrightarrow{h} \in \overset{\circ}{\Omega}$ and L may not be defined over $\overrightarrow{E}$. If E=$[a, b] \subset \mathbb{R}$ one can define right derivative at a and left derivative at b because L is a scalar.

Theorem 1217 *(Schwartz II p.83) A map $f : \Omega \to F$ defined on an open subset Ω of the normed affine space $\left(E, \overrightarrow{E}\right)$ and valued in the normed affine space $\left(F, \overrightarrow{F}\right) = \prod_{i=1}^{r} \left(F_i, \overrightarrow{F}_i\right)$, both on the same field K, is differentiable at $a \in \Omega$ iff each of its components $f_k : E \to F_k$ is differentiable at a and its derivative f'(a) is the linear map in $\mathcal{L}\left(\overrightarrow{E}; \prod_{i=1}^{r} \overrightarrow{F}_i\right)$ defined by $f'_k(a)$.*

Continuously differentiable in an open subset

Definition 1218 *A map $f : \Omega \to F$ defined on an open subset Ω of the normed affine space $\left(E, \overrightarrow{E}\right)$ and valued in the normed affine space $\left(F, \overrightarrow{F}\right)$ both on the same field K, is differentiable in Ω if it is differentiable at each point of Ω. Then the map : $f\prime : \Omega \to \mathcal{L}\left(\overrightarrow{E}; \overrightarrow{F}\right)$ is the derivative map or more simply derivative, of f in Ω. If f' is continuous f is said to be **continuously differentiable** or of class 1 (or C_1).*

Notation 1219 *f' is the derivative of f : $f\prime : \Omega \to \mathcal{L}\left(\overrightarrow{E}; \overrightarrow{F}\right)$*

$f\prime(a) = f\prime(x)\mid_{x=a}$ is the value of the derivative in a . So f'(a)$\in \mathcal{L}\left(\overrightarrow{E}; \overrightarrow{F}\right)$
$C_1(\Omega; F)$ is the set of continuously differentiable maps $f : \Omega \to F$.

If E,F are vector spaces then $C_1(\Omega; F)$ is a vector space and the map which associates to each map $f : \Omega \to F$ its derivative is a linear map on the space $C_1(\Omega; F)$.

Theorem 1220 *(Schwartz II p.87) If the map $f : \Omega \to F$ defined on an open subset Ω of the normed affine space $\left(E, \overrightarrow{E}\right)$ and valued in the normed affine space $\left(F, \overrightarrow{F}\right)$ both on the same field K, is continuously differentiable in Ω then the map :$\Omega \times \overrightarrow{E} \to \overrightarrow{F} :: f'(x)\overrightarrow{u}$ is continuous.*

Differentiable along a vector

Definition 1221 *A map $f : \Omega \to F$ defined on an open subset Ω of the normed affine space $\left(E, \overrightarrow{E}\right)$ and valued in the normed affine space $\left(F, \overrightarrow{F}\right)$ on the same field K, is **differentiable** at $a \in \Omega$ **along the vector** $\overrightarrow{u} \in \overrightarrow{E}$ if there is $\overrightarrow{v} \in \overrightarrow{F}$ such that : $\lim_{z \to 0}\left(\frac{1}{z}\left(f(a + z\overrightarrow{u}) - f(a)\right)\right) = \overrightarrow{v}$. $\overrightarrow{v}$ is the derivative of f in a with respect to the vector $\overrightarrow{u}$*

Notation 1222 *$D_u f(a) \in \overrightarrow{F}$ is the derivative of f in a with respect to the vector $\overrightarrow{u}$*

Definition 1223 *A map $f : \Omega \to F$ defined on an open subset Ω of the normed affine space $\left(E, \overrightarrow{E}\right)$ and valued in the normed affine space $\left(F, \overrightarrow{F}\right)$ on the same field K, is **Gâteaux differentiable** at $a \in \Omega$ if there is $L \in L\left(\overrightarrow{E}; \overrightarrow{F}\right)$ such that $: \forall \overrightarrow{u} \in \overrightarrow{E} : \lim_{z \to 0}\left(\frac{1}{z}\left(f(a + z\overrightarrow{u}) - f(a)\right)\right) = L\overrightarrow{u}$.*

Theorem 1224 *If f is differentiable at a, then it is Gâteaux **differentiable** and $D_u f = f'(a)\overrightarrow{u}$.*

But *the converse is not true* : there are maps which are Gâteaux differentiable and not even continuous ! But if $\forall \varepsilon > 0, \exists r > 0, \forall \overrightarrow{u} \in \overrightarrow{E} : \|\overrightarrow{u}\|_E < r :$ $\|\varphi(z) - \overrightarrow{v}\| < \varepsilon$ then f is differentiable in a.

Partial derivatives

Definition 1225 *A map $f : \Omega \to F$ defined on an open subset Ω of the normed affine space $\left(E, \overrightarrow{E}\right) = \prod\limits_{i=1}^{r}\left(E_i, \overrightarrow{E}_i\right)$ and valued in the normed affine space $\left(F, \overrightarrow{F}\right)$, all on the same field K, has a **partial derivative** at $a=(a_1, .a_r) \in \Omega$ with respect to the variable k if the map: $f_k : \Omega_k = \pi_k(\Omega) \to F :: f_k(x_k) = f(a_1, ...a_{k-1}, x_k, a_{k+1}, ..a_r)$, where π_k is the canonical projection $\pi_k : E \to E_k$, is differentiable at a*

Notation 1226 $\frac{\partial f}{\partial x_k}(a) = f'_{x_k}(a)$ *denotes the value of the partial derivative at a with respect to the variable x_k. $\frac{\partial f}{\partial x_k}(a) \in \mathcal{L}\left(\overrightarrow{E}_k; \overrightarrow{F}\right)$*

Definition 1227 *If f has a **partial derivative** with respect to the variable x_k at each point $a \in \Omega$, and if the map : $a_k \to \frac{\partial f}{\partial x_k}(a)$ is continuous, then f is said to be continuously differentiable with respect to the variable x_k in Ω*

Notice that a partial derivative does not necessarily refers to a basis.
If f is differentiable at a then it has a partial derivative with respect to each of its variable and :
$f'(a)(\overrightarrow{u}_1, ...\overrightarrow{u}_r) = \sum_{i=1}^{r} f'_{x_i}(a)(\overrightarrow{u}_i)$
But the *converse is not true*. We have the following :

Theorem 1228 *(Schwartz II p.118) A map $f : \Omega \to F$ defined on an open subset Ω of the normed affine space $\left(E, \overrightarrow{E}\right) = \prod\limits_{i=1}^{r}\left(E_i, \overrightarrow{E}_i\right)$ and valued in the normed affine space $\left(F, \overrightarrow{F}\right)$, all on the same field K, which is continuously differentiable in Ω with respect to each of its variable is continuously differentiable in Ω*

So f continuously differentiable in $\Omega \Leftrightarrow$ f has continuous partial derivatives in Ω

but f has partial derivatives in a $\nRightarrow$ f is differentiable in a

Notice that the E_i and F can be infinite dimensional. We just need a finite product of normed vector spaces.

Coordinates expressions

Let f be a map $f : \Omega \to F$ defined on an open subset Ω of the normed affine space $\left(E, \overrightarrow{E}\right)$ and valued in the normed affine space $\left(F, \overrightarrow{F}\right)$ on the same field K.

1. If E is a m dimensional affine space, it can be seen as the product of n one dimensional affine spaces and, with a basis $(\overrightarrow{e}_i)_{i=1}^m$ of $\overrightarrow{E}$ we have :

The value of f'(a) along the basis vector $\overrightarrow{e}_i$ is $D_{\overrightarrow{e}_i} f(a) = f'(a)\left(\overrightarrow{e}_i\right) \in \overrightarrow{F}$

The partial derivative with respect to x_i is : $\frac{\partial f}{\partial x_i}(a)$ and : $D_{\overrightarrow{e}_i} f(a) = \frac{\partial f}{\partial x_i}(a)\left(\overrightarrow{e}_i\right)$

The value of f'(a) along the vector $\overrightarrow{u} = \sum_{i=1}^m u_i \overrightarrow{e}_i$ is $D_{\overrightarrow{u}} f(a) = f'(a)\left(\overrightarrow{u}\right) = \sum_{i=1}^m u_i D_{\overrightarrow{e}_i} f(a) = \sum_{i=1}^m u_i \frac{\partial f}{\partial x_i}(a)\left(\overrightarrow{e}_i\right) \in \overrightarrow{F}$

2. If F is a n dimensional affine space, with a basis $\left(\overrightarrow{f}_i\right)_{i=1}^n$ we have :

$f(x) = \sum_{k=1}^n f_k(x)$ where $f_k(x)$ are the coordinates of f(x) in a frame $\left(O, \left(\overrightarrow{f}_i\right)_{i=1}^n\right)$.

$f'(a) = \sum_{k=1}^n f'_k(a)\overrightarrow{f}_k$ where $f'_k(a) \in K$

3. If E is m dimensional and F n dimensional, the map f'(a) is represented by a matrix J with n rows and m columns, each column being the matrix of a partial derivative, called the **jacobian** of f:

$$[f'] = J = \left\{\overbrace{\left[\frac{\partial f_j}{\partial x_i}\right]}^{m}\right\} n = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & .. & \frac{\partial f_1}{\partial x_m} \\ ... & .. & ... \\ \frac{\partial f_n}{\partial x_1} & .. & \frac{\partial f_n}{\partial x_m} \end{bmatrix}$$

If E=F the determinant of J is the determinant of the linear map f'(a), thus it does not depend on the basis.

14.1.2   Properties of the derivative

Derivative of linear maps

Theorem 1229 *A continuous affine map $f : \Omega \to F$ defined on an open subset Ω of the normed affine space $\left(E, \overrightarrow{E}\right)$ and valued in the normed affine space $\left(F, \overrightarrow{F}\right)$, both on the same field K, is continuously differentiable in Ω and f' is the linear map $\overrightarrow{f} \in \mathcal{L}\left(\overrightarrow{E}; \overrightarrow{F}\right)$ associated to f.*

So if f=constant then f'=0

Theorem 1230 *(Schwartz II p.86) A continuous r multilinear map $f \in \mathcal{L}^r \left(\overrightarrow{E}_1, ... \overrightarrow{E}_r; \overrightarrow{F} \right)$*

defined on the normed vector space $\prod_{i=1}^{r} \overrightarrow{E}_r$ and valued in the normed vector space

$\overrightarrow{F}$ *,all on the same field K, is continuously differentiable and its derivative at*
$\overrightarrow{u} = (\overrightarrow{u}_1, ..., \overrightarrow{u}_r)$ *is :*
$$f'\left(\overrightarrow{u}\right)\left(\overrightarrow{v}_1, ..., \overrightarrow{v}_r\right) = \sum_{i=1}^{r} f\left(\overrightarrow{u}_1, ., \overrightarrow{u}_{i-1}, \overrightarrow{v}_i, \overrightarrow{u}_{i+1}.., \overrightarrow{u}_r\right)$$

Chain rule

Theorem 1231 *(Schwartz II p.93) Let $\left(E, \overrightarrow{E}\right), \left(F, \overrightarrow{F}\right), \left(G, \overrightarrow{G}\right)$ be affine normed spaces on the same field K, Ω an open subset of E,*

If the map $f : \Omega \to F$ is differentiable at $a \in E$, and the map : $g : F \to G$ is differentiable at $b=f(a)$, then the map $g \circ f : \Omega \to G$ is differentiable at a and :
$(g \circ f)'(a) = g'(b) \circ f'(a) \in \mathcal{L}\left(\overrightarrow{E}; \overrightarrow{G}\right)$

Let us write : $y = f(x), z = g(y)$. Then $g'(b)$ is the differential of g with respect to y, computed in b=f(a), and f'(a) is the differential of f with respect to x, computed in x=a.

If the spaces are finite dimensional then the jacobian of g∘f is the product of the jacobians.

Special case : let E an affine normed space and $f \in \mathcal{L}(E; E)$ continuously differentiable. Consider the iterate $F_n = (f)^n = (f \circ f \circ ...f) = F_{n-1} \circ f$. By recursion : $F_n'(a) = (f'(a))^n$ the n iterate of the linear map f'(a)

Derivatives on the spaces of linear maps
The definition of derivative holds for any normed vector spaces, in particular for spaces of linear maps.

1. Derivative of the compose of linear maps:

Theorem 1232 *If E is a normed vector space, then the set $\mathcal{L}(E; E)$ of continuous endomorphisms is a normed vector space and the composition : $M :$ $\mathcal{L}(E; E) \times \mathcal{L}(E; E) \to \mathcal{L}(E; E) :: M(f, g) = f \circ g$ is a bilinear, continuous map $M \in \mathcal{L}^2 (\mathcal{L}(E; E); \mathcal{L}(E; E))$ thus it is differentiable and the derivative of M at (f,g) is: $M'(f, g)(\delta f, \delta g) = \delta f \circ g + f \circ \delta g$*

This is the application of the previous theorem.

2. Derivative of the inverse of a linear map:

Theorem 1233 *(Schwartz II p.181) Let E,F be Banach vector spaces, U the subset of invertible elements of $\mathcal{L}$(E;F), U^{-1} the subset of invertible elements of $\mathcal{L}$(F;E), then :*

i) U,U^{-1} are open subsets

ii) the map $\Im : U \to U^{-1} :: \Im(f) = f^{-1}$ is a C_∞−diffeomorphism (bijective, continuously differentiable at any order as its inverse). Its derivative at f is :
$$\delta f \in U \subset \mathcal{L}(E;F) : (\Im(f))'(\delta f) = -f^{-1} \circ (\delta f) \circ f^{-1}$$

3. As a consequence:

Theorem 1234 *The set $G\mathcal{L}(E;E)$ of continuous automorphisms of a Banach vector space E is an open subset of $\mathcal{L}(E;E)$.*

i) the composition law : $M : \mathcal{L}(E;E) \times \mathcal{L}(E;E) \to \mathcal{L}(E;E) :: M(f,g) = f \circ g$ is differentiable and

$$M'(f,g)(\delta f, \delta g) = \delta f \circ g + f \circ \delta g$$

ii) the map : $\Im : G\mathcal{L}(E;E) \to G\mathcal{L}(E;E)$ is differentiable and $(\Im(f))'(\delta f) = -f^{-1} \circ \delta f \circ f^{-1}$

Diffeomorphism

Definition 1235 *A map $f : \Omega \to \Omega'$ between open subsets of the affine normed spaces on the same field K, is a **diffeomorphism** if f is bijective, continuously differentiable in Ω, and f^{-1} is continuously differentiable in Ω'.*

Definition 1236 *A map $f : \Omega \to \Omega'$ between open subsets of the affine normed spaces on the same field K, is a **local diffeomorphism** if for any $a \in \Omega$ there are a neighborhood n(a) of a and n(b) of b=f(a) such that f is a diffeomorphism from n(a) to n(b)*

A diffeomorphism is a homeomorphism, thus if E,F are finite dimensional we have necessarily dimE=dimF. Then the jacobian of f^{-1} is the inverse of the jacobian of f and $\det(f\prime(a)) \neq 0$.

Theorem 1237 *(Schwartz II p.96) If $f : \Omega \to \Omega'$ between open subsets of the affine normed spaces on the same field K, is a diffeomorphism then $\forall a \in \Omega, b = f(a) : (f'(a))^{-1} = (f^{-1})'(b)$*

Theorem 1238 *(Schwartz II p.190) If the map $f : \Omega \to F$ from the open subset Ω of the Banach affine space $\left(E, \overrightarrow{E}\right)$ to the Banach affine space $\left(F, \overrightarrow{F}\right)$ is continuously differentiable in Ω then :*

i) if for $a \in \Omega$ the derivative $f'(a)$ is invertible in $\mathcal{L}\left(\overrightarrow{E}; \overrightarrow{F}\right)$ then there A open in E, B open in F, $a \in A, b = f(a) \in B$, such that f is a diffeormorphism from A to B and $(f'(a))^{-1} = (f^{-1})'(b)$

ii) If for any $a \in \Omega$ $f'(a)$ is invertible in $\mathcal{L}\left(\overrightarrow{E}; \overrightarrow{F}\right)$ then f is an open map and a local diffeomorphism in Ω.

iii) If f is injective and for any $a \in \Omega$ $f'(a)$ is invertible then f is a diffeomorphism from Ω to $f(\Omega)$

Theorem 1239 *(Schwartz II p.192) If the map $f : \Omega \to F$ from the open subset Ω of the Banach affine space $\left(E, \overrightarrow{E} \right)$ to the normed affine space $\left(F, \overrightarrow{F} \right)$ is continuously differentiable in Ω and $\forall x \in \Omega$ f'(x) is invertible then f is a local homeomorphism on Ω. As a consequence f is an open map and $f(\Omega)$ is open.*

Immersion, submersion

Definition 1240 *A continuously differentiable map $f : \Omega \to F$ between an open subset of the affine normed space E to the affine normed space F, both on the same field K, is an **immersion** at $a \in \Omega$ if f'(a) is injective.*

Definition 1241 *A continuously differentiable map $f : \Omega \to F$ between an open subset of the affine normed space E to the affine normed space F, both on the same field K, is a **submersion** at $a \in \Omega$ if f'(a) is surjective.*

A submersion (resp.immersion) on Ω is a submersion (resp.immersion) at every point of Ω

Theorem 1242 *(Schwartz II p.193) If $f : \Omega \to F$ between an open subset of the affine Banach E to the affine Banach F is a submersion at $a \in \Omega$ then the image of a neighborhood of a is a neighborhood of f(a). If f is a submersion on Ω then it is an open map.*

Theorem 1243 *(Lang p.18) If the continuously differentiable map $f : \Omega \to F$ between an open subset of E to F, both Banach vector spaces on the same field K, is such that f'(p) is an isomorphism, continuous as its inverse, from E to a closed subspace F_1 of F and $F = F_1 \oplus F_2$, then there is a neighborhood n(p) such that $\pi_1 \circ f$ is a diffeomorphism from n(p) to an open subset of F_1, with π_1 the projection of F to F_1.*

Theorem 1244 *(Lang p.19) If the continuously differentiable map $f : \Omega \to F$ between an open subset of $E = E_1 \oplus E_2$ to F, both Banach vector spaces on the same field K, is such that the partial derivative $\partial_{x_1} f(p)$ is an isomorphism, continuous as its inverse, from E_1 to F, then there is a neighborhood n(p) where $f = f \circ \pi_1$ with π_1 the projection of E to E_1.*

Theorem 1245 *(Lang p.19) If the continuously differentiable map $f : \Omega \to F$ between an open subset of E to F, both Banach vector spaces on the same field K, is such that f'(p) is surjective and $E = E_1 \oplus \ker f'(p)$, then there is a neighborhood n(p) where $f = f \circ \pi_1$ with π_1 the projection of E to E_1.*

Rank of a map

The derivative is a linear map, so it has a rank $= \dim f'(a) \left(\overrightarrow{E} \right)$

Definition 1246 *The **rank** of a differentiable map is the rank of its derivative.*

$$\dim f\prime(a)\overrightarrow{E} \leq \min\left(\dim \overrightarrow{E}, \dim \overrightarrow{F}\right)$$

If E,F are finite dimensional the rank of f in a is the rank of the jacobian.

Theorem 1247 *Constant rank (Schwartz II p.196) Let f be a continuously differentiable map $f : \Omega \to F$ between an open subset of the affine normed space $\left(E, \overrightarrow{E}\right)$ to the affine normed space $\left(F, \overrightarrow{F}\right)$, both finite dimensional on the same field K. Then:*

i) If f has rank r at $a \in \Omega$, there is a neighborhood n(a) such that f has rank $\geq r$ in n(a)

ii) if f is an immersion or a submersion at $a \in \Omega$ then f has a constant rank in a neighborhood n(a)

iii) if f has constant rank r in Ω then there are a bases in $\overrightarrow{E}$ and $\overrightarrow{F}$ such that f can be expressed as :

$F(x_1, ..., x_m) = (x_1, .., x_r, 0, ...0)$

Derivative of a map defined by a sequence

Theorem 1248 *(Schwartz II p.122) If the sequence $(f_n)_{n \in \mathbb{N}}$ of differentiable (resp.continuously differentiable) maps : $f_n : \Omega \to F$ from an open subset Ω of the normed affine space E, to the normed affine space F, both on the same field K, converges to f and if for each $a \in \Omega$ there is a neighborhood where the sequence f'_n converges uniformly to g, then f is differentiable (resp.continuously differentiable) in Ω and f'=g*

We have also the slightly different theorem :

Theorem 1249 *(Schwartz II p.122) If the sequence $(f_n)_{n \in \mathbb{N}}$ of differentiable (resp.continuously differentiable) maps : $f_n : \Omega \to F$ from an open connected subset Ω of the normed affine space E, to the Banach affine space F, both on the same field K, converges to f(a) at least at a point $b \in \Omega$, and if for each $a \in \Omega$ there is a neighborhood where the sequence f'_n converges uniformly to g, then f_n converges locally uniformly to f in Ω, f is differentiable (resp.continuously differentiable) in Ω and f'=g*

Theorem 1250 *Logarithmic derivative (Schwartz II p.130) If the sequence $(f_n)_{n \in \mathbb{N}}$ of continuously differentiable maps : $f_n : \Omega \to \mathbb{C}$ on an open connected subset Ω of the normed affine space E are never null on Ω, and for each $a \in \Omega$ there is a neighborhood where the sequence $(f'_n(a) / f_n(a))$ converges uniformly to g, if there is $b \in \Omega$ such that a $(f_n(b))_{n \in \mathbb{N}}$ converges to a non zero limit, then $(f_n)_{n \in \mathbb{N}}$ converges to a function f which is continuously differentiable over Ω , never null and g=f'/f*

Remark : f'/f is called the **logarithmic derivative**

Derivative of a function defined by an integral

Theorem 1251 *(Schwartz IV p.107) Let E be an affine normed space, μ a Radon measure on a topological space T, $f \in C(E \times T; F)$ with F a banach vector space. If f(.,t) is x differentiable for almost every t, if for almost every a in T $\frac{\partial f}{\partial x}(a,t)$ is $\mu-$measurable and there is a neighborhood n(a) in E such that $\left\|\frac{\partial f}{\partial x}(x,t)\right\| \leq k(t)$ in n(a) where $k(t) \geq 0$ is integrable on T, then the map : $u(x) = \int_T f(x,t)\,\mu(t)$ is differentiable in E and its derivative is : $\frac{du}{dx}(a) = \int_T \frac{\partial f}{\partial x}(x,t)\,\mu(t)$. If f(.,t) is continuously x differentiable then u is continuously differentiable.*

Theorem 1252 *(Schwartz IV p.109) Let E be an affine normed space, μ a Radon measure on a topological space T, f a continuous map from ExT in a Banach vector space F. If f has a continuous partial derivative with respect to x, if for almost every a in T there is a compact neighborhood K(a) in E such that the support of $\frac{\partial f}{\partial x}(x,t)$ is in K(a) then the function : $u(x) = \int_T f(x,t)\,\mu(t)$ is continuously differentiable in E and its derivative is : $\frac{du}{dx}(a) = \int_T \frac{\partial f}{\partial x}(x,t)\,\mu(t)$.*

Gradient

If $f \in C_1(\Omega; K)$ then $f\prime(a) \in \overrightarrow{E}'$ the topological dual of $\overrightarrow{E}$. If E is finite dimensional and there is, either a bilinear symmetric or an hermitian form g, non degenerate on $\overrightarrow{E}$, then there is an isomorphism between $\overrightarrow{E}$ and $\overrightarrow{E}'$. To f'(a) we can associate a vector, called **gradient** and denoted $\text{grad}_a f$ such that : $\forall \overrightarrow{u} \in \overrightarrow{E} : f'(a)\overrightarrow{u} = g(grad_a f, \overrightarrow{u})$. If f is continuously differentiable then the map : $\text{grad}: \Omega \to \overrightarrow{E}$ defines a vector field on E.

14.2   Higher order derivatives

14.2.1   Definitions

Definition

If f is continuously differentiable, its derivative $f' : \Omega \to \mathcal{L}\left(\overrightarrow{E}; \overrightarrow{E}\right)$ can be differentiable and its derivative is f''=(f')'.

Theorem 1253 *(Schwartz II p.136) If the map $f : \Omega \to F$ from the open subset Ω of the normed affine space $\left(E, \overrightarrow{E}\right)$ to the normed affine space $\left(F, \overrightarrow{F}\right)$ is continuously differentiable in Ω and its derivative map f' is differentiable in $a \in \Omega$ then f''(a) is a **continuous symmetric bilinear** map in $\mathcal{L}^2(\overrightarrow{E}; \overrightarrow{F})$*

We have the map $f' : \Omega \to \mathcal{L}\left(\overrightarrow{E}; \overrightarrow{E}\right)$ and its derivative in a f''(a)=(f'(x))$|_{x=a}$ is a continuous linear map :f''(a): $\overrightarrow{E} \to \mathcal{L}\left(\overrightarrow{E}; \overrightarrow{F}\right)$. Such a map is equivalent to a continuous bilinear map in $\mathcal{L}^2(\overrightarrow{E}; \overrightarrow{F})$. Indeed : $\overrightarrow{u}, \overrightarrow{v} \in \overrightarrow{E} : f''(a)(\overrightarrow{u}) \in$

$\mathcal{L}\left(\overrightarrow{E}; \overrightarrow{F}\right) \Leftrightarrow (f"(a)(\overrightarrow{u}))(\overrightarrow{v}) = B(\overrightarrow{u}, \overrightarrow{v}) \in \overrightarrow{F}$. So we usually consider the map f'(a) as a bilinear map valued in $\overrightarrow{F}$. This bilinear map is symmetric : $f"(a)(\overrightarrow{u}, \overrightarrow{v}) = f"(a)(\overrightarrow{v}, \overrightarrow{u})$

This definition can be extended by recursion to the derivative of order r.

Definition 1254 *The map $f : \Omega \to F$ from the open subset Ω of the normed affine space $\left(E, \overrightarrow{E}\right)$ to the normed affine space $\left(F, \overrightarrow{F}\right)$ is **r continuousy differentiable** in Ω if it is continuously differentiable and its derivative map f' is r-1 differentiable in Ω . Then its r order derivative $f^{(r)}(a)$ in $a \in \Omega$ is a continuous symmetric r linear map in $\mathcal{L}^r(\overrightarrow{E}; \overrightarrow{F})$.*

If f is r-continuously differentiable, whatever r, it is said to be **smooth**

Notation 1255 *$C_r(\Omega; F)$ is the set of continuously r-differentiable maps $f : \Omega \to F$.*

$C_\infty(\Omega; F)$ is the set of smooth maps $f : \Omega \to F$

$f^{(r)}$ is the r order derivative of f : $f^{(r)} : \Omega \to \mathcal{L}_S^r(\overrightarrow{(E)^t}; \overrightarrow{F})$

$f"$ is the 2 order derivative of f : $f" : \Omega \to \mathcal{L}_S^2(\overrightarrow{(E)^t}; \overrightarrow{F})$

$f^{(r)}(a)$ is the value at a of the r order derivative of f : $f^{(r)}(a) \in \mathcal{L}_S^r(\overrightarrow{(E)^t}; \overrightarrow{F})$

Partial derivatives

Definition 1256 *A map $f : \Omega \to F$ defined on an open subset Ω of the normed affine space $\left(E, \overrightarrow{E}\right) = \prod_{i=1}^r \left(E_i, \overrightarrow{E}_i\right)$ and valued in the normed affine space $\left(F, \overrightarrow{F}\right)$, all on the same field K, has a **partial derivative of order 2** in Ω with respect to the variables $x_k = \pi_k(x), x_l = \pi_l(x)$ where $\pi_k : E \to E_k$ is the canonical projection, if f has a partial derivative with respect to the variable x_k in Ω and the map f'_{x_k} has a partial derivative with respect to the variable x_l.*

The partial derivatives must be understood as follows :

1. Let $E = E_1 \times E_2$ and $\Omega = \Omega_1 \times \Omega_2$. We consider the map $f : \Omega \to F$ as a two variables map $f(x_1, x_2)$.

For the first derivative we proceed as above. Let us fix $x_1 = a_1$ so we have a map : $f(a_1, x_2) : \Omega_2 \to F$ for $\Omega_2 = \{x_2 \in E_2 : (a_1, x_2) \in \Omega\}$. Its partial derivative with respect to x_2 at a_2 is the map $f'_{x_2}(a_1, a_2) \in \mathcal{L}\left(\overrightarrow{E}_2; \overrightarrow{F}\right)$

Now allow $x_1 = a_1$ to move in E_1 (but keep a_2 fixed). So we have now a map : $f'_{x_2}(x_1, a_2) : \Omega_1 \to \mathcal{L}\left(\overrightarrow{E}_2; \overrightarrow{F}\right)$ for $\Omega_1 = \{x_1 \in E_1 : (x_1, a_2) \in \Omega\}$. Its partial derivative with respect to x_1 is a map : $f"_{x_1 x_2}(a_1, a_2) : \overrightarrow{E}_1 \to \mathcal{L}\left(\overrightarrow{E}_2; \overrightarrow{F}\right)$ that we assimilate to a map $f"_{x_1 x_2}(a_1, a_2) \in \mathcal{L}^2\left(\overrightarrow{E}_1, \overrightarrow{E}_2; \overrightarrow{F}\right)$

If f is 2 times differentiable in (a_1, a_2) the result does not depend on the order for the derivations : $f"_{x_1 x_2}(a_1, a_2) = f"_{x_2 x_1}(a_1, a_2)$

We can proceed also to $f"_{x_1 x_1}(a_1, a_2), f"_{x_2 x_2}(a_1, a_2)$ so we have 3 distinct partial derivatives with respect to all the combinations of variables.

2. The partial derivatives are symmetric bilinear maps which act on different vector spaces:

$$f"_{x_1 x_2}(a_1, a_2) \in \mathcal{L}^2\left(\overrightarrow{E}_1, \overrightarrow{E}_2; \overrightarrow{F}\right)$$

$$f"_{x_1 x_1}(a_1, a_2) \in \mathcal{L}^2\left(\overrightarrow{E}_1, \overrightarrow{E}_1; \overrightarrow{F}\right)$$

$$f"_{x_2 x_2}(a_1, a_2) \in \mathcal{L}^2\left(\overrightarrow{E}_2, \overrightarrow{E}_2; \overrightarrow{F}\right)$$

A vector in $\overrightarrow{E} = \overrightarrow{E}_1 \times \overrightarrow{E}_2$ can be written as : $\overrightarrow{u} = (\overrightarrow{u}_1, \overrightarrow{u}_2)$

The action of the first derivative map $f'(a_1, a_2)$ is just : $f'(a_1, a_2)(\overrightarrow{u}_1, \overrightarrow{u}_2) = f'_{x_1}(a_1, a_2)\overrightarrow{u}_1 + f'_{x_2}(a_1, a_2)\overrightarrow{u}_2$

The action of the second derivative map $f"(a_1, a_2)$ is now on the two vectors $\overrightarrow{u} = (\overrightarrow{u}_1, \overrightarrow{u}_2), \overrightarrow{v} = (\overrightarrow{v}_1, \overrightarrow{v}_2)$

$f"(a_1, a_2)((\overrightarrow{u}_1, \overrightarrow{u}_2), (\overrightarrow{v}_1, \overrightarrow{v}_2))$
$= f"_{x_1 x_1}(a_1, a_2)(\overrightarrow{u}_1, \overrightarrow{v}_1) + f"_{x_1 x_2}(a_1, a_2)(\overrightarrow{u}_1, \overrightarrow{v}_2) + f"_{x_2 x_1}(a_1, a_2)(\overrightarrow{u}_2, \overrightarrow{v}_1) + f"_{x_2 x_2}(a_1, a_2)(\overrightarrow{u}_2, \overrightarrow{v}_2)$

Notation 1257 $f^{(r)}_{x_{i_1}...x_{i_r}} = \frac{\partial^r f}{\partial x_{i_1}...\partial x_{i_r}} = D_{i_1...i_r}$ *is the r order partial derivative map* : $\Omega \to \mathcal{L}^r\left(\overrightarrow{E}_{i_1}, ..\overrightarrow{E}_{i_r}; \overrightarrow{F}\right)$ *with respect to* $x_{i_1}, ...x_{i_r}$

$f^{(r)}_{x_{i_1}...x_{i_r}}(a) = \frac{\partial^r f}{\partial x_{i_1}...\partial x_{i_r}}(a) = D_{i_1...i_r}(a)$ *is the value of the partial derivative map at* $a \in \Omega$

Condition for r-differentiability

The theorem for differentiabillity is extended as follows:

Theorem 1258 *(Schwartz II p.142) A map* $f : \Omega \to F$ *defined on an open subset* Ω *of the normed affine space* $\left(E, \overrightarrow{E}\right) = \prod_{i=1}^{r}\left(E_i, \overrightarrow{E}_i\right)$ *and valued in the normed affine space* $\left(F, \overrightarrow{F}\right)$, *all on the same field K, is continuously r differentiable in* Ω *iff it has continuous partial derivatives of order r with respect to every combination of r variables.in* Ω.

Coordinates expression

If E is m dimensional and F n dimensional, the map $f^{(r)}_{x_{i_1}...x_{i_r}}(a)$ for $r > 1$ is no longer represented by a matrix. This is a r covariant and 1 contravariant tensor in $\otimes_r \overrightarrow{E}^* \otimes F$.

With a basis $\left(e^i\right)_{i=1}^{m}$ of $\overrightarrow{E}^*$ and $(f_j)_{i=1}^{n}$: $\sum_{i_1...i_r=1}^{m}\sum_{j=1}^{n} T^j_{i_1...i_r} e^{i_1} \otimes ... \otimes e^{i_r} \otimes f_j$ and $T^j_{i_1...i_r}$ is symmetric in all the lower indices.

14.2.2 Properties of higher derivatives

Polynomial

Theorem 1259 *If f is a continuous affine map $f : E \to F$ with associated linear map $\overrightarrow{f} \in \mathcal{L}\left(\overrightarrow{E}; \overrightarrow{F}\right)$ then f is smooth and $f'=\overrightarrow{f}, f^{(r)} = 0, r > 1$*

A polynomial P of degree p in n variables over the field K, defined in an open subset $\Omega \subset K^n$ is smooth. $f^{(r)} \equiv 0$ if p < r

Theorem 1260 *(Schwartz II p.164) A map $f : \Omega \to K$ from an open connected subset in K^n has a r order derivative $f^{(r)} \equiv 0$ in Ω iff it is a polynomial of order < r.*

Leibniz's formula

Theorem 1261 *(Schwartz II p.144) Let E, E_1, E_2, F be normed vector spaces, Ω an open subset of E, B a continuous bilinear map in $\mathcal{L}^2(E_1, E_2; F), U_1, U_2$ r-continuously differentiable maps $U_k : \Omega \to E_k$ maps, then the map : $B(U_1, U_2) : \Omega \to F :: B(U_1(x), U_2(x))$ is r-continuously differentiable in Ω .*

If E is n-dimensional,.with the notation above it reads :
$D_{i_1...i_r} B(U_1, U_2) = \sum_{J \sqsubseteq (i_1...i_r)} B\left(D_J U_1, D_{(i_1...i_r)\backslash J} U_2\right)$
the sum is extended to all combinations J of indices in I=$(i_1...i_r)$
This is a generalization of the rule for the product of real functions : $(fg)' = f'g + fg'$

Differential operator
(see Functional analysis for more)
1. Let E a n-dimensional normed affine space with open subset Ω, F a normed *vector* space, a differential operator of order m$\leq r$ is a map :
$P : C_r(\Omega; F) \to C_r(\Omega; F) :: P(f) = \sum_{I, \|I\| \leq m} a_I D_I f$
the sum is taken over any I set of m indices in (1,2,...n), the coefficients are scalar functions $a_I : \Omega \to K$
Example : laplacian : $P(f) = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}$
Linear operators are linear maps on the vector space $C_r(\Omega; F)$
2. If the coefficients are constant scalars differential operators can be composed: $(P \circ Q)(f) = P(Q(f))$ as long as the resulting maps are differentiable, and the composition is commutative : $(P \circ Q)(f) = (Q \circ P)(f)$

Taylor's formulas

Theorem 1262 *(Schwartz II p.155) If f is a r-1 continuously differentiable map : $f : \Omega \to \overrightarrow{F}$ from an open Ω of an affine space E to a normed vector space $\overrightarrow{F}$, both on the same field K, and has a derivative $f^{(r)}$ in $a \in \Omega$, then for $h \in \overrightarrow{E}$ such that the segment $[a, a+h] \subset \Omega$:*

    i). $f(a+h) = f(a) + \sum_{k=1}^{r-1} \frac{1}{k!} f^{(k)}(a) h^k + \frac{1}{r!} f^{(r)}(a+\theta h) h^r$ with $\theta \in [0,1]$

    ii) $f(a+h) = f(a) + \sum_{k=1}^{r} \frac{1}{k!} f^{(k)}(a) h^k + \frac{1}{r!} \varepsilon(h) \|h\|^r$ with $\varepsilon(h) \in \overrightarrow{F}, \varepsilon(h)_{h \to 0} \to 0$

    iii) If $\forall x \in]a, a+h[: \exists f^{(r)}(x), \|f^{(r)}(x)\| \le M$ then : $\left\| f(a+h) - \sum_{k=0}^{r-1} \frac{1}{k!} f^{(k)}(a) h^k \right\| \le M \frac{1}{r!} \|h\|^r$

    with the notation : $f^{(k)}(a) h^k = f^{(k)}(a)(h, ...h)$ k times

    If E is m dimensional, in a basis : $\sum_{k=0}^{r} \frac{1}{k!} f^{(k)}(a) h^k = \sum_{(\alpha_1...\alpha_m)} \frac{1}{\alpha_1!...\alpha_m!} \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} ... \left(\frac{\partial}{\partial x_m}\right)^{\alpha_m} f(a) h_1^{\alpha_1}..h_m^{\alpha_m}$
where the sum is extended to all combinations of integers such that $\sum_{k=1}^{m} \alpha_k \le r$

Chain rule

    The formula only when f,g are real functions $f, g : \Omega \sqsubseteq \mathbb{R} \to \mathbb{R}$ is :

    $(g \circ f)^{(r)}(a) = \sum_{I_r} \frac{r!}{i_1! i_2! ... i_r!} g^{(r)}(f(a)) (f'(a))^{i_1} ... \left(f^{(r)}(a)\right)^{i_r}$ where $I_r = (i_1, ..i_r) : i_1 + i_2 + .. + i_r = r$

    to be understood as : $f^{(p)} \in \mathcal{L}^p \left(\overrightarrow{\mathbb{R}}; \mathbb{R}\right)$

Convex functions

Theorem 1263 *(Berge p.209) A 2 times differentiable function $f : C \to \mathbb{R}$ on a convex subset of $\mathbb{R}^m$ is convex iff $\forall a \in C : f''(a)$ is a positive bilinear map.*

14.3 Extremum of a function

14.3.1 Definitions

E set, Ω subset of E, $f:\Omega \to \mathbb{R}$

    f has a **maximum** in $a \in \Omega$ if $\forall x \in \Omega : f(x) \le f(a)$

    f has a **minimum** in $a \in \Omega$ if $\forall x \in \Omega : f(x) \ge f(a)$

    f has an **extremum** in $a \in \Omega$ if it has either a maximum of a minimum in a

    The extremum is local if it is an extremum in a neighborhood of a. It is global if it an extremum in the whole of Ω

14.3.2 General theorems

Continuous functions

Theorem 1264 *A continuous real valued function $f:C \to \mathbb{R}$ on a compact subset C of a topological space E has both a global maximum and a global minimum.*

Proof. $f(\Omega)$ is compact in $\mathbb{R}$, thus bounded and closed, so it has both an upper bound and a lower bound, and on $\mathbb{R}$ this entails that is has a greatest lower bound and a least upper bound, which must belong to $f(\Omega)$ because it is closed. ∎

Remark: if f is continuous and C connected then f(C) is connected, thus is an interval |a,b| with a,b possibly infinite. But it is possible that a or b are not met by f.

Convex functions

There are many theorems about extrema of functions involving convexity properties. This is the basis of linear programming. See for example Berge for a comprehensive review of the problem.

Theorem 1265 *If $f : C \to \mathbb{R}$ is a stricly convex function defined on a convex subset of a real affine space E, then a maximum of f is an extreme point of C.*

Proof. C,f stricly convex : $\forall M, P \in C, t \in [0,1]$:$f(tM + (1-t)P) < tf(M) + (1-t)f(P)$

If a is not an extreme point of C : $\exists M, P \in C, t \in]0,1[$:$tM + (1-t)P = a \Rightarrow f(a) < tf(M) + (1-t)f(P)$

If f is a maximum : $\forall M, P : f(a) \geq f(M), f(a) \geq f(P)$

$t \in]0,1[$: $tf(a) \geq tf(M), (1-t)f(a) \geq (1-t)f(P) \Rightarrow f(a) \geq tf(M) + (1-t)f(P)$ ∎

This theorem shows that for many functions the extrema do not lie in the interior of the domain but at its border. So this limits seriously the interest of the following theorems, based upon differentiability, which assume that the domain is an open subset.

Another classic theorem (which has many versions) :

Theorem 1266 *Minimax (Berge p.220) If f is a continuous functions :$f : \Omega \times \Omega' \to \mathbb{R}$ where Ω, Ω' are convex compact subsets of $\mathbb{R}^p, \mathbb{R}^q$,and f is concave in x and convex in y, then :$\exists (a,b) \in \Omega \times \Omega' : f(a,b) = \max_{x \in \Omega} f(x,b) = \min_{y \in \Omega'} f(a,y)$*

14.3.3 Differentiable functions

Theorem 1267 *If a function f:$\Omega \to \mathbb{R}$,differentiable in the open subset Ω of a normed affine space, has a local extremum in $a \in \Omega$ then f'(a)=0.*

The proof is immediate with the Taylor's formula.

The converse is not true. It is common to say that f is **stationary** (or that a is a **critical point**) in a if f'(a)=0, but this does not entail that f has an extremum in a (but if f'(a)≠0 it is *not* an extremum). The condition open on Ω is mandatory.

With the Taylor's formula the result can be precised :

Theorem 1268 *If a function f:$\Omega \to \mathbb{R}$,r differentiable in the open subset Ω of a normed affine space, has a local extremum in a$\in \Omega$ and $f^{(p)}(a) = 0, 1 \leq p < s \leq r, f^{(s)}(a) \neq 0$, then :*

if a is a local maximum, then s is even and $\forall h \in \overrightarrow{E} : f^{(s)}(a) h^m \leq 0$

if a is a local minimum, then s is even and $\forall h \in \overrightarrow{E} : f^{(s)}(a) h^m \geq 0$

The condition is necessary, not sufficient. If s is odd then a cannot be an extremum.

14.3.4 Maximum under constraints

They are the problems, common in engneering, to find the extremum of a map belonging to some set defined through relations, which may be or not strict.

Extremum with strict constraints

Theorem 1269 *(Schwartz II p.285) Let Ω be an open subset of a real affine normed space, $f, L_1, L_2, ...L_m$ real differentiable functions in Ω, A the subset $A = \{x \in \Omega : L_k(x) = 0, k = 1...m\}$. If a$\in A$ is a local extremum of f in A and the maps $L'_k(a) \in \overrightarrow{E}'$ are linearly independant, then here is a unique family of scalars $(\lambda_k)_{k=1}^m$ such that : $f'(a) = \sum_{k=1}^m \lambda_k L'_k(a)$*
if Ω is a convex set and the map $f(x) + \sum_{k=1}^m \lambda_k L_k(x)$ are concave then the condition is sufficient.

The λ_k are the **Lagrange multipliers**. In physics they can be interpreted as forces, and in economics as prices.

Notice that E can be of infinite dimensional. This theorem can be restated as follows :

Theorem 1270 *Let Ω be an open subset Ω of a real affine normed space E, $f : \Omega \to \mathbb{R}$,$L : \Omega \to F$ real differentiable functions in Ω, F a m dimensional real vector space, A the set $A = \{x \in \Omega : L(x) = 0\}$. If a$\in A$ is a local extremum of f in A and if the map $L'(a)$ is surjective, then : $\exists \lambda \in F^*$ such that : $f'(a) = \lambda \circ L'(a)$*

Kuhn and Tucker theorem

Theorem 1271 *(Berge p.236) Let Ω be an open subset of $\mathbb{R}^n$, $f, L_1, L_2, ...L_m$ real differentiable functions in Ω, A the subset $A = \{x \in \Omega : L_k(x) \leq 0, k = 1...m\}$. If a$\in A$ is a local extremum of f in A and the maps $L'_k(a) \in \overrightarrow{E}'$ are linearly independant, then here is a family of scalars $(\lambda_k)_{k=1}^m$ such that :*

k=1...p: $L_k(a) = 0, \lambda_k L_k(a) = 0$
k=p+1...m: $L_k(a) \leq 0, \lambda_k \geq 0$
$f'(a) + \sum_{k=1}^m \lambda_k L'_k(a) = 0$

If f is linear and L are affine functions this is the linear programming problem
:

Problem : find a$\in \mathbb{R}^n$ extremum of $[C]^t [x]$ with $[A] [x] \leq [B], [A]$ mxn matrix, $[B]$ mx1 matrix, $[x] \geq 0$

An extremum point is necessarily on the border, and there are many computer programs for solving the problem (simplex method).

14.4   Implicit maps

One classical problem in mathematics is to solve the equation f(x,y)=0 : find x with respect to y. If there is a function g such that f(x,g(x))=0 then y=g(x) is called the implicit function defined by f(x,y)=0. The fixed point theorem in a Banach space is a key ingredient to resolve the problem. These theorems are the basis of many other results in Differential Geometry and funcitonal Analysis. One important feature of the theorems below is that they apply on infinite dimensional vector spaces (when they are Banach).

In a neighborhood of a solution

The first theorems apply when a specific solution of the equation f(a,b)=c is known.

Theorem 1272 *(Schwartz II p.176) Let E be a topological space, $\left(F, \overrightarrow{F}\right)$ an affine Banach space, $\left(G, \overrightarrow{G}\right)$ an affine normed space, Ω an open in ExF, f a continuous map $f : \Omega \to G, (a, b) \in E \times F, c \in G$ such that c=f(a,b)*

if $\forall (x, y) \in \Omega$ f has a partial derivative map $f\prime_y (x, y) \in \mathcal{L}\left(\overrightarrow{F}; \overrightarrow{G}\right)$ and $(x, y) \to f'_y (x, y)$ is continuous in Ω

if $Q = f'_y (a, b)$ is invertible in $\mathcal{L}\left(\overrightarrow{F}; \overrightarrow{G}\right)$

then there are neighborhoods $n(a) \subset E, n(b) \subset F$ of a,b such that for any $x \in n(a)$ there is a unique $y = g(x) \in n (b)$ such that f(x,y)=c and g is continuous in n(a).

Theorem 1273 *(Schwartz II p.180) Let $\left(E, \overrightarrow{E}\right), \left(F, \overrightarrow{F}\right), \left(G, \overrightarrow{G}\right)$ be affine normed spaces, Ω an open in ExF, f a continuous map f:$\Omega \to G, (a, b) \in E \times F, c \in G$ such that c=f(a,b), and the neighborhoods $n(a) \subset E, n(b) \subset F$ of a,b,*

if there is a map g:n(a)→n(b) continuous at a and such that $\forall x \in n(a)$: $f(x, g(x)) = c$

if f is differentiable at (a,b) and $f\prime_y (a, b)$ invertible

then g is differentiable at a, and its derivative is : $g' (a) = - \left(f'_y (a, b)\right)^{-1} \circ (f'_x (a, b))$

Implicit map theorem

Theorem 1274 *(Schwartz II p.185) Let* $\left(E,\overrightarrow{E}\right),\left(F,\overrightarrow{F}\right),\left(G,\overrightarrow{G}\right)$ *be affine normed spaces,* Ω *an open in ExF, f:* $\Omega \to G$ *a continuously differentiable map in* Ω,

i) If there are A open in E, B open in F such that $AxB \sqsubseteq \Omega$ *and* $g : A \to B$ *such that f(x,g(x))=c in A,*

if $\forall x \in A : f_y'(x, g(x))$ *is invertible in* $\mathcal{L}\left(\overrightarrow{F};\overrightarrow{G}\right)$ *then g is continuously differentiable in A*

if f is r-continuously differentiable then g is r-continuously differentiable

ii) If there are $(a, b) \in E \times F, c \in G$ *such that c=f(a,b), F is complete and* $f_y'(a, b)$ *is invertible in* $\mathcal{L}\left(\overrightarrow{F};\overrightarrow{G}\right)$, *then there are neighborhoods* $n(a) \subset A, n(b) \subset B$ *of a,b such that n(a)xn(b)* $\subset \Omega$ *and for any* $x \in n(a)$ *there is a unique* $y = g(x) \in n(b)$ *such that f(x,y)=c. g is continuously differentiable in n(a) and its derivative is :* $g'(x) = -\left(f_y'(x, y)\right)^{-1} \circ \left(f_x'(x, y)\right)$. *If f is r-continuously differentiable then g is r-continuously differentiable*

14.5 Holomorphic maps

In algebra we have imposed for any linear map $f \in L(E; F)$ that E and F shall be vector spaces over the *same field* K. Indeed this is the condition for the definition of linearity $f(ku) = kf(u)$ to be consistent. Everything that has been said previously (when K was not explicitly $\mathbb{R}$) stands for complex vector spaces. But differentiable maps over complex affine spaces have surprising properties.

14.5.1 Differentiability

Definitions

1. Let E,F be two *complex* normed affine spaces with underlying vector spaces $\overrightarrow{E}, \overrightarrow{F}$, Ω an open subset in E.

i) If f is differentiable in a$\in E$ then f is said to be **C-differentiable**, and f'(a) is a C-linear map $\in \mathcal{L}\left(\overrightarrow{E};\overrightarrow{F}\right)$ so :

$\forall \overrightarrow{u} \in \overrightarrow{E} : f'(a)i\overrightarrow{u} = if'(a)\overrightarrow{u}$

ii) If there is a R-linear map L:$\overrightarrow{E} \to \overrightarrow{F}$ such that :

$\exists r > 0, \forall h \in \overrightarrow{E}, \left\|\overrightarrow{h}\right\|_E < r : f(a + \overrightarrow{h}) - f(a) = L\overrightarrow{h} + \varepsilon(h)\left\|\overrightarrow{h}\right\|_F$ where $\varepsilon(h) \in \overrightarrow{F}$ is such that $\lim_{h\to 0} \varepsilon(h) = 0$

then f is said to be **R-differentiable** in a. So the only difference is that L is R-linear.

A R-linear map is such that f(u+v)=f(u)+f(v), f(kv)=kf(v) for any real scalar k

iii) If E is a real affine space, and F a complex affine space, one cannot (without additional structure on E such as complexification) speak of C-differentiability of a map $f : E \to F$ but it is still fully legitimate to speak of R-differentiability. This is a way to introduce derivatives for maps with real domain valued in a complex codomain.

2. A C-differentiable map is R-differentiable, but a R-differentiable map is C-differentiable iff $\forall \overrightarrow{u} \in \overrightarrow{E} : f'(a)(i\overrightarrow{u}) = if'(a)(\overrightarrow{u})$

3. Example : take a real structure on a complex vector space $\overrightarrow{E}$. This is an antilinear map $\sigma : \overrightarrow{E} \to \overrightarrow{E}$. Apply the criterium for differentiability : $\sigma\left(\overrightarrow{u} + \overrightarrow{h}\right) - \sigma\left(\overrightarrow{u}\right) = \sigma\left(\overrightarrow{h}\right)$ so the derivative σ' would be σ but this map is R-linear and not C-linear. It is the same for the maps : $\mathrm{Re} : \overrightarrow{E} \to \overrightarrow{E} :: \mathrm{Re}\,\overrightarrow{u} = \frac{1}{2}\left(\overrightarrow{u} + \sigma\left(\overrightarrow{u}\right)\right)$ and $\mathrm{Im} : \overrightarrow{E} \to \overrightarrow{E} :: \mathrm{Im}\,\overrightarrow{u} = \frac{1}{2i}\left(\overrightarrow{u} - \sigma\left(\overrightarrow{u}\right)\right)$. Thus it is not legitimate to use the chain rule to C-differentiate a map such that $f\left(\mathrm{Re}\,\overrightarrow{u}\right)$.

4. The extension to differentiable and continuously differentiable maps over an open subset are obvious.

Definition 1275 *A **holomorphic** map is a map $f : \Omega \to F$ .continuously differentiable in Ω, where Ω is an open subset of E, and E,F are complex normed affine spaces.*

Cauchy-Riemann equations

Theorem 1276 *Let f be a map : $f : \Omega \to F$, where Ω is an open subset of E, and $\left(E, \overrightarrow{E}\right), \left(F, \overrightarrow{F}\right)$ are complex normed affine spaces. For any real structure on E, f can be written as a map $\widetilde{f}(x,y)$ on the product $E_{\mathbb{R}} \times iE_{\mathbb{R}}$ of two real affine spaces. f is holomorphic iff $\widetilde{f}'_y = i\widetilde{f}'_x$.*

Proof. 1) Complex affine spaces can be considered as real affine spaces (see Affine spaces) by using a real structure on the underlying complex vector space. Then a point in E is identified by a couple of points in two real affine spaces. Indeed it sums up to distinguish the real and the imaginary part of the coordinates. The operation is always possible but the real structures are not unique. With real structures on E and F, f can be written as a map :

$f\left(\mathrm{Re}\,z + i\,\mathrm{Im}\,z\right) = P\left(\mathrm{Re}\,z, \mathrm{Im}\,z\right) + iQ\left(\mathrm{Re}\,z, \mathrm{Im}\,z\right)$

$\widetilde{f} : \Omega_{\mathbb{R}} \times i\Omega_{\mathbb{R}} \to F_{\mathbb{R}} \times iF_{\mathbb{R}} : \widetilde{f}(x,y) = P(x,y) + iQ(x,y)$

where $\Omega_{\mathbb{R}} \times i\Omega_{\mathbb{R}}$ is the embedding of Ω in $E_{\mathbb{R}} \times iE_{\mathbb{R}}$, P,Q are maps valued in $F_{\mathbb{R}}$

2) If f is holomorphic in Ω then at any point a$\in \Omega$ the derivative f'(a) is a linear map between two complex vector spaces endowed with real structures. So for any vector u$\in \overrightarrow{E}$ it can be written : f'(a)u$= \widetilde{P}_x(a)(\mathrm{Re}\,u) + \widetilde{P}_y(a)(\mathrm{Im}\,u) + i\left(\widetilde{Q}_x(a)(\mathrm{Re}\,u) + \widetilde{Q}_y(a)(\mathrm{Im}\,u)\right)$ where $\widetilde{P}_x(a), \widetilde{P}_y(a), \widetilde{Q}_x(a), \widetilde{Q}_y(a)$ are real linear maps between the real kernels $E_{\mathbb{R}}, F_{\mathbb{R}}$ which satifsfy the identities : $\widetilde{P}_y(a) = -\widetilde{Q}_x(a); \widetilde{Q}_y(a) = \widetilde{P}_x(a)$ (see Complex vector spaces in the Algebra part).

On the other hand f'(a)u reads :

$f'(a)u = \widetilde{f}(x_a, y_a)'(\mathrm{Re}\,u, \mathrm{Im}\,u) = \widetilde{f}'_x(x_a, y_a)\,\mathrm{Re}\,u + \widetilde{f}'_y(x_a, y_a)\,\mathrm{Im}\,u = P'_x(x_a, y_a)\,\mathrm{Re}\,u + P'_y(x_a, y_a)\,\mathrm{Im}\,u + i\left(Q'_x(x_a, y_a)\,\mathrm{Re}\,u + Q'_y(x_a, y_a)\,\mathrm{Im}\,u\right)$

$P'_y(x_a, y_a) = -Q'_x(x_a, y_a); Q'_y(x_a, y_a) = P'_x(x_a, y_a)$

Which reads : $\widetilde{f}'_x = P'_x + iQ'_x; \widetilde{f}'_y = P'_y + iQ'_y = -Q'_x + iP'_x = i\widetilde{f}'_x$

3) Conversely if there are partial derivatives P'_x, P'_y, Q'_x, Q'_y continuous on $\Omega_{\mathbb{R}} \times i\Omega_{\mathbb{R}}$ then the map (P, Q) is R-differentiable. It will be C-differentiable if $f'(a) i \overrightarrow{u} = if'(a) \overrightarrow{u}$ and that is just the Cauchy-Rieman equations. The result stands for a given real structure, but we have seen that there is always such a structure, thus if C-differentiable for a real structure it will be C-differentiable in any real structure. ∎

The equations $f'_y = if'_x$ are the **Cauchy-Rieman equations**.

Remarks :

i) The partial derivatives depend on the choice of a real structure σ. If one starts with a basis $(e_i)_{i \in I}$ the simplest way is to define $\sigma(e_j) = e_j, \sigma(ie_j) = -ie_j$ so $\overrightarrow{E}_{\mathbb{R}}$ is generated by $(e_i)_{i \in I}$ with real components. In a frame of reference $\left(O, (e_j, ie_j)_{j \in I}\right)$ the coordinates are expressed by two real set of scalars (x_j, y_j). Thus the Cauchy-Rieman equations reads ; $\frac{\partial f}{\partial y_j} = i\frac{\partial f}{\partial x_j}$. It is how they are usually written but we have proven that the equations hold for E infinite dimensional.

ii) We could have thought to use $f(z) = f(x + iy)$ and the chain rule but the maps : $z \to \operatorname{Re} z, z \to \operatorname{Im} z$ are not differentiable.

iii) If F=$\overrightarrow{F}$ Banach then the condition f has continuous R-partial derivatives can be replaced by $\|f\|^2$ locally integrable.

Differential

The notations are the same as above, E and F are assumed to be complex Banach affine spaces, endowed with real structures.

Take a fixed origin O' for a frame in F. f reads : $f(x + iy) = O' + \overrightarrow{P}(x, y) + i\overrightarrow{Q}(x, y)$ with $\left(\overrightarrow{P}, \overrightarrow{Q}\right) : E_{\mathbb{R}} \times iE_{\mathbb{R}} \to \overrightarrow{F}_{\mathbb{R}} \times i\overrightarrow{F}_{\mathbb{R}}$

1. As an affine Banach space, E is a manifold, and the open subset Ω is still a manifold, modelled on $\overrightarrow{E}$. A frame of reference $\left(O, (\overrightarrow{e}_i)_{i \in I}\right)$ of E gives a map on E, and a holonomic basis on the tangent space, which is $\overrightarrow{E}_{\mathbb{R}} \times \overrightarrow{E}_{\mathbb{R}}$, and a 1-form (dx, dy) which for any vector $(\overrightarrow{u}, \overrightarrow{v}) \in \overrightarrow{E}_{\mathbb{R}} \times \overrightarrow{E}_{\mathbb{R}}$ gives the componants in the basis : $(dx, dy)(\overrightarrow{u}, \overrightarrow{v}) = (u^j, v^k)_{j,k \in I}$.

2. $\left(\overrightarrow{P}, \overrightarrow{Q}\right)$ can be considered as a 0-forms defined on a manifold and valued in a fixed vector space. They are R-differentiable, so one can define the exterior derivatives :
$$\left(d\overrightarrow{P}, d\overrightarrow{Q}\right) = \left(\overrightarrow{P}'_x dx + \overrightarrow{P}'_y dy, \overrightarrow{Q}'_x dx + \overrightarrow{Q}'_y dy\right) \in \Lambda_1\left(\Omega'; \overrightarrow{F}_{\mathbb{R}} \times \overrightarrow{F}_{\mathbb{R}}\right)$$
and the 1-form valued in $\overrightarrow{F}$:
$$\varpi = d\overrightarrow{P} + id\overrightarrow{Q} \in \Lambda_1\left(\Omega'; \overrightarrow{F}\right)$$
$$\varpi = \left(\overrightarrow{P}'_x dx + \overrightarrow{P}'_y dy\right) + i\left(\overrightarrow{Q}'_x dx + \overrightarrow{Q}'_y dy\right) = \left(\overrightarrow{P}'_x + i\overrightarrow{Q}'_x\right) dx + \left(\overrightarrow{P}'_y + i\overrightarrow{Q}'_y\right) dy$$

3. f is holomorphic iff : $\overrightarrow{Q}'_x = -\overrightarrow{P}'_y; \overrightarrow{Q}'_y = \overrightarrow{P}'_x$ that is iff
$$\varpi = \left(\overrightarrow{P}'_x - i\overrightarrow{P}'_y\right) dx + \left(\overrightarrow{P}'_y + i\overrightarrow{P}'_x\right) dy = \left(\overrightarrow{P}'_x - i\overrightarrow{P}'_y\right) dx + i\left(-i\overrightarrow{P}'_y + \overrightarrow{P}'_x\right) dy =$$
$$\left(\overrightarrow{P}'_x - i\overrightarrow{P}'_y\right)(dx + idy)$$

4. From (dx,dy) one can define the 1-forms valued in $\overrightarrow{F}$:

$dz = dx + idy$, $d\overline{z} = dx - idy$

thus : $dx = \frac{1}{2}\left(dz + d\overline{z}\right), dy = \frac{1}{2i}\left(dz - d\overline{z}\right)$

ϖ then can be written :

$\varpi = \left(\overrightarrow{P}'_x + i\overrightarrow{Q}'_x\right)\frac{1}{2}(dz + d\overline{z}) + \left(\overrightarrow{P}'_y + i\overrightarrow{Q}'_y\right)\frac{1}{2i}(dz - d\overline{z}) = \left(\overrightarrow{P}'_x + \frac{1}{i}\overrightarrow{P}'_y + i\overrightarrow{Q}'_x + \overrightarrow{Q}'_y\right)\frac{1}{2}(dz) + \left(\overrightarrow{P}'_x + i\overrightarrow{Q}'_x - \frac{1}{i}\overrightarrow{P}'_y - \overrightarrow{Q}'_y\right)\frac{1}{2}(d\overline{z})$

It is customary to denote :

$\overrightarrow{P}'_z = \overrightarrow{P}'_x + \frac{1}{i}\overrightarrow{P}'_y; \overrightarrow{P}'_{\overline{z}} = \overrightarrow{P}'_x - \frac{1}{i}\overrightarrow{P}'_y;$

$\overrightarrow{Q}'_z = \overrightarrow{Q}'_x + \frac{1}{i}\overrightarrow{Q}'_y; \overrightarrow{Q}'_{\overline{z}} = \overrightarrow{Q}'_x - \frac{1}{i}\overrightarrow{Q}'_y;$

and $f'_z = \overrightarrow{P}'_z + i\overrightarrow{Q}'_z, f'_{\overline{z}} = \overrightarrow{P}'_{\overline{z}} + i\overrightarrow{Q}'_{\overline{z}}$

so : $\varpi = \frac{1}{2}\left(\left(\overrightarrow{P}'_z + i\overrightarrow{Q}'_z\right)(dz) + \left(\overrightarrow{P}'_{\overline{z}} + i\overrightarrow{Q}'_{\overline{z}}\right)(d\overline{z})\right) = \frac{1}{2}\left(f'_z(dz) + f'_{\overline{z}}(d\overline{z})\right)$

and f is holomorphic iff $\overrightarrow{P}'_{\overline{z}} + i\overrightarrow{Q}'_{\overline{z}} = 0$ that is interpreted as $f'_{\overline{z}} = 0$:

Theorem 1277 *A map is continuously C-differentiable iff it does not depend explicitly on the conjugates $\overline{z}^j$.*

If so the differential of f reads : df=f'(z)dz

5. If all this is legitimate it is clear that $dz, d\overline{z}$ are not differential or derivatives. As $\frac{\partial \widetilde{f}}{\partial z^j}, \frac{\partial \widetilde{f}}{\partial \overline{z}^j}$ they are ad hoc notations and cannot be deduced from the chain rule on f(z)=f(x+iy). My personal experience is that they are far less convenient than it seems. Anyway the important result is that a differential, that can be denoted df=f'(z)dz, can be defined if f is continuously C-differentiable.

Derivatives of higher order

Theorem 1278 *A holomorphic map f:$\Omega \rightarrow F$ from an open subset of an affine normed space to an affine normed space F has C-derivatives of any order.*

If f' exists then $f^{(r)}$ exists $\forall r$. So this is in stark contrast with maps in real affine spaces. The proof is done by differentiating the Cauchy-Rieman equations

Extremums

A non constant holomorphic map cannot have an extremum.

Theorem 1279 *(Schwartz III p.302, 307, 314) If f:$\Omega \rightarrow F$ is a holomorphic map on the open Ω of the normed affine space E, valued in the normed affine space F, then:*

i) $\|f\|$ has no strict local maximum in Ω

ii) If Ω is bounded in E, f continuous on the closure of Ω, then:

$\sup_{x\in\Omega}\|f(x)\| = \sup_{x\in\overset{\circ}{\Omega}}\|f(x)\| = \sup_{x\in\overline{\Omega}}\|f(x)\|$

iii) If E is finite dimensional $\|f\|$ has a maximum on $\partial\left(\overline{\Omega}\right)$.

iv) if Ω is connected and f $\exists a \in \Omega : f(a) = 0, \forall n : f^{(n)}(a) = 0$ then f=0 in

Ω

v) if Ω is connected and f is constant in an open in Ω then f is constant in Ω

If f is never zero take $1/\|f\|$ and we get the same result for a minimum.

One consequence is that any holomorphic map on a compact holomorphic manifold is constant (Schwartz III p.307)

Theorem 1280 *(Schwartz III p.275, 312) If f:$\Omega \to \mathbb{C}$ is a holomorphic function on the connected open Ω of the normed affine space E, then:*

i) if $\operatorname{Re} f$ is constant then f is constant

ii) if $|f|$ is constant then f is constant

iii) If there is a$\in \Omega$ local extremum of $\operatorname{Re} f$ or $\operatorname{Im} f$ then f is constant

iv) If there is a$\in \Omega$ local maximum of $|f|$ then f is constant

v) If there is a$\in \Omega$ local minimum of $|f|$ then f is constant or f(a)=0

Sequence of holomorphic maps

Theorem 1281 *Weierstrass (Schwartz III p.326) Let Ω be an open bounded in an affine normed space, F a Banach vector space, if the sequence $(f_n)_{n \in \mathbb{N}}$ of maps : $f_n : \Omega \to F$,holomorphic in Ω and continuous on the closure of $\overline{\Omega}$, converges uniformly on $\partial \Omega$ it converges uniformly on $\overline{\Omega}$. Its limit f is holomorphic in Ω and continuous on the closure of Ω , and the higher derivatives $f_n^{(r)}$ converges locally uniformly in Ω to $f^{(r)}$.*

Theorem 1282 *(Schwartz III p.327) Let Ω be an open in an affine normed space, F a Banach vector space, if the sequence $(f_n)_{n \in \mathbb{N}}$ of maps : $f_n : \Omega \to F$,holomorphic in Ω and continuous on the closure of Ω, converges locally uniformly in Ω, then it converges locally uniformly on Ω, its limit is holomorphic and the higher derivatives $f_n^{(r)}$ converges locally uniformly in Ω to $f^{(r)}$.*

14.5.2 Maps defined on $\mathbb{C}$

The most interesting results are met when f is defined in an open of $\mathbb{C}$. But for most of them cannot be extended to higher dimensions.

In this subsection : Ω is an open subset of $\mathbb{C}$ and F a Banach *vector* space (in the following we will drop the arrow but F is a vector space and not an affine space). And f is a map : $f : \Omega \to F$

Cauchy differentiation formula

Theorem 1283 *The map $f : \Omega \to F, \Omega$ from an open in $\mathbb{C}$ to a Banach vector space F, continuously R-differentiable, is holomorphic iff the 1-form $\lambda = f'(z)dz$ is closed : $d\lambda = 0$*

Proof. This is a direct consequence of the previous subsection. Here the real structure of E=$\mathbb{C}$ is obvious : take the "real axis" and the "imaginary axis" of the plane $\mathbb{R}^2$. $\mathbb{R}^2$ as $\mathbb{R}^n, \forall n$, is a manifold and the open subset Ω is itself a manifold (with canonical maps). We can define the differential $\lambda = f'(z)dx + if'(z)dy = f'(z)dz$ ∎

Theorem 1284 *Morera (Schwartz III p.282): Let Ω be an open in $\mathbb{C}$ and $f : \Omega \to F$ be a continuous map valued in the Banach vector space F. If for any smooth compact manifold X with boundary in Ω we have $\int_{\partial X} f(z)dz = 0$ then f is holomorphic*

Theorem 1285 *(Schwartz III p.281) Let Ω be a simply connected open subset in $\mathbb{C}$ and $f : \Omega \to F$ be a holomorphic map valued in the Banach vector space F. Then*
i) for any class 1 manifold with boundary X in Ω $\int_{\partial X} f(z)dz = 0$
ii) f has indefinite integrals which are holomorphic maps $\varphi \in H(\Omega; F)$: $\varphi'(z) = f(z)$ defined up to a constant

Theorem 1286 *(Schwartz III p.289,294) Let Ω be a simply connected open subset in $\mathbb{C}$ and $f : \Omega \to F$ be a holomorphic map valued in the Banach vector space F. Then for any class 1 manifold X with boundary in Ω*
i) if $a \notin X : \int_{\partial X} \frac{f(z)}{z-a} = 0$ and if $a \in \overset{\circ}{X} : \int_{\partial X} \frac{f(z)}{z-a}dz = 2i\pi f(a)$
ii) If X is compact and if $a \in \overset{\circ}{X} : f^{(n)}(a) = \frac{n!}{2i\pi} \int_{\partial X} \frac{f(z)}{(z-a)^{n+1}}dz$

The proofs are a direct consequence of the Stockes theorem applied to Ω.
So we have : $\int_a^b f(z)dz = \varphi(b) - \varphi(a)$ the integral being computed on any continuous curve from a to b in Ω
These theorems are the key to the computation of many definite integrals $\int_a^b f(z)dz$
i) f being holomorphic depends only on z, and the indefinite integral (or antiderivative) can be computed as in elementary analysis
ii) as we can choose any curve we can take γ such that f or the integral is obvious on some parts of the curve
iii) if f is real we can consider some extension of f which is holomorphic

Taylor's series

Theorem 1287 *(Schwartz III p.303) If the map $f : \Omega \to F, \Omega$ from an open in $\mathbb{C}$ to a Banach vector space F is holomorphic, then the series : $f(z) = f(a) + \sum_{n=1}^{\infty} \frac{(z-a)^n}{n!} f^{(n)}(a)$ is absolutely convergent in the largest disc B(a,R) centered in a and contained in Ω and convergent in any disc B(a,r), r<R.*

Algebra of holomorphic maps

Theorem 1288 *The set of holomorphic maps from an open in $\mathbb{C}$ to a Banach vector space F is a complex vector space.*

The set of holomorphic functions on an open in $\mathbb{C}$ is a complex agebra with pointwise multiplication.

Theorem 1289 *Any polynomial is holomorphic, the exponential is holomorphic,*

The complex logarithm is defined as the indefinite integral of $\int \frac{dz}{z}$. We have $\int_R \frac{dz}{z} = 2i\pi$ where R is any circle centered in 0. Thus complex logarithms are defined up to $2i\pi n$

Theorem 1290 *(Schwartz III p.298) If the function $f : \Omega \to \mathbb{C}$ is holomorphic on the simply connected open Ω is such that $\forall z \in \Omega : f(z) \neq 0$ then there is a holomorphic function $g\, f : \Omega \to \mathbb{C}$ such that $f = \exp g$*

Meromorphic maps

Theorem 1291 *If f is a non null holomorphic function $f : \Omega \to \mathbb{C}$ on an open subset of $\mathbb{C}$, then all zeros of f are isolated points.*

Definition 1292 *The map $f : \Omega \to F, \Omega$ from an open in $\mathbb{C}$ to a Banach vector space F is **meromorphic** if it is holomorphic except at a set of isolated points, which are called the **poles** of f. A point a is a pole of order $r{>}0$ if there is some constant C such that $f(z) \simeq C/(z-a)^r$ when $z \to a$. If a is a pole and there is no such r then a is an **essential pole**.*

if $F = \mathbb{C}$ then a meromorphic function can be written as the ratio u/v of two holomorphic functions:

Warning ! the poles must be isolated, thus $\sin\frac{1}{z}, \ln z, ..$ are not meromorphic

Theorem 1293 *(Schwartz III p.330) If f is a non null holomorphic function $f : \Omega \to \mathbb{C}$ on an open subset of $\mathbb{C}$: $\Omega = R_1 < |z - a| < R_2$ then there is a family of complex scalars $(c_n)_{n=-\infty}^{+\infty}$ such that : $f(z) = \sum_{n=-\infty}^{n=+\infty} c_n (z-a)^n$. The coefficients are uniquely defined by : $c_n = \frac{1}{2i\pi} \int_\gamma \frac{f(z)}{(z-a)^{n+1}} dz$ where $\gamma \subset \Omega$ is a loop which wraps only once around a.*

This formula is of interest if f is not holomorphic in a.

Theorem 1294 *Weierstrass (Schwartz III p.337) : If f: $\Omega \to \mathbb{C}$ is holomorphic in $\Omega = \{0 < |z - a| < R\}$ and a is an essential pole for f, then the image by f of any subset $\{0 < |z - a| < r < R\}$ is dense in $\mathbb{C}$.*

It means that f(z) can be arbitrarily close to any complex number.

14.5.3 Analytic maps

Harmonic maps are treated in the Functional Analysis - Laplacian part.

Definition 1295 *A map f:$\Omega \rightarrow F$ from an open of a normed affine space and valued in a normed affine space F, both on the field K, is **K-analytic** if it is K-differentiable at any order and $\forall a \in \Omega, \exists n\,(a) : \forall x \in n\,(a) : f(x) - f(a) = \sum_{n=1}^{\infty} \frac{1}{n!} f^{(n)}\,(a)\,(x-a)^n$*

Warning ! a K-analytic function is smooth (indefinitely K-differentiable) but the converse is not true in general.

Theorem 1296 *(Schwartz III p.307) For a K-analytic map f:$\Omega \rightarrow F$ from a connected open of a normed affine space and valued in a normed affine space F, both on the field K, the following are equivalent :*
  i) f is constant in Ω
  ii) $\exists a \in \Omega : \forall n \geq 1 : f^{(n)}\,(a) = 0$
  iii) f is constant in an open in Ω

Theorem 1297 *Liouville (Schwartz III p.322): For a K-analytic map f:$E \rightarrow F$ from a normed affine space and valued in a normed affine space F, both on the field K:*
  i) if f, Ref or Imf is bounded then f is constant
  ii) if $\exists a \in E, n \in \mathbb{N}, n > 0, C > 0 : \|f\,(x)\| \leq C\,\|x-a\|^n$ then f is a polynomial of order $\leq n$

Theorem 1298 *(Schwartz III p.305) A holomorphic map f:$\Omega \rightarrow F$ on an open of a normed affine space to a Banach vector space F is $\mathbb{C}-analytic$*

Theorem 1299 *(Schwartz III p.322) If $f \in C_{\infty}\,(\Omega; \mathbb{R}), \Omega$ open in $\mathbb{R}$, then the following are equivalent*
  i) f is R-analytic
  ii) there is a holomorphic (complex analytic) extension of f in $D \subset \mathbb{C}$ such that $\Omega \subset D$
  iii) for every compact set $C \subset \Omega$ there exists a constant M such that for every $a \in C$ and every $n \in \mathbb{N}$:
  $\left| f^{(n)}\,(a) \right| \leq M^{n+1} n!$

15 MANIFOLDS

15.1 Manifolds

A manifold can be viewed as a "surface" of elementary geometry, and it is customary to introduce manifolds as some kind of sets embedded in affine spaces. However useful it can be, it is a bit misleading (notably when we consider Lie groups) and it is good to start looking at the key ingredient of the theory, which is the concept of charts. Indeed charts are really what the name calls for : a way to locate a point through a set of figures. The beauty of the concept is that we do not need to explicitly give all the procedure for going to the place so designated by the coordinates : all we need to know is that it is possible (and indeed if we have the address of a building we can go there). Thus to be mathematically useful we add a condition : the procedures from coordinates to points must be consistent with each other. If we have two charts, giving different coordinates for the same point, there must be some way to pass from one set of coordinates to the other, and this time we deal with figures, that is mathematical objects which can be precisely dealt with. So this "interoperability" of charts becomes the major feature of manifolds, and enables us to forget most of the time the definition of the charts.

15.1.1 Definitions

Definition of a manifold

The most general definition of a manifold is the following (Maliavin and Lang) :

Definition 1300 *An **atlas** $A\big(E, (O_i, \varphi_i)_{i \in I}\big)$ of class r on a set M is comprised of:*
i) a Banach vector space E on a field K
ii) a cover $(O_i)_{i \in I}$ of M (each O_i is a subset of M and $\cup_{i \in I} O_i = M$)
*iii) a family $(\varphi_i)_{i \in I}$, called **charts,** of bijective maps : $\varphi_i : O_i \to U_i$ where U_i is an open subset of E*
*iii) $\forall i, j \in I, O_i \cap O_j \neq \varnothing : \varphi_i (O_i \cap O_j)$ is an open subset of E, and the map $\varphi_j \circ \varphi_i^{-1} : U_i \to U_j$, called **transition map,** is a r continuous differentiable diffeomorphism on the domain $U_i \cap U_j$*

Definition 1301 *A **manifold** modeled on a Banach E is a set endowed with an atlas $A\big(E, (O_i, \varphi_i)_{i \in I}\big)$ of class r and there is at least another atlas $A\,'\big(E, (O'_i, \varphi'_i)_{i \in I}\big)$ of class r such that the union $A \cup A'$ is still an atlas of class r.*

Comments

1. M is said to be modeled on E. If E' is another Banach such that there is a continuous bijective map between E and E', then this map is a smooth diffeomorphism and E' defines the same manifold structure. So it is simpler to

assume that E is always the same. If E is over the field K M is a K-manifold. We will specify real or complex manifold when necessary. If E is a Hilbert space M is a Hilbert manifold. There is also a concept of manifold modelled on Fréchet spaces (example in the infinite jet prolongation of a bundle). Not all Banach spaces are alike, and the properties of E are crucial for those of M.

2. The dimension of E is the **dimension** of the manifold (possibly infinite). If E is finite n dimensional it is always possible to take E=K^n.

3. The charts of the atlas A are said to be compatible with the charts of the other r-atlas A'. The maps $\varphi_j \circ \varphi_i^{-1}$ in E give the rule when the domains of two charts overlap. There can be a unique chart in an atlas, but if there are than one it is mandatory that the O_i are open subsets.

4. r is the **class of the manifold**, if r=∞ the manifold is said to be smooth, if r=0 (the transition maps are continuous only) M is said to be a topological manifold. If the transition charts are K-analytic the manifold is said to be K-analytic.

5. To a point p$\in M$ a chart associes a vector u=$\varphi_i(p)$ in E and, through a basis in E, a set of numbers $(x_j)_{j\in J}$ in K which are the **coordinates** of p in the chart. If the manifold is finite dimensional the canonical basis of K^n is used and the coordinates are given by :j=1...n: $[x_j] = [\varphi_i(p)]$ matrices nx1

6. The condition b) could be seen as useless, it is not. Indeed the key point in manifolds is the interoperability of charts thus, if an atlas is comprised of a unique chart the existence of other atlases, defining the same manifold structure, is necessary.

7. The property for two atlas to be compatible is an equivalence relation. A class of equivalence in this relation defines a structure of manifold on a set, and one can have different manifold structures on a given set. For $\mathbb{R}^n$ n$\neq$ 4 all the smooth structures are euivalent (diffeomorphic), but on $\mathbb{R}^4$ there are uncountably many non equivalent smooth manifold structures (exotic !).

8. From the definition it is clear that any open subset of a manifold is itself a manifold.

9. Notice that no assumption is done about a topological structure on M. This important point is addressed below.

15.1.2   Examples

1. Any Banach vector space, any Banach affine space, and any open subsets of these spaces have a canonical structure of smooth differential manifold (with an atlas comprised of a unique chart), and we will always refer to this structure when required.

2. A finite n dimensional subspace of a topological vector space or affine space has a manifold structure (homeomorphic to K^n).

3. An atlas for the sphere S_n, n dimensional manifold defined by $\sum_{i=1}^{n+1} x_i^2 = 1$ in $\mathbb{R}^{n+1}$ is the stereographic projection. Choose a south pole $a \in S_n$ and a north pole $-a \in S_n$. The atlas is comprised of 2 charts :

$O_1 = S_n \setminus \{a\}, \varphi_1(p) = \frac{p - \langle p,a \rangle a}{1 - \langle p,a \rangle}$

$O_2 = S_n \setminus \{-a\}, \varphi_2(p) = \frac{p - \langle p,a \rangle a}{1 + \langle p,a \rangle}$

with the scalar product : $\langle p, a \rangle = \sum_{i=1}^{n+1} p_i a_i$

6. For a manifold embedded in $\mathbb{R}^n$, passing from cartesian coordinates to curvilinear coordinates (such that polar, spheric, cylindrical coordinates) is just a change of chart on the same manifold (see in the section Tensor bundle below).

Grassmanian

Definition 1302 *The **Grassmanian** denoted Gr(E;r) of a n dimensional vector space E over a field K is the set of all r dimensional vector subspaces of E.*

Theorem 1303 *(Schwartz II p.236).The Grassmanian Gr(E;r) has a structure of smooth manifold of dimension r(n-r), isomorphic to Gr(E;n-r) and homeomorphic to the set of matrices M in K(n) such that : M^2=M ; M*=M ; Trace(M)=r*

The Grassmanian for r=1 is the projective space P(E) associated to E. It is a n-1 smooth manifold, which is compact if K=$\mathbb{R}$..

15.1.3 Topology

The key point is that a manifold structure is defined by an atlas, and this atlas defines a topology on M. Conversely if M has already a topology, and a manifold structure is added, then there are compatibility conditions, which are quite obvious but not always met.

The topology associated to an atlas

The principe is simple : as a minimum, all the charts should be continuous.

Theorem 1304 *An atlas $A\big(E, (O_i, \varphi_i)_{i \in I}\big)$ on a manifold M defines a topology on M for which the sets O_i are open in M and the charts are continuous. This topology is the same for all equivalent atlas.*

Proof. i) Take a base Ω of the topology on E, then the collection of sets $\big\{\varphi_i^{-1}(\varpi), \varpi \in \Omega, i \in I\big\}$ is a base of a topology on M. Each O_i is open and each φ_i is an homeomorphism between O_i and U_i.

ii) If we have two compatible atlas A=$\big(E, (O_i, \varphi_i)_{i \in I}\big)$,A'=$\Big(E, \big(O'_j, \varphi'_j\big)_{i \in I}\Big)$ then $A \cup A'$ is still a r-atlas. So at the intersections $\widetilde{O}_{ij} = O_i \cap O'_j$ we have :$\varphi_i\big(O_i \cap O'_j\big), \varphi'_j\big(O_i \cap O'_j\big)$ are open subsets in E, and the transition maps : $\varphi'_j \circ \varphi_i^{-1} : U_i \to U'_j$ are r continuous differentiable diffeomorphism on the domain $U_i \cap U'_j$

Consider the topology defined by A. With this topology an open is the union of sets such $\varphi_i^{-1}(\varpi), \varpi \in \Omega$. :It suffices to prove that $\varphi_i^{-1}(\varpi)$ is open in the topology defined by A'.

$\forall \varpi \in \Omega : \varphi_i^{-1}(\varpi) = \varphi_i^{-1}(\varpi) \cap O_i = \varphi_i^{-1}(\varpi) \cap O_i \cap M = \cup_{j \in J} \big(\varphi_i^{-1}(\varpi) \cap O_i \cap O'_j\big)$

$$\varphi'_j\left(\varphi_i^{-1}\left(\varpi\right)\cap O_i\cap O'_j\right)=\varphi'_j\circ\varphi_i^{-1}\left(\varpi\cap U_i\cap U'_j\right)$$

$\varpi\cap U_i\cap U'_j$ is open in E, and because $\varphi'_j\circ\varphi_i^{-1}$ is a homeomorphism, $\varphi'_j\circ\varphi_i^{-1}\left(\varpi\cap U_i\cap U'_j\right)$ is open in E, and $\varphi_i^{-1}\left(\varpi\right)\cap O_i\cap O'_j=\varphi'^{-1}_j\left(\varphi_i^{-1}\left(\varpi\right)\cap O_i\cap O'_j\right)$ is an open for A'. ■

Conversely if M is endowed with a given topological structure, and then with a manifold structure, how do the two topologies coincide ?

Theorem 1305 *(Malliavin p.20) The topology induced by a manifold structure through an atlas $A\left(E,(O_i,\varphi_i)_{i\in I}\right)$ on a topological space (M,Ω) coincides with this latter topology iff $\forall i\in I, O_i\in\Omega$ and φ_i is an homeomorphism on O_i.*

A manifold modelled on E is locally homeomorphic to E

Theorem 1306 *If a manifold M is modelled on E, then every point of M has a neighborhood which is homeomorphic to an open of E. Conversely a topological space M such that every point of M has a neighborhood which is homeomorphic to an open of E can be endowed with a structure of manifold modelled on E.*

Proof. i) Take an atlas $A=\left(E,(O_i,\varphi_i)_{i\in I}\right)$. Let $p\in M$ so $\exists i\in I:p\in O_i$. with the topology defined by A, O_i is a neighborhood of p, which is homeomorphic to U_i, which is a neighborhood of $\varphi_i\left(p\right)\in E$.

ii) Conversely let (M,Ω) be a topological space, such that for each $p\in M$ there is a neighborhood n(p), an open subset $\mu\left(p\right)$ of E, and a homeomorphism φ_p between n(p) and μ. The family $\left(n\left(p\right),\varphi_p\right)_{p\in M}$ is an atlas for M :

$\forall p\in M:\varphi_p\left(n\left(p\right)\right)=\mu\left(p\right)$ is open in E

$\varphi_p\left(n\left(p\right)\cap n\left(q\right)\right)=\mu\left(p\right)\cap\mu\left(q\right)$ is an open in E (possibly empty).

$\varphi_p\circ\varphi_q^{-1}$ is the compose of two homeomorphisms, so a homeomorphism ■

Warning ! usually there is no global homeomorphism between M and E

Topological properties of manifolds

To sum up :

- if M has no prior topology, it gets one, uniquely defined by a class of atlas, and it is locally homeomorphic to E.

- if M is a topological space, its topology defines the conditions which must be met by an atlas so that it can define a manifold structure compatible with M. So a set, endowed with a given topology, may not accept some structures of manifolds (this is the case with structures involving scalar products).

1. Locally compact manifold:

Theorem 1307 *A manifold is locally compact iff it is finite dimensional. It is then a Baire space.*

Proof. i) If a manifold M modelled on a Banach E is locally compact, then E is locally compact, and is necessarily finite dimensional, and so is M.

ii) If E is finite dimensional, it is locally compact. Take p in M, and a chart $\varphi_i\left(p\right) = x \in E$. x has a compact neighborhood n(x), its image by the continous map φ_i^{-1} is a compact neighborhood of p. ∎

It implies that a compact manifold is *never* infinite dimensional.

2. Paracompactness, metrizability

Theorem 1308 *A second countable, regular manifold is metrizable.*

Proof. It is semi-metrizable, and metrizable if it is T1, but any manifold is T1 ∎

Theorem 1309 *A regular, Hausdorff manifold with a σ-locally finite base is metrizable*

Theorem 1310 *A metrizable manifold is paracompact.*

Theorem 1311 *For a finite dimensional manifold M the following properties are equivalent:*
i) M is paracompact
ii) M is metrizable
iii) M admits an inner product on its vector bundle

Proof. The final item of the proof is the following theorem :

(Kobayashi 1 p.116, 166) The vector bundle of a finite dimensional paracompact manifold M can be endowed with an inner product (a definite positive, either symmetric bilinear or hermitian sequilinear form) and M is metrizable. ∎

It implies that :

Theorem 1312 *For a finite dimensional, paracompact manifold M it is always possible to choose an atlas $A\left(E, (O_i, \varphi_i)_{i \in I}\right)$ such that the cover is relatively compact ($\overline{O}_i$ is compact) and locally finite (each points of M meets only a finite number of O_i)*

If M is a Hausdorf m dimensional class 1 real manifold then we can also have that any non empty finite intersection of O_i is diffeomorphic with an open of $\mathbb{R}^m$ (Kobayashi p.167).

Theorem 1313 *A finite dimensional, Hausdorff, second countable manifold is paracompact, metrizable and can be endowed with an inner product.*

Proof. The final item of the proof is the following theorem :

(Lang p.35) For every open covering $(\Omega_j)_{j \in J}$ of a locally compact, Hausdorff, second countable manifold M modelled on a Banach E, there is an atlas $(O_i, \varphi_i)_{i \in I}$ of M such that $(O_i)_{i \in I}$ is a locally finite refinement of $(\Omega_j)_{j \in J}$, $\varphi_i(O_i)$ is an open ball $B(x_i, 3) \subset E$ and the open sets $\varphi_i^{-1}(B(x_i, 1))$ covers M. ∎

3. Countable base:

Theorem 1314 *A metrizable manifold is first countable.*

Theorem 1315 *For a semi-metrizable manifold separable is equivalent to second countable.*

Theorem 1316 *A semi-metrizable manifold has a $\sigma-$locally finite base.*

Theorem 1317 *A connected, finite dimensional, metrizable manifold is separable and second countable.*

Proof. It is locally compact so the result follows the Kobayashi p.269 theorem (see General topology) ∎

4. Separability:

Theorem 1318 *A manifold is a T1 space*

Proof. A Banach is a T1 space, so each point is a closed subset, and its preimage by a chart is closed ∎

Theorem 1319 *A metrizable manifold is a Hausdorff, normal, regular topological space*

Theorem 1320 *A semi-metrizable manifold is normal and regular.*

Theorem 1321 *A paracompact manifold is normal*

Theorem 1322 *A finite dimensional manifold is regular.*

Proof. because it is locally compact ∎

5. To sum up:

Theorem 1323 *(Kobayashi 1 p.271) For a finite dimensional, connected, Hausdorff manifold M the following are equivalent :*
i) M is paracompact
ii) M is metrizable
iii) M admits an inner product
iv) M is second countable

6. A finite dimensional class 1 manifold has an equivalent smooth structure (Kolar p.4) thus one can usually assume that a finite dimensional manifold is smooth

Infinite dimensional manifolds

Infinite dimensional manifolds which are not too exotic have a simple structure : they are open subsets of Hilbert spaces.

Theorem 1324 *(Henderson) A separable metric manifold modelled on a separable infinite dimensional Fréchet space can be embedded as an open subset of an infinite dimensional, separable Hilbert space defined uniquely up to linear isomorphism.*

Of course the theorem applies to a manifold modeled on Banach space E, which is a Fréchet space. E is separable iff it is second countable, because this is a metric space. Then M is second countable if it has an atlas with a finite number of charts. If so it is also separable. It is metrizable if it is regular (because it is T1). Then it is necessarily Hausdorff.

Theorem 1325 *A regular manifold modeled on a second countable infinite dimensional Banach vector space, with an atlas comprised of a finite number of charts, can be embedded as an open subset of an infinite dimensional, separable Hilbert space, defined uniquely up to linear isomorphism.*

15.2 Differentiable maps

Manifolds are the only structures, other than affine spaces, upon which differentiable maps are defined.

15.2.1 Definitions

Definition 1326 *A map $f : M \to N$ between the manifolds M,N is said to be continuously differerentiable at the order r if, for any point p in M, there are charts (O_i, φ_i) in M, and (Q_j, ψ_j) in N, such that $p \in O_i, f(p) \in Q_j$ and that $\psi_j \circ f \circ \varphi_i^{-1}$ is r continuously differentiable in $\varphi_i \left(O_i \cap f^{-1}(Q_j) \right).$*

If so then $\psi_k \circ f \circ \varphi_l^{-1}$ is r continuously differentiable with any other charts meeting the same conditions.
Obviously r is less or equal to the class of both M and N. If the manifolds are smooth and f is of class r for any r then f is said to be smooth. In the following we will assume that the classes of the manifolds and of the maps match together.

Definition 1327 *A **r-diffeomorphism** between two manifolds is a bijective map, r-differentiable and with a r-differentiable inverse.*

Definition 1328 *A **local diffeomorphism** between two manifolds M,N is a map such that for each p∈ M there is a open subsets n(p) and n(f(p)) such that the restriction $\widehat{f} : n(p) \to f(n(p))$ is a diffeomorphism*

If there is a diffeomorphism between two manifolds they are said to be **diffeomorphic**. They have necessarily same dimension (possibly infinite).

The maps of charts (O_i, φ_i) of a class r manifold are r-diffeomorphism : $\varphi_i \in C_r(O_i; \varphi_i(O_i))$. Indeed whenever $O_i \cap O_j \neq \varnothing$ $\varphi_j \circ \varphi_i^{-1}$ is r continuously differentiable. And we have the same result with any other atlas.

If a manifold is an open of an affine space then its maps are smooth.

Let $A(E, (O_i, \varphi_i))$ be an atlas of M, and $B(G, (Q_j, \psi_j))$ of N. To any map $f : M \rightarrow N$ is associated maps between coordinates : if $x = \varphi_i(p)$ then $y = \psi_j(f(p))$ they read :

$F : O_i \rightarrow Q_j :: y = F(x)$ with $F = \psi_j \circ f \circ \varphi_i^{-1}$

$$
\begin{array}{ccccc}
\text{M} & & f & & \text{N} \\
O_i & \rightarrow & \rightarrow & \rightarrow & Q_i \\
\downarrow & & & & \downarrow \\
\downarrow & \varphi_i & & & \downarrow \quad \psi_j \\
\downarrow & & \text{F} & & \downarrow \\
\widehat{O}_i & \rightarrow & \rightarrow & \rightarrow & \widehat{Q}_j
\end{array}
$$

Then $F'(a) = \left(\psi_j \circ f \circ \varphi_i^{-1}\right)'(a)$ is a continuous linear map $\in \mathcal{L}(E; G)$.

If f is a diffeomorphism $F = \psi_j \circ f \circ \varphi_i^{-1}$ is a diffeomorphism between Banach vector spaces, thus :

i) $F'(a)$ is inversible and $\left(F^{-1}(b)\right)' = (F'(a))^{-1} \in \mathcal{L}(G; E)$

ii) F is an open map (it maps open subsets to open subsets)

Definition 1329 *The **jacobian** of a differentiable map between two finite dimensional manifolds is the matrix* $F'(a) = \left(\psi_j \circ f \circ \varphi_i^{-1}(a)\right)'$

If M is m dimensional defined over K^m, N is n dimensional defined over K^n, then $F(a) = \psi_j \circ f \circ \varphi_i^{-1}(a)$ can be written in the canonical bases of K^m, K^n : j=1...n : $y^j = F^j(x^1, ...x^m)$ using tensorial notation for the indexes

And $F'(a) = \left(\psi_j \circ f \circ \varphi_i^{-1}(a)\right)'$ is expressed in bases as a nxp matrix (over K)

$$
J = [F'(a)] = \left\{ \overbrace{\left[\frac{\partial F^\alpha}{\partial x^\beta}\right]}^{m} \right\} n
$$

If f is a diffeomorphism the jacobian of F^{-1} is the inverse of the jacobian of F.

15.2.2   General properties

Set of r differentiable maps

Notation 1330 $C_r(M;N)$ *is the set of class r maps from the manifold M to the manifold N (both on the same field)*

Theorem 1331 *The set $C_r(M;F)$ of r differentiable map from a manifold M to a Banach vector space F, both on the same field K, is a vector space. The set $C_r(M;K)$ is a vector space and an algebra with pointwise multiplication.*

Categories of differentiable maps

Theorem 1332 *(Schwartz II p.224) If $f \in C_r(M;N), g \in C_r(N;P)$ then $g \circ f \in C_r(M.P)$ (if the manifolds have the required class)*

Theorem 1333 *The class r K-manifolds and the class r differentiable maps constitute a category. The smooth K-manifolds and the smooth differentiable maps constitute a subcategory.*

There is more than the obvious : functors will transform manifolds into fiber bundles.

Product of manifolds
The product MxN of two class r manifolds on the same field K is a manifold with dimension = dim(M)+dim(N) and the projections $\pi_M : M \times N \to M$, $\pi_N : M \times N \to N$ are of class r.

For any class r maps : $f : P \to M, g : P \to N$ the mapping : $(f,g) :$ $P \to M \times N :: (f,g)(p) = (f(p), g(p))$ is the unique class r mapping with the property : $\pi_M((f,g)(p)) = f(p), \pi_N((f,g)(p)) = g(p)$

Space $\mathcal{L}$(E;E) for a Banach vector space

Theorem 1334 *The set $\mathcal{L}(E;E)$ of continuous linear map over a Banach vector space E is a Banach vector space, so this is a manifold. The subset $G\mathcal{L}(E;E)$ of inversible map is an open subset of $\mathcal{L}$(E;E), so this is also a manifold.*
The composition law and the inverse are differentiable maps :

i) the composition law : $M : \mathcal{L}(E;E) \times \mathcal{L}(E;E) \to \mathcal{L}(E;E) :: M(f,g) = f \circ g$ is differentiable and
$M'(f,g)(\delta f, \delta g) = \delta f \circ g + f \circ \delta g$
ii) the map : $\Im : G\mathcal{L}(E;E) \to G\mathcal{L}(E;E)$ is differentiable and $(\Im(f))'(\delta f) = -f^{-1} \circ \delta f \circ f^{-1}$

15.2.3 Partition of unity

Partition of unity is a powerful tool to extend local properties to local ones. They exist for paracompact Hausdorff spaces, and so for any Hausdorff finite dimensional manifold. However we will need maps which are are not only continuous but also differentiable. Furthermore difficulties arise with infinite dimensional manifolds.

Definition

Definition 1335 *A **partition of unity** of class r subordinate to an open covering $(\Omega_i)_{i \in I}$ of a manifold M is a family $(f_i)_{i \in I}$ of maps $f_i \in C_r(M; \mathbb{R}_+)$, such that the support of f_i is contained in Ω_i and :*

$\forall p \in M : f_i(p) \neq 0$ *for at most finite many i*
$\forall p \in M : \sum_{i \in I} f_i(p) = 1$

As a consequence the family $(Supp(f_i))_{i \in I}$ of the supports of the functions is locally finite

If the support of each function is compact then the partition is compactly supported

A manifold is said to admit partitions of unity if it has a partition of unity subordinate to any open cover.

Conditions for the existence of a partition of unity   From the theorems of general topology:

Theorem 1336 *A paracompact Hausdorff manifold admits continuous partitions of unity*

Theorem 1337 *(Lang p.37, Bourbaki) For any paracompact Hausdorff manifold and locally finite open cover $(\Omega_i)_{i \in I}$ of M there is a localy finite open cover $(U_i)_{i \in I}$ such that $\overline{U}_i \subset \Omega_i$*

Theorem 1338 *(Kobayashi I p.272) For any paracompact, finite dimensional manifold, and locally finite open cover $(\Omega_i)_{i \in I}$ of M sch that each Ω_i has compact closure, there is a partition of unity subodinate to $(\Omega_i)_{i \in I}$.*

Theorem 1339 *(Lang p.38) A class r paracompact manifold modeled on a separable Hilbert space admits class r partitions of unity subordinate to any locally finite open covering.*

Theorem 1340 *(Schwartz II p.242) For any class r finite dimensional second countable real manifold M , open cover $(\Omega_i)_{i \in I}$ of M there is a family $(f_i)_{i \in I}$ of functions $f_i \in C_r(M; \mathbb{R}_+)$ with support in Ω_i , such that :$\forall p \in K : \sum_{i \in I} f_i(p) = 1$, and $\forall p \in M$ there is a neighborhood n(p) on which only a finite number of f_i are not null.*

Theorem 1341 *(Schwartz II p.240) For any class r finite dimensional real manifold M, Ω open in M, $p \in \Omega$, there is a r continuously differentiable real function f with compact support included in Ω such that : f(p) > 0 and $\forall m \in \Omega : 0 \leq f(m) \leq 1$*

Theorem 1342 *(Schwartz II p.242) For any class r finite dimensional real manifold M, open cover $(\Omega_i)_{i \in I}$ of M, compact K in M, there is a family $(f_i)_{i \in I}$ of functions $f_i \in C_r(M; \mathbb{R}_+)$ with compact support in Ω_i such that : $\forall p \in M : f_i(p) \neq 0$ for at most finite many i and $\forall p \in K : \sum_{i \in I} f_i(p) > 0$*

Prolongation of a map

Theorem 1343 *(Schwartz II p.243) Let M be a class r finite dimensional second countable real manifold, C a closed subset of M, F real Banach vector space, then a map $f \in C_r (C; F)$ can be extended to a map $\widehat{f} \in C_r (M; F) : \forall p \in C :$ $\widehat{f}(p) = f(p)$*

Remark : the definition of a class r map on a closed set is understood in the Whitney sense : there is a class r map g on M such that the derivatives for any order $s \leq r$ of g are equal on C to the approximates of f by the Taylor's expansion.

15.3   The tangent bundle

15.3.1   Tangent vector space

Definition

Theorem 1344 *A each point p on a class 1 differentiable manifold M modelled on a Banach E on the field K, there is a set, called the **tangent space** to M at p, which has the structure of a vector space over K, isomorphic to E*

There are several ways to construct the tangent vector space. The simplest is the following:

Proof. i) two differentiable functions f,g$\in C_1 (M; K)$ are said to be equivalent if for a chart φ covering p, for every differentiable path : $c : K \to E$ such that $\varphi^{-1} \circ c(0) = p$ we have : $\left(f \circ \varphi^{-1} \circ c\right)' |_{t=0} = \left(g \circ \varphi^{-1} \circ c\right)' |_{t=0}$. The derivative is well defined because this is a map : $K \to K$. This is an equivalence relation $\sim$. Two maps equivalent with a chart are still equivalent with a chart of a compatible atlas.

ii) The value of the derivative $\left(f \circ \varphi^{-1}\right)' |_x$ for $\varphi(p) = x$ is a continuous map from E to K, so this is a form in E'.

The set of these values $\widetilde{T}^* (p)$ is a vector space over K, because $C_1 (M; K)$ is a vector space over K. If we take the sets $T^* (p)$ of these values for each class of equivalence we still have a vector space.

iii) $T^* (p)$ is isomorphic to E'.

The map : $T^* (p) \to E'$ is injective :

If $\left(f \circ \varphi^{-1}\right)' |_x = \left(g \circ \varphi^{-1}\right)' |_x$ then $\left(f \circ \varphi^{-1} \circ c\right)' |_{t=0} = \left(f \circ \varphi^{-1}\right)' |_x \circ c' |_{t=0} = \left(g \circ \varphi^{-1}\right)' |_x \circ c' |_{t=0} \Rightarrow f \sim g$

It is surjective : for any $\lambda \in E'$ take $f(q) = \lambda(\varphi(y))$

iv) The tangent space is the topological dual of $T^* (p)$. If E is reflexive then (E')' is isomorphic to E. ∎

Remarks :

i) It is crucial to notice that the tangent spaces at two different points have no relation with each other. To set up a relation we need special tools, such as connections.

ii) T*(p) is the 1-jet of the functions on M at p.

iii) the demonstration fails if E is not reflexive, but there are ways around this issue by taking the weak dual.

Notation 1345 T_pM *is the tangent vector space at p to the manifold M*

Properties of the tangent space

Theorem 1346 *The charts $\varphi_i'(p)$ of an atlas $\left(E, (O_i, \varphi_i)_{i \in I}\right)$ are continuous inversible linear map $\varphi_i'(p) \in G\mathcal{L}(T_pM; E)$*

A Banach vector space E has a smooth manifold structure. The tangent space at any point is just E itself.

A Banach affine space $\left(E, \overrightarrow{E}\right)$ has a smooth manifold structure. The tangent space at any point p is the affine subspace $\left(p, \overrightarrow{E}\right)$.

Definition 1347 *A **holonomic basis** of a manifold M with atlas $\left(E, (O_i, \varphi_i)_{i \in I}\right)$ is the image of a basis of E by $\varphi_i'^{-1}$. A each point p it is a basis of the tangent space T_pM.*

Notation 1348 $(\partial x_\alpha)_{\alpha \in A}$ *is the holonomic basis at $p = \varphi_i^{-1}(x)$ associated to the basis $(e_\alpha)_{\alpha \in A}$ by the chart (O_i, φ_i)*

When it is necessary to distinguish the basis given by different charts we will use $(\partial x_\alpha)_{\alpha \in A}, (\partial y_\alpha)_{\alpha \in A}, ...$

$\partial x_\alpha = \left(\varphi_i(p)'\right)^{-1} e_\alpha \in T_pM \Leftrightarrow \varphi_i'(p) \partial x_\alpha = e_\alpha \in E$

So a vector u_p of the tangent vector space can be written : $u_p = \sum_{\alpha \in A} u_p^\alpha \partial x_\alpha$ and at most finitely many u^α are non zero. Its image by the maps $\varphi_i'(p)$ is a vector of E : $\varphi_i'(p) u_p = \sum_{\alpha \in A} u_p^\alpha e_\alpha$ which has the same components in the basis of E.

The holonomic bases are not the only bases on the tangent space. Any other basis can be defined from a holonomic basis. They are called **non holonomic bases** . An example is the orthonormal basis if M is endowed wih a metric. But for a non holonomic basis the simple relation above : $u_p = \sum_{\alpha \in A} u_p^\alpha \partial x_\alpha = \sum_{\alpha \in A} v_p^\alpha \delta_\alpha \Rightarrow \varphi_i'(p) u_p = \sum_{\alpha \in A} v_p^\alpha \varphi_i'(p) \delta_\alpha$ does not hold any longer. The vector has not the same components in the tangent space and in E. The equality happens only if $\varphi_i'(p) \delta_\alpha = e_\alpha$ meaning that the basis is holonomic.

Theorem 1349 *The tangent space at a point of a manifold modelled on a Banach space has the structure of a Banach vector space. Different compatible charts give equivalent norm.*

Proof. Let $\left(E, (O_i, \varphi_i)_{i \in I}\right)$ be an atlas of M, and $p \in O_i$

The map : $\tau_i : T_pM \to E :: \tau_i(u) = \varphi_i'(p) u = v$ is a continuous isomorphism.

Define : $\|u\|_i = \|\varphi_i'(p) u\|_E = \|v\|_E$

With another map :

$$\|u\|_j = \left\|\varphi_j'(p) u\right\|_E = \left\|\varphi_j'(p) \circ (\varphi_i'(p))^{-1} v\right\|_E \leq \left\|\varphi_j'(p) \circ (\varphi_i'(p))^{-1}\right\|_{\mathcal{L}(E;E)} \|v\|_E \leq$$
$$\left\|\varphi_j'(p) \circ (\varphi_i'(p))^{-1}\right\|_{\mathcal{L}(E;E)} \|u\|_i$$

and similarly : $\|u\|_i \leq \left\|\varphi_i'(p) \circ \left(\varphi_j'(p)\right)^{-1}\right\|_{\mathcal{L}(E;E)} \|u\|_j$

So the two norms are equivalent (but not identical), they define the same topology. Moreover with these norms the tangent vector space is a Banach vector space. ∎

Derivative of a map

1. Definition:

Definition 1350 *For a map $f \in C_r(M; N)$ between two manifolds with atlas $\left(E, (O_i, \varphi_i)_{i \in I}\right), \left(G, (Q_j, \psi_j)_{j \in J}\right)$ there is, at each point $p \in M$ a unique continuous linear map $f'(p) \in \mathcal{L}\left(T_p M; T_{f(p)} N\right)$, called the derivative of f at p, such that : $\psi_j' \circ f'(p) \circ \left(\varphi_i^{-1}\right)'(x) = \left(\psi_j \circ f \circ \varphi_i^{-1}\right)'(x)$ with $x = \varphi_i(p)$*

f'(p) is defined in a holonomic basis by : $f'(p)\partial x_\alpha = f'(p) \circ \varphi_i'(x)^{-1} e_\alpha \in T_{f(p)} N$

We have the following commuting diagrams :

$$
\begin{array}{ccccccc}
 & f & & & & f'(p) & \\
M & \to \to \to & N & \quad & T_p M & \to \quad \to \quad \to & T_{f(p)} N \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
\downarrow \ \varphi_i & & \downarrow \ \psi_j & & \downarrow \ \varphi_i'(p) & & \downarrow \ \psi_j'(q) \\
\downarrow & F & \downarrow & & \downarrow & F'(x) & \downarrow \\
E & \to \to \to & G & & E & \to \quad \to \quad \to & G \\
\end{array}
$$

$\psi_j' \circ f'(p)\partial x_\alpha = \psi_j' \circ f'(p) \circ \varphi_i'(a)^{-1} e_\alpha \in G$

If M is m dimensional with coordinates x, N n dimensional with coordinates y the jacobian is just the matrix of f'(p) in the holonomic bases in $T_p M, T_{f(p)} N$:

$$y = \psi_j \circ f \circ \varphi_i^{-1}(x) \Leftrightarrow \alpha = 1...n : y^\alpha = F^\alpha\left(x^1, ...x^m\right)$$

$$[f'(p)] = [F'(x)] = \overbrace{\left\{\left[\frac{\partial F^\alpha}{\partial x^\beta}\right]\right\}}^{m} n = \left[\frac{\partial y^\alpha}{\partial x^\beta}\right]_{n \times m}$$

Whatever the choice of the charts in M,N there is always a derivative map f'(p), but its expression depends on the coordinates (as for any linear map). The rules when in a change of charts are given in the Tensorial bundle subsection.

Remark : usually the use of f'(p) is the most convenient. But for some demonstrations it is simpler to come back to maps between fixed vector spaces by using $\left(\psi_j \circ f \circ \varphi_i^{-1}\right)'(x)$.

2. Partial derivatives:

The partial derivatives $\frac{\partial f}{\partial x^\alpha}(p) = f'_\alpha(p)$ wih respect to the coordinate x^α is the maps $\mathcal{L}\left(\mathfrak{E}_\alpha; T_{f(p)}N\right)$ where $\mathfrak{E}_\alpha$ is the one dimensional vector subspace in T_pM generated by ∂x_α

To be consistent with the notations for affine spaces :

Notation 1351 *$f'(p)$ is the derivative $f'(p) \in \mathcal{L}\left(T_pM; T_{f(p)}N\right)$*

Notation 1352 *$\frac{\partial f}{\partial x^\alpha}(p) = f'_\alpha(p)$ are the partial derivative with respect to the coordinate $x^\alpha =$ the maps $\mathcal{L}\left(\mathfrak{E}_\alpha; T_{f(p)}N\right)$*

2. Composition of maps :

Theorem 1353 *If $f \in C_1(M; N), g \in C_1(N; P)$ then $(g \circ f)'(p) = g'(f(p)) \circ f'(p) \in \mathcal{L}\left(T_pM; T_{g \circ f(p)}P\right)$*

3. Higher derivative :

With maps on affine spaces the derivative f'(a) is a linear map depending on a, but it is still a map on fixed affine spaces, so we can consider f''(a). This is no longer possible with maps on manifolds : if f is of class r then this is the map $F(a) = \psi_j \circ f \circ \varphi_i^{-1}(a) \in C_r(E; G)$ which is r differentiable, and thus for higher derivatives we have to account for ψ_j, φ_i^{-1}. In other words f'(p) is a linear map between vector spaces which themselves depend on p, so there is no easy way to compare f'(p) to f'(q). Thus we need other tools, such as connections, to go further (see Higher tangent bundle for more).

4. Diffeomorphisms are very special maps :

i) This is a bijective map $f \in C_r(M; N)$ such that $f^{-1} \in C_r(M; N)$

ii) $f'(p)$ is inversible and $\left(f^{-1}(q)\right)' = (f'(p))^{-1} \in \mathcal{L}\left(T_{f(p)}N; T_pM\right)$: this is a continuous linear isomorphism between the tangent spaces

iii) f is an open map (it maps open subsets to open subsets)

5. Rank:

Definition 1354 *The **rank of a differentiable map** $f : M \to N$ between manifolds at a point p is the rank of its derivative f'(p). It does not depend on the choice of the maps in M,N and is necessarily $\leq \min(\dim M, \dim N)$*

Cotangent space

Definition 1355 *The **cotangent space** to a manifold M at a point p is the topological dual of the tangent space T_pM*

To follow on long custom we will not use the prime notation in this case:

Notation 1356 T_pM^* *is the cotangent space to the manifold M at p*

Definition 1357 *The transpose of the derivative of $f \in C_r(M; N)$ at p is the map : $f'(p)^t \in \mathcal{L}\left(T_{f(p)}N^*; T_pN^*\right)$*

The transpose of the derivative $\varphi_i'(p) \in \mathcal{L}(T_pM; E)$ of a chart is : $\varphi_i'(p)^t \in \mathcal{L}\left(E'; (T_pM)^*\right)$

If e^α is a basis of E' such that $e^\alpha(e_\beta) = \delta_\beta^\alpha$ (it is not uniquely defined by e_α if E is infinite dimensional) then $\varphi_i'(p)^t(e^\alpha)$ is a (holonomic) basis of T_pM^*.

Notation 1358 $dx^\alpha = \varphi_i'(p)^*(e^\alpha)$ *is the holonomic basis of T_pM^* associated to the basis $(e^\alpha)_{\alpha \in A}$ of E' by the atlas $\left(E, (O_i, \varphi_i)_{i \in I}\right)$*

So : $dx^\alpha(\partial x_\beta) = \delta_\beta^\alpha$

For a function $f \in C_1(M; K)$: $f\prime(a) \in T_pM^*$ so $f\prime(a) = \sum_{\alpha \in A} \varpi_\alpha dx^\alpha$

The partial derivatives $f_\alpha'(p) \in \mathcal{L}(\mathfrak{E}_\alpha; K)$ are scalars functions so : $f\prime(a) = \sum_{\alpha \in A} f_\alpha'(p) dx^\alpha$

The action of f'(a) on a vector $u \in T_pM$ is $f\prime(a)u = \sum_{\alpha \in A} f_\alpha'(p) u^\alpha$

The exterior differential of f is just $df = \sum_{\alpha \in A} f_\alpha'(p) dx^\alpha$ which is consistent with the usual notation (which justifies the notation dx^α)

Extremum of a function

The theorem for affine spaces can be generalized .

Theorem 1359 *If a function $f \in C_1(M; \mathbb{R})$ on a class 1 real manifold has a local extremum in $p \in M$ then f'(p)=0*

Proof. Take an atlas $\left(E, (O_i, \varphi_i)_{i \in I}\right)$ of M. If p is a local extremum on M it is a local extremum on any $O_i \ni p$. Consider the map with domain an open subset of E : $F\prime : \varphi_i(O_i) \to \mathbb{R} :: F'(a) = f' \circ \varphi_i'^{-1}$. If p=$\varphi_i(a)$ is a local extremum on O_i then $a \in \varphi_i(O_i)$ is a local extremum for $f \circ \varphi_i$ so $F'(a) = 0 \Rightarrow f'(\varphi_i(a)) = 0$. ∎

Morse's theory

A real function $f : M \to \mathbb{R}$ on a manifold can be seen as a map giving the heigth of some hills drawn above M.If this map is sliced for different elevations appear figures (in two dimensions) highlighting characteristic parts of the landscape (such that peaks or lakes). Morse's theory studies the topology of a manifold M through real functions on M (corresponding to "elevation"), using the special points where the elevation "vanishes".

1. Subsets of critical points :

Definition 1360 *For a differentiable map $f : M \to N$ a point p is* **critical** *is f'(p)=0 and regular otherwise.*

Theorem 1361 *(Lafontaine p.77) For any smooth maps $f \in C_\infty(M;N)$, M finite dimensional manifold, union of countably many compacts, N finite dimensional, the set of critical points is negligible.*

A subset X is neligible means that, if M is modelled on a Banach E, $\forall p \in M$ there is a chart (O, φ) such that p$\in$ M and $\varphi(O \cap X)$ has a null Lebesgue measure in E.

In particular :

Theorem 1362 *Sard Lemna : the set of critical values of a function defined on an open set of $\mathbb{R}^m$ has a null Lebesgue measure*

Theorem 1363 *Reeb : For any real function f defined on a compact real manifold M:*

i) if f is continuous and has exactly two critical points then M is homeomorphic to a sphere

ii) if M is smooth then the set of non critical points is open and dense in M

2. Degenerate points:

For a class 2 real function on an open subset of $\mathbb{R}^m$ the Hessian of f is the matrix of f'(p) which is a bilinear symmetric form. A critical point is degenerate if f'(a) is degenerate (then det $[F"(a)]$=0)

Theorem 1364 *Morse's lemna: If a is a critical non degenerate point of the function f on an open subset M of $\mathbb{R}^m$, then in a neighborhood of a there is a chart of M such that : $f(x) = f(a) - \sum_{\alpha=1}^{p} x_\alpha^2 + \sum_{\alpha=p+1}^{m} x_\alpha^2$*

The integer p is the index of a (for f). It does not depend on the chart, and is the dimension of the largest tangent vector subspace over which f'(a) is definite negative.

A **Morse function** is a smooth real function with no critical degenerate point. The set of Morse functions is dense in $C_\infty(M;\mathbb{R})$.

One extension of this theory is "catastroph theory", which studies how real valued functions on $\mathbb{R}^n$ behave around a point. René Thom has proven that there are no more than 14 kinds of behaviour (described as polynomials around the point).

15.3.2   The tangent bundle

Definitions

Definition 1365 *The tangent bundle over a class 1 manifold M is the set :* $TM = \cup_{p \in M} \{T_p M\}$

So an element of TM is comprised of a point p of M and a vector u of $T_p M$

Theorem 1366 *The tangent bundle over a class r manifold M with the atlas $\left(E, (O_i, \varphi_i)_{i \in I}\right)$ is a class r-1 manifold*

The cover of TM is defined by : $O'_i = \cup_{p \in O_I} \{T_p M\}$
The maps : $O'_i \to U_i \times E :: (\varphi_i(p), \varphi'_i(p) u_p)$ define a chart of TM
If M is finite dimensional, TM is a 2xdimM dimensional manifold.

Theorem 1367 *The tangent bundle over a manifold M with the atlas $\left(E, (O_i, \varphi_i)_{i \in I}\right)$ is a fiber bundle TM(M,E,π)*

TM is a manifold
Define the projection : $\pi : TM \to M :: \pi(u_p) = p$. This is a smooth surjective map and $\pi^{-1}(p) = T_p M$
Define (called a trivialization) : $\Phi_i : O_i \times E \to TM :: \Phi_i(p, u) = \varphi'_i(p)^{-1} u \in T_p M$
If $p \in O_i \cap O_j$ then $\varphi'_j(p)^{-1} u$ and $\varphi'_i(p)^{-1} u$ define the same vector of $T_p M$
All these conditions define the structure of a vector bundle with base M, modelled on E (see Fiber bundles).
A vector u_p in TM can be seen as the image of a couple $(p, u) \in M \times E$ through the maps Φ_i defined on the open cover given by an atlas.

Theorem 1368 *The tangent bundle of a Banach vector space $\overrightarrow{E}$ is the set $TM = \cup_{p \in M} \{u_p\}$. As the tangent space at any point p is $\overrightarrow{E}$ then $TM = \overrightarrow{E} \times \overrightarrow{E}$*

Similarly the tangent bundle of a Banach affine space $\left(E, \overrightarrow{E}\right)$ is $E \times \overrightarrow{E}$ and can be considered as E iself.

Theorem 1369 *If f is a differentiable map between the manifolds M,N, then f' is the map $f\prime : TM \to TN : F(U) = f'(\pi(u_p)) u_p$*

We have the following diagram with the atlas $\left(E, (O_i, \varphi_i)_{i \in I}\right), \left(G, (Q_j, \psi_j)_{j \in J}\right)$

$TM \to \to \to \to \to \to f' \to \to \to \to \to TN$
$\downarrow \varphi'_i \qquad\qquad\qquad \downarrow \psi'_j$
$\varphi_i(O_i) \times E \to \to F' \to \to \psi_j(Q_j) \times G$
$F'(x, u) = \left(\psi_j(f(p)), \psi'_j \circ f'(p) \circ \left(\varphi_i^{-1}\right)' u\right)$

Theorem 1370 *The product MxN two class r manifolds has a structure of manifold of class r with the projections $\pi_M : M \times N \to M, \pi_N : M \times N \to N$ and the tangent bundle of MxN is T(MxN)=TMxTN, $\pi'_M : T(M \times N) \to TM, \pi'_N : T(M \times N) \to TN$*

Similarly the **cotangent bundle** TM* is defined with $\pi^{-1}(p) = T_p M^*$

Notation 1371 *TM is the tangent bundle over the manifold M*
TM* is the cotangent bundle over the manifold M

Vector fields

Definition 1372 *A **vector field** over the manifold M is a map* $V : M \to TM :: V(p) = v_p$ *which associes to each point p of M a vector of the tangent space T_pM at p*

In fiber bundle parlance this is a section of the vector bundle.

Warning ! With an atlas : $\left(E, (O_i, \varphi_i)_{i \in I}\right)$ of M a holonomic basis is defined as the preimage of fixed vectors of a basis in E. So this not the same vector at the intersections : $\partial x_{i\alpha} = \varphi'_i(x)^{-1}(\varepsilon_\alpha) \neq \partial x_{j\alpha} = \varphi'_j(x)^{-1}(\varepsilon_\alpha)$

But a vector field V is always the same, whatever the open O_i. So it must be defined by a collection of maps :

$V_i : O_i \to K :: V(p) = \sum_{\alpha \in A} V_i^\alpha(p) \partial x_{i\alpha}$

If $p \in O_i \cap O_j : V(p) = \sum_{\alpha \in A} V_i^\alpha(p) \partial x_{i\alpha} = \sum_{\alpha \in A} V_j^\alpha(p) \partial x_{j\alpha}$ and $\partial x_{i\alpha} = \varphi'_i(x)^{-1} \circ \varphi_j(x)(\partial x_{j\alpha})$

In a finite dimensional manifold $\varphi'_i(x)^{-1} \circ \varphi_j(x)^{-1}$ is represented (in the holonomic bases) by a matrix : $[J_{ij}]$ and $\partial x_{i\alpha} = [J_{ij}]_\alpha^\beta (\partial x_{j\beta})$ so : $V_j^\alpha(p) = \sum_{\beta \in A} V_i^\beta(p) [J_{ij}]_\beta^\alpha$

If M is a class r manifold, TM is a class r-1 manifold, so vector fields can be defined by class r-1 maps.

Notation 1373 $\mathfrak{X}_r(TM)$ *is the set of class r vector fields on M. If r is omitted it will mean smooth vector fields*

With the structure of vector space on T_pM the usual operations : V+W, kV are well defined, so the set of vector fields on M has a vector space structure. It is infinite dimensional : the components at each p are functions (and not constant scalars) in K.

Theorem 1374 *If $V \in \mathfrak{X}(TM), W \in \mathfrak{X}(TN)$ then $X \in \mathfrak{X}(TM) \times \mathfrak{X}(TN)$: $X(p) = (V(p), W(q)) \in \mathfrak{X}(T(M \times N))$*

Theorem 1375 *(Kolar p. 16) For any manifold M modelled on E, and a family of isolated points and vectors of E : $(p_j, u_j)_{j \in J}$ there is always a vector field V such that $V(p_j) = \Phi_i(p_j, u_j)$*

Definition 1376 *The **support** of a vector field $V \in \mathfrak{X}(TM)$ is the support of the map : $V : M \to TM$.*

It is the closure of the set : $\{p \in M : V(p) \neq 0\}$

Definition 1377 *A **critical point** of a vector field V is a point p where V(p)=0*

Topology : if M is finite dimensional, the spaces of vector fields over M can be endowed with the topology of a Banach or Fréchet space (see Functional analysis). But there is no such topology available if M is infinite dimensional, even for the vector fields with compact support (as there is no compact if M is infinite dimensional).

Commutator of vector fields

Theorem 1378 *The set of of class $r \geq 1$ functions $C_r(M;K)$ over a manifold on the field K is a commutative algebra with pointwise multiplication as internal operation : $f \cdot g(p) = f(p)g(p)$.*

Theorem 1379 *(Kolar p.16) The space of vector fields $\mathfrak{X}_r(TM)$ over a manifold on the field K coincides with the set of derivations on the algebra $C_r(M;K)$*

i) A derivation over this algebra (cf Algebra) is a linear map : $D \in L(C_r(M;K); C_r(M;K))$ such that
$\forall f, g \in C_r(M;K) : D(fg) = (Df)g + f(Dg)$
ii) Take a function $f \in C_1(M;K)$ we have $f\prime(p) = \sum_{\alpha \in A} f'_\alpha(p) dx^\alpha \in T_pM^*$
A vector field can be seen as a differential operator DV acting on f :
$DV(f) = f\prime(p)V = \sum_{\alpha \in A} f'_\alpha(p) V^\alpha = \sum_{\alpha \in A} V^\alpha \frac{\partial}{\partial x^\alpha} f$
DV is a derivation on $C_r(M;K)$

Theorem 1380 *The vector space of vector fields over a manifold is a Lie algebra with the bracket, called **commutator** of vector fields : $\forall f \in C_r(M;K)$: $[V,W](f) = DV(DW(f))) - DW(DV(f))$*

Proof. If r>1, take : $DV(DW(f))) - DW(DV(f))$ it is still a derivation, thus there is a vector field denoted $[V,W]$ such that :
$\forall f \in C_r(M;K) : [V,W](f) = DV(DW(f))) - DW(DV(f))$
The operation : $[] : VM \to VM$ is bilinear and antisymmetric, and :
$[V,[W,X]] + [W,[X,V]] + [X,[V,W]] = 0$
With this operation the vector space V_rM of vector fields becomes a Lie algebra (of infinite dimension). ■
The bracket $[]$ is often called "Lie bracket", but as this is a generic name we will use the -also common - name commutator.
The components of the commutator (which is a vector field) in a holonomic basis are given by :
$[V,W]^\alpha = \sum_{\beta \in A} \left(V^\beta W'^\alpha_\beta - W^\beta V'^\alpha_\beta \right)$
By the symmetry of the second derivative of the φ_i for *holonomic* bases :
$\forall \alpha, \beta \in A : [\partial x_\alpha, \partial x_\beta] = 0$

Commutator of vectors fields on a Banach:
Let M be an open subset of a Banach vector space E. A vector field is a map
: $V : M \to E : V(u)$ with derivative : $V'(u) : E \to E \in \mathcal{L}(E;E)$
With $f \in C_r(M;K) : f'(u) \in \mathcal{L}(E;K), DV : C_r(M;K) \to K :: DV(f) = f'(u)(V(u))$
$(DV(DW(f))) - DW(DV(f)))(u) = \left(\frac{d}{du} (f'(u)(W(u))) \right) V(u) - \left(\frac{d}{du} (f'(u)(V(u))) \right) W(u)$
$= f"(u)(W(u), V(u)) + f'(u)(W'(u)(V(u))) - f"(u)(V(u), W(u)) - f'(u)(V'(u)(W(u)))$
$= f'(u)(W'(u)(V(u)) - V'(u)(W(u)))$
that we can write : $[V,W](u) = W'(u)(V(u)) - V'(u)(W(u)) = (W' \circ V - V' \circ W)(u)$

Let now M be either the set $\mathcal{L}$(E;E) of continuous maps, or its subset of inversible maps G$\mathcal{L}$(E;E), which are both manifolds, with vector bundle the set $\mathcal{L}$(E;E). A vector field is a differentiable map :

$V : M \to \mathcal{L}(E;E)$ and $f \in M : V'(f) : \mathcal{L}(E;E) \to \mathcal{L}(E;E) \in \mathcal{L}(\mathcal{L}(E;E);\mathcal{L}(E;E))$

$[V,W](f) = W'(f)(V(f)) - V'(f)(W(f)) = (W' \circ V - V' \circ W)(f)$

f related vector fields

Definition 1381 *The **push forward** of vector fields by a differentiable map $f \in C_1(M;N)$ is the linear map :*

$f_* : \mathfrak{X}(TM) \to \mathfrak{X}(TN) :: f_*(V)(f(p)) = f'(p)V(p)$

We have the following diagram :

$TM \to f' \to TN$
$\downarrow \pi_M \qquad \downarrow \pi_N$
$M \to\to f \to N$

which reads : $f_*V = f'V$
In components :
$V(p) = \sum_\alpha v^\alpha(p)\,\partial x_\alpha(p)$
$f'(p)V(p) = \sum_{\alpha\beta} [J(p)]^\alpha_\beta\, v^\beta(p)\,\partial y_\alpha(f(p))$ with $[J(p)]^\alpha_\beta = \frac{\partial y^\alpha}{\partial x^\beta}$
The vector fields v_i can be seen as : $v_i = (\varphi_i)_* V :: v_i(\varphi_i(p)) = \varphi_i'(p)V(p)$
and $\varphi_{i*}\partial x_\alpha = e_\alpha$

Theorem 1382 *(Kolar p.20)The map $f_* : TM \to TN$ has the following properties :*

i) it is a linear map : $\forall a,b \in K : f_*(aV_1 + bV_2) = af_*V_1 + bf_*V_2$

ii) it preserves the commutator : $[f_*V_1, f_*V_2] = f_*[V_1, V_2]$

iii) if f is a diffeomorphism then f_* is a Lie algebra morphism between the Lie algebras $V_r M$ and $V_r N$.

Definition 1383 *Two vector fields $V \in VM, W \in TN$ are said to be **f related** if :$W(f(p)) = f_*V(p)$*

Theorem 1384 *(Kolar p.19) If $V \in \mathfrak{X}(TM), W \in \mathfrak{X}(TN), X \in \mathfrak{X}(T(M \times N)) :$ $X(p) = (V(p), W(q))$ then X and V are π_M related, X and W are π_N related with the projections $\pi_M : M \times N \to M, \pi_N : M \times N \to N$.*

Definition 1385 *The **pull back** of vector fields by a diffeomorphism $f \in C_1(M;N)$ is the linear map :*

$f^* : \mathfrak{X}(TN) \to \mathfrak{X}(TM) :: f^*(W)(p) = (f'(p))^{-1} W(f(p))$

So: $f^* = (f^{-1})_* , \varphi_i^* e_\alpha = \partial x_\alpha$

Frames

1. A non holonomic basis in the tangent bundle is defined by : $\delta_\alpha = \sum_{\beta \in A} F_\alpha^\beta \partial_\beta$ where $F_\alpha^\beta \in K$ depends on p, and as usual if the dimension is infinite at most a finite number of them are non zero. This is equivalent to define vector fields $(\delta_\alpha)_{\alpha \in A}$ which at each point represent a basis of the tangent space. Such a set of vector fields is a (non holonomic) **frame**. One can impose some conditions to these vectors, such as being orthonormal. But of course we need to give the F_α^β and we cannot rely upon a chart : we need additional information.

2. If this operation is always possible locally (roughly in the domain of a chart - which can be large), it is usually impossible to have a unique frame of vector fields covering the whole manifold (even in finite dimensions). When this is possible the manifold is said to be **parallelizable** . For instance the only parallelizable spheres are S_1, S_3, S_7. The tangent bundle of a parallelizable manifold is trivial, in that it can be written as the product MxE. For the others, TM is in fact made of parts of MxE glued together in some complicated manner.

15.3.3 Flow of a vector field

Integral curve

Theorem 1386 *(Kolar p.17) For any manifold M, point p$\in$ M and vector field $V \in \mathfrak{X}_1(M)$ there is a map : $c : J \to M$ where J is some interval of $\mathbb{R}$ such that : c(0)=p and c'(t)=V(c(t)) for t$\in$ M. The set $\{c(t), t \in J\}$ is an **integral curve** of V.*

With an atlas $\left(E, (O_i, \varphi_i)_{i \in I}\right)$ of M, and in the domain i, c is the solution of the differential equation :

To find $x : \mathbb{R} \to U_i = \varphi_i(O_i) \subset E$ such that : $\frac{dx}{dt} = v(x(t)) = \varphi_i'(c(t)) V(c(t))$ and $x(0) = \varphi_i(p)$

The map v(x) is locally Lipschitz on U_i : it is continuously differentiable and:

$v(x+h) - v(x) = v'(x)h + \varepsilon(h)\|h\|$ and $\|v'(x)h\| \le \|v'(x)\|\|h\|$

$\varepsilon(h) \to 0 \Rightarrow \forall \delta > 0, \exists r : \|h\| \le r \Rightarrow \|\varepsilon(h)\| < \delta$

$\|v(x+h) - v(x)\| \le (\|v'(x)\| + \|\varepsilon(h)\|)\|h\| \le (\|v'(x)\| + \delta)\|h\|$

So the equation has a unique solution in a neighborhood of p.

The interval J can be finite, and the curve may not be defined on a the whole of M.

Theorem 1387 *(Lang p.94) If for a class 1 vector field V on the manifold V, and V(p)=0 for some point p, then any integral curve of V going through p is constant, meaning that $\forall t \in \mathbb{R} : c(t) = p$.*

Flow of a vector field

1. Definition:

Theorem 1388 *(Kolar p.18) For any class 1 vector field V on a manifold M and $p \in M$ there is a maximal interval $J_p \subset \mathbb{R}$ such that there is an integral curve $c: J_p \to M$ passing at p for t=0. The map : $\Phi_V : D(V) \times M \to M$, called the* **flow of the vector field**, *is smooth, $D(V)=\cup_{p \in M} J_p \times \{p\}$ is an open neighborhood of $\{0\} \times M$, and $\Phi_V(s+t,p) = \Phi_V(s, \Phi_V(p,t))$*

The last equality has the following meaning: if the right hand side exists, then the left hand side exists and they are equal, if s,t are both ≥ 0 or ≤ 0 and if the left hand side exists, then the right hand side exists and are equal.

Notation 1389 $\Phi_V(t,p)$ *is the flow of the vector field V, defined for $t \in J$ and $p \in M$*

The theorem from Kolar can be extended to infinite dimensional manifolds (Lang p.89)

As $\Phi_V(0,p) = p$ always exist, whenever $t, -t \in J_p$ then $\Phi_V(t, \Phi_V(-t,p)) = p$

Φ_V is differentiable *with respect to t* and : $\frac{\partial}{\partial t}\Phi_V(t,p)\,|_{t=0} = V(p); \frac{\partial}{\partial t}(\Phi_V(t,p)\,|_{t=\theta} = V(\Phi_V(\theta,p))$

Warning ! the partial derivative of $\Phi_V(t,p)$ with respect to p is more complicated (see below)

Theorem 1390 *For t fixed $\Phi_V(t,p)$ is a class r local diffeomorphism : there is a neighborhood n(p) such that $\Phi_V(t,p)$ is a diffeomorphism from n(p) to its image.*

2. Examples on $M=\mathbb{R}^n$

i)if V(p)=V a constant vector field. Then the integral curves are straigth lines parallel to V and passing by a given point. Take the point $A=(a_1,...a_n)$. $\Phi_V(a,t) = (y_1,...y_n)$ such that : $\frac{\partial y_i}{\partial t} = V_i, y_i(a,0) = a_i \Leftrightarrow y_i = tV_i + a_i$ so the flow of V is the affine map : $\Phi_V(a,t) = Vt + a$

ii) if V(p)=Ap $\Leftrightarrow V_i(x_1,...x_n) = \sum_{,j=1}^n A_i^j x_j$ where A is a constant matrix. Then $\Phi_V(a,t) = (y_1,...y_n)$ such that :

$\frac{\partial y_i}{\partial t}\,|_{t=\theta} = \sum_{,j=1}^n A_i^j y_j(\theta) \Rightarrow y(a,t) = (\exp tA) a$

iii) in the previous example, if $A = rI$ then $y(a,t) = (\exp tr) a$ and we have a radial flow

3. Complete flow:

Definition 1391 *The flow of a vector field is said to be* **complete** *if it is defined on the whole of $\mathbb{R} \times M$. Then $\forall t$ $\Phi_V(t, \cdot)$ is a diffeomorphism on M.*

Theorem 1392 *(Kolar p.19) Every vector field with compact support is complete.*

So on compact manifold every vector field is complete.
There is an extension of this theorem :

Theorem 1393 *(Lang p.92) For any class 1 vector field V on a manifold with atlas $(E, (O_i, \varphi_i))$ $v_i = (\varphi_i)_* V$, if :*
$\quad \forall p \in M, \exists i \in I, \exists k, r \in \mathbb{R} :$
$\quad p \in O_i, \max\left(\|v_i\|, \left\|\frac{\partial v_i}{\partial x}\right\|\right) \leq k, B\left(\varphi_i(p), r\right) \subset \varphi_i(O_i)$
$\quad$ *then the flow of V is complete.*

4. Properties of the flow:

Theorem 1394 *(Kolar p.20,21) For any class 1 vector fields V,W on a manifold M:*
$\quad \frac{\partial}{\partial t}\left(\Phi_V(t,p)_* W\right)|_{t=0} = \Phi_V(t,p)_* [V,W]$
$\quad \frac{\partial}{\partial t}\Phi_W\left(-t, \Phi_V\left(-t, \Phi_W\left(t, \Phi_V(t,p)\right)\right)\right)|_{t=0} = 0$
$\quad \frac{1}{2}\frac{\partial^2}{\partial t^2}\Phi_W\left(-t, \Phi_V\left(-t, \Phi_W\left(t, \Phi_V(t,p)\right)\right)\right)|_{t=0} = [V,W]$
$\quad$ *the following are equivalent :*
$\quad$ *i) $[V,W] = 0$*
$\quad$ *ii) $(\Phi_V)^* W = W$ whenever defined*
$\quad$ *iii) $\Phi_V(t, \Phi_W(s,p)) = \Phi_W(s, \Phi_V(t,p))$ whenever defined*

Theorem 1395 *(Kolar p.20) For a differentiable map $f \in C_1(M; N)$ between the manifolds M,N manifolds , any vector field $V \in \mathfrak{X}_1(TM) : f \circ \Phi_V = \Phi_{f_* V} \circ f$ whenever both sides are defined. If f is a diffeomorphism then similarly for $W \in \mathfrak{X}_1(TN) : f \circ \Phi_{f^* W} = \Phi_W \circ f$*

Theorem 1396 *(Kolar p.24) For any vector fields $V_k \in \mathfrak{X}_1(TM), k = 1...n$ on a real n-dimensional manifold M such that :*
$\quad$ *i) $\forall k,l : [V_k, V_l] = 0$*
$\quad$ *i) $V_k(p)$ are linearly independent at p*
$\quad$ *there is a chart centered at p such that $V_k = \partial x_k$*

5. Remarks:
i) $\frac{\partial}{\partial t}\Phi_V(t,p)|_{t=0} = V(p) \Rightarrow \frac{\partial}{\partial t}(\Phi_V(t,p)|_{t=\theta} = V(\Phi_V(\theta,p))$
Proof. Let be $T = t + \theta, \theta$ fixed
$\quad \Phi_V(T,p) = \Phi_V(t, \Phi_V(\theta,p))$
$\quad \frac{\partial}{\partial t}\Phi_V(T,p)|_{t=0} = \frac{\partial}{\partial t}(\Phi_V(t,p)|_{t=\theta} = \frac{\partial}{\partial t}\Phi_V(t, \Phi_V(\theta,p))|_{t=0} = V(\Phi_V(\theta,p))$
∎

So the flow is fully defined by the equation : $\frac{\partial}{\partial t}(\Phi_V(t,p)|_{t=0} = V(p)$
ii) If we proceed to the change of parameter : $s \to t = f(s)$ with $f : J \to J$ some function such that f(0)=0,f'(s)$\neq$ 0
$\quad \Phi_V(t,p) = \Phi_V(f(s),p) = \widehat{\Phi}_V(s,p)$
$\quad \frac{\partial}{\partial s}(\widehat{\Phi}_V(s,p)|_{s=0} = \frac{\partial}{\partial t}(\Phi_V(t,p)|_{t=f(0)}\frac{df}{ds}|_{s=0} = V(\Phi_V(f(0),p))\frac{df}{ds}|_{s=0} = V(p)\frac{df}{ds}|_{s=0}$
So it sums up to replace the vector field V by $\widehat{V}(p) = V(p)\frac{df}{ds}|_{s=0}$
iii) the Lie derivative (see next sections)
$\quad \pounds_V W = [V,W] = \frac{\partial}{\partial t}\left(\frac{\partial}{\partial p}\Phi_V(-t,p) \circ W \circ \frac{\partial}{\partial p}\Phi_V(t,p)\right)|_{t=0}$

One parameter group of diffeomorphisms

Definition 1397 *A **one parameter group of diffeomorphims** on a manifold M is a map : $F : \mathbb{R} \times M \to M$ such that for each t fixed $F(t,.)$ is a diffeomorphism on M and $\forall t, s \in \mathbb{R}, p \in M : F(t+s,p) = F(t, F(s,p)) = F(s, F(t,p)); F(0,p) = p$*

$\mathbb{R} \times M$ has a manifold structure so F has partial derivatives.

For p fixed $F(.,p) : \mathbb{R} \to M$ and $F_t'(t,p) \in T_{F(t,p)}M$ so $F_t'(t,p)|_{t=0} \in T_p M$ and there is a vector field

$V(p) = \Phi_i(p, v(p))$ with $v(p) = \varphi_i'(p)(F_t'(t,p)|_{t=0})$

So V is the **infinitesimal generator** of F : $F(t,p) = \Phi_V(t,p)$

Warning ! If M has the atlas $\left(E, (O_i, \varphi_i)_{i \in I}\right)$ the partial derivative with respect to p : $F_p'(t,p) \in \mathcal{L}\left(T_p M; T_{F(t,p)}M\right)$ and $U(t,p) = \varphi_i' \circ F_p' \circ \varphi_i'^{-1}(a) \in \mathcal{L}(E;E)$

$U(t+s,p) = \varphi_i' \circ F_p'(t+s,p) \circ \varphi_i'^{-1}(a) = \varphi_i' \circ F_p'(t, F(s,p)) \circ F_p'(s,p) \circ \varphi_i'^{-1}(a)$

$= \varphi_i' \circ F_p'(t, F(s,p)) \circ \varphi_i'^{-1} \circ \varphi_i' \circ F_p'(s,p) \circ \varphi_i'^{-1}(a) = U(t, F(s,p)) \circ U(s,p)$

So we *do not* have a one parameter group on the Banach E which would require : U(t+s,p)=U(t,p)oU(s,p).

15.4   Submanifolds

A submanifold is a part of a manifold that is itself a manifold, meaning that there is an atlas to define its structure. This can be conceived in several ways. The choice that has been made is that the structure of a submanifold must come from its "mother". Practically this calls for a specific map which injects the submanifold structure into the manifold : an embedding. But there are other ways to relate two manifolds, via immersion and submersion. The definitions vary according to the authors. We have chosen the definitions which are the most illuminating and practical, without loss of generality. The theorems cited have been adjusted to account for these differences.

The key point is that most of the relations between the manifolds M,N stem from the derivative of the map $f : M \to N$ which is linear and falls into one of 3 cases : injective, surjective or bijective.

For finite dimensional manifolds the results sum up in the following :

Theorem 1398 *(Kobayashi I p.8) For a differentiable map f from the m dimensional manifold M to the n dimensional manifold N, at any point p in M:*

i) if f'(p) is bijective there is a neighborhood n(p) such that f is a diffeomorphism from n(p) to f(n(p))

ii) if f'(p) is injective from a neighborhood n(p) to n(f(p)), f is a homeomorphism from n(p) to f(n(p)) and there are maps φ of M, ψ of N such that $F = \psi \circ f \circ \varphi^{-1}$ reads : i=1..m : $y^i(f(p)) = x^i(p)$

iii) if f'(p) is surjective from a neighborhood n(p) to n(f(p)), $f : n(p) \rightarrow N$ is open, and there are maps φ of M, ψ of N such that $F = \psi \circ f \circ \varphi^{-1}$ reads :
$i = 1..n : y^i (f(p)) = x^i (p)$

15.4.1 Submanifolds

Submanifolds

Definition 1399 *A subset M of a manifold N is a **submanifold** of N if :*
i) $G = G_1 \oplus G_2$ where G_1, G_2 are vector subspaces of G
ii) there is an atlas $\left(G, (Q_i, \psi_j)_{j \in J} \right)$ of N such that M is a manifold with atlas $\left(G_1, (M \cap Q_i, \psi_j|_{M \cap Q_i})_{i \in I} \right)$

The key point is that the manifold structure of M is defined through the structure of manifold of N. M has no manifold structure of its own. The dimension of M is $\leq$ dimension N. But it is clear that not any subset can be a submanifold.

Topologically M can be any subset, so it can be closed in N and so we have the concept of closed manifold.

Theorem 1400 *For any point p of the submanifold M in N, the tangent space $T_p M$ is a subspace of $T_p N$*

Proof. $\forall q \in N$, $\psi_j (q)$ can be uniquely written as : $\psi_j (q) = \sum_{\alpha \in B_1} x^\alpha e_\alpha + \sum_{\beta \in B_2} x^\beta e_\beta$ with $(e_\alpha)_{\alpha \in B_1}, (e_\beta)_{\beta \in B_2}$ bases of G_1, G_2

$q \in M \Leftrightarrow \forall \beta \in B_2 : x^\beta = 0$
For any vector $u_q \in T_q N : u_q = \sum_{\alpha \in B} u_q^\alpha \partial x_\alpha$
$\psi_j' (q) u_q = \sum_{\alpha \in B_1} u_q^\alpha \partial x_\alpha + \sum_{\beta \in B_2} u_q^\beta \partial x_\beta$
and $u_q \in T_q M \Leftrightarrow \forall \beta \in B_2 : u_q^\beta = 0$
So : $\forall p \in M : T_p M \subset T_p N$ and a vector tangent to N at p can be written uniquely :

$u_p = u_1 + u_2 : u_1 \in T_p M$ with $u_p \in T_p M \Leftrightarrow u_2 = 0$ ■
The vector u_2 is said to be **transversal** to M at p
If N is n finite dimensional and M is a submanifold of dimension n-1 then M is called an **hypersurface**.

Theorem 1401 *Extension of a map (Schwartz II p.442) A map $f \in C_r (M; E)$, $r \geq 1$ from a m dimensional class r submanifold M of a real manifold N which is the union of countably many compacts, to a Banach vector space E can be extended to $\widehat{f} \in C_r (N; E)$*

Conditions for a subset to be a manifold

Theorem 1402 *An open subset of a manifold is a submanifold with the same dimension.*

Theorem 1403 *A connected component of a manifold is a a submanifold with the same dimension.*

Theorem 1404 *(Schwartz II p.261) For a subset M of a n dimensional class r manifold N of a field K with atlas $\left(E, (Q_i, \psi_i)_{i \in I}\right)$, if,, $\forall p \in M$, there is, in a neighborhood of p, a chart (Q_i, ψ_i) of N such that :*
i) either $\psi_i (M \cap Q_i) = \{x \in K^n : x_{m+1} = .. = x_n = 0\}$ and M is closed
i) or $\psi_i (M \cap Q_i) = \psi (Q_i) \cap K^m$
then M is a m dimensional class r submanifold of N

Theorem 1405 *Smooth retract (Kolar p.9): If M is a class r connected finite dimensional manifold, $f \in C_r (M; M)$ such that $f \circ f = f$ then f(M) is a submanifold of M*

Embedding

The previous definition is not practical in many cases. It is more convenient to use a map, as it is done in a parametrized representation of a submanifold in $\mathbb{R}^n$. There are different definitions of an embedding. The simplest if the following.

Definition 1406 *An **embedding** is a map $f : C_r (M; N)$ between two manifolds M,N such that:*
i) f is a diffeomorphism from M to f(M)
ii) f(M) is a submanifold of N

M is the origin of the parameters, f(M) is the submanifold. So M must be a manifod, and we must know that f(M) is a submanifold. To be the image by a diffeomorphism is not sufficient. The next subsection deals with this issue.

dim M = dim f(M) $\leq$ dim N

If M,N are finite dimensional, F can be written in a neighborhood of $q \in f(M)$ and adapted charts :

$\beta = 1...m : y^\beta = F^\beta \left(x^1, ...x^m\right)$

$\beta = m + 1...n : y^\beta = 0$

The image of a vector $u_p \in M$ is $f'(p)u_p = v_1 + v_2 : v_1 \in T_p f(M)$ and $v_2 = 0$

The jacobian $[f'(p)]_m^n$ is of rank m.

If M is a m dimensional embedded submanifold of N then it is said that M has codimension n-m.

Example :

Theorem 1407 *Let $c : J \to N$ a path in the manifold N with J an interval in $\mathbb{R}$. The curve C=$\{c(t), t \in J\} \subset N$ is a connected 1 dimensional submanifold iff c is class 1 and c'(t) is never zero. If J is closed then C is compact.*

Proof. c'(t) $\neq 0$: then c is injective and a homeomorphism in N

$\psi_j' \circ c'(t)$ is a vector in G and there is an isomorphism between $\mathbb{R}$ as a vector space and the 1 dimensional vector space generated by $\psi_j' \circ c'(t)$ in G ■

Submanifolds defined by embedding

The following important theorems deal with the pending issue : is f(M) a submanifold of N ?

Theorem 1408 *Theorem of constant rank (Schwartz II .263) : If the map $f \in C_1(M; N)$ on a m dimensional manifold M to a manifold N has a constant rank s in M then :*

i) $\forall p \in M$, there is a neighborhood n(p) such that f(n(p)) is a s dimensional submanifold of N. For any $m \in n(p)$ we have : $T_{f(m)} f(n(p)) = f'(m) T_m M$.

ii) $\forall q \in f(M)$, the set $f^{-1}(q)$ is a closed m-s submanifold of M and $\forall m \in f^{-1}(q) : T_m f^{-1}(q) = \ker f'(m)$

Theorem 1409 *(Schwartz II p.263) If the map $f \in C_1(M; N)$ on a m dimensional manifold M is such that f is injective and $\forall p \in M$ f'(p) is injective :*

i) if M is compact then f(M) is a submanifold of N and f is an embedding.

ii) if f is an homeomorphism of M to f(M) then f(M) is a submanifold of N and f is an embedding.

Theorem 1410 *(Schwartz II p.264) If, for the map $f \in C_1(M; N)$ on a m dimensional manifold M, f'(p) is injective at some point p, there is a neighborhood n(p) such that f(n(p)) is a submanifold of N and f an embedding of n(p) into f(n(p)).*

Remark : L.Schwartz used a slightly different definition of an embedding. His theorems are adjusted to our definition.

Theorem 1411 *(Kolar p.10) A smooth n dimensional real manifold can be embedded in $\mathbb{R}^{2n+1}$ and $\mathbb{R}^{2n}$*

Immersion

Definition 1412 *A map $f \in C_1(M; N)$ from the manifold M to the manifold N is an* **immersion** *at p is f'(p) is injective. It is an immersion of M into N if it is an immersion at each point of M.*

In an immersion $\dim M \leq \dim N$ (f(M) is "smaller" than N so it is immersed in N)

Theorem 1413 *(Kolar p.11) If the map $f \in C_1(M; N)$ from the manifold M to the manifold N is an immersion on M, both finite dimensional, then for any p in M there is a neighborhood n(p) such that f(n(p)) is a submanifold of N and f an embedding from n(p) to f(n(p)).*

Theorem 1414 *(Kolar p.12) If the map $f \in C_1(M; N)$ from the manifold M to the manifold N is an immersion on M, both finite dimensional, if f is injective and a homeomorphism on f(M), then f(M) is a submanifold of N.*

Theorem 1415 *(Kobayashi I p.178) If the map f$\in C_1 (M; N)$ from the manifold M to the manifold N is an immersion on M, both connected and of the same dimension, if M is compact then N is compact and a covering space for M and f is a projection.*

Real submanifold of a complex manifold

We always assume that M,N are defined, as manifolds or other structure, on the same field K. However it happens that a subset of a complex manifold has the structure of a real manifold. For instance the matrix group U(n) is a real manifold comprised of complex matrices and a subgroup of GL($\mathbb{C}, n$). To deal with such situations we define the following :

Definition 1416 *A real manifold M with atlas $\left(E, (O_i, \varphi_i)_{i \in I}\right)$ is an immersed submanifold of the complex manifold N with atlas $\left(G, (Q_i, \psi_i)_{i \in I'}\right)$ if there is a map : $f : M \to N$ such that the map : $F = \psi_j \circ f \circ \varphi_i^{-1}$, whenever defined, is R-differentiable and its derivative is injective.*

The usual case is f=Identity.

Submersions

Submersions are the converse of immersions. Here M is "larger" than N so it is submersed by M. They are mainly projections of M on N and used in fiber bundles.

Definition 1417 *A map f$\in C_1 (M; N)$ from the manifold M to the manifold N is a **submersion** at p is f'(p) is surjective. It is an submersion of M into N if it is an submersion at each point of M.*

In an submersion $\dim N \leq \dim M$

Theorem 1418 *(Kolar p.11) A submersion on finite dimensional manifolds is an open map*

A fibered manifold $M(N, \pi)$ is a triple of two manifolds M, N and a map $\pi : M \to N$ which is both surjective and a submersion. It has the universal property : if f is a map f$\in C_r (N; P)$ in another manifold P then $f \circ \pi$ is class r iff f is class r (all the manifolds are assumed to be of class r).

Independant maps

This an aplication of the previous theorems to the following problem : let $f \in C_1 (\Omega; K^n), \Omega$ open in K^m. We want to tell when the n scalar maps f_i are "independant".

We can give the following meaning to this concept. f is a map between two manifolds. If $f(\Omega)$ is a p$\leq n$ dimensional submanifold of K^n , any point q in $f(\Omega)$ can be coordinated by p scalars y. If p<m we could replace the m variables x by y and get a new map which can meet the same values with fewer variables.

1) Let m$\geq n$. If f'(x) has a constant rank p then the maps are independant

2) If f'(x) has a constant rank r<m then locally $f(\Omega)$ is a r dimensional submanifold of K^n and we have n-r independent maps.

15.4.2   Distributions

Given a vector field, it is possible to define an integral curve such that its tangent at any point coincides with the vector. A distribution is a generalization of this idea : taking several vector fields, they define at each point a vector space and we look for a submanifold which admits this vector space as tangent space. We address here mainly the finite dimensional case, a more general formulation is given in the Fiber bundle part.

Distributions of Differential Geometry are not related in any way to the distributions of Functional Analysis.

Definitions
   1. Distribution:

Definition 1419 *A r dimensional **distribution** on the manifold M is a map :*
$W : M \to (TM)^r$ *such that W(p) is a r dimensional vector subspace of T_pM*

If M is an open in K^m a r dimensional distribution is a map between M and the grassmanian $Gr(K^m;r)$ which is a (m-r)r dimensional manifold.

The definition can be generalized : W(p) can be allowed to have different dimensions at different points, and even be infinite dimensional. We will limit ourself to more usual conditions.

Definition 1420 *A family $(V_j)_{j \in J}$ of vector fields on a manifold M generates a distribution W if for any point p in M the vector subspace spanned by the family is equal to W(p) : $\forall p \in M : W(p) = Span\left(V_j\left(p\right)\right)$*

So two families are equivalent with respect to a distribution if they generate the same distribution. To generate a m dimensional distribution the family must be comprised at least of m pointwise linearly independent vector fields.

   2. Integral manifold:

Definition 1421 *A connected submanifold L of M is an **integral manifold** for the distribution W on M if $\forall p \in L : T_pL = W\left(p\right)$*

So dimL=dimW. A distribution is not always integrable, and the submanifolds are usually different in each point.

An integral manifold is said to be maximal if it is not strictly contained in another integral manifold. If there is an integral manifold, there is always a unique maximal integral manifold. Thus we will assume in the following that the integral manifolds are maximal.

Definition 1422 *A distribution   W on M is **integrable** if there is a family $(L_\lambda)_{\lambda \in \Lambda}$ of maximal integral manifolds of W such that : $\forall p \in M : \exists \lambda : p \in L_\lambda$. This family defines a partition of M, called a **folliation**, and each L_λ is called a **leaf** of the folliation.*

Notice that the condition is about points of M.

$p \sim q \Leftrightarrow (p \in L_\lambda) \,\&\, (q \in L_\lambda)$ is a relation of equivalence for points in M which defines the partition of M.

Example : take a single vector field. An integral curve is an integral manifold. If there is an integral curve passing through each point then the distribution given by the vector field is integrable, but we have usually many integral submanifolds. We have a folliation, whose leaves are the curves.

3. Stability of a distribution:

Definition 1423 *A distribution W on a manifold M is **stable** by a map $f \in C_1(M; M)$ if : $\forall p \in M : f'(p)W(p) \subset W(f(p))$*

Definition 1424 *A vector field V on a manifold M is said to be an **infinitesimal automorphism of the distribution** W on M if W is stable by the flow of V*

meaning that : $\frac{\partial}{\partial p}\Phi_V(t, p)(W(p)) \subset W(\Phi_V(t, p))$ whenever the flow is defined.

The set Aut(W) of vector fields which are infinitesimal generators of W is stable.

4. Family of involutive vector fields:

Definition 1425 *A subset $V \subset \mathfrak{X}_1(TM)$ is **involutive** if $\forall V_1, V_2 \in V, \exists V_3 \in V : [V_1, V_2] = V_3$*

Conditions for integrability of a distribution

There are two main formulations, one purely geometric, the other relying on forms.

1. Geometric formulation:

Theorem 1426 *(Maliavin p.123) A distribution W on a finite dimensional manifold M is integrable iff there is an atlas $\left(E, (O_i, \varphi_i)_{i \in I}\right)$ of M such that, for any point p in M and neighborhood $n(p)$: $\forall q \in n(p) \cap O_i : \varphi_i'(q)W(q) = E_{i1}$ where $E = E_{i1} \oplus E_{i2}$*

Theorem 1427 *(Kolar p.26) For a distribution W on a finite dimensional manifold M the following conditions are equivalent:*
i) the distribution W is integrable
ii) the subset $V_W = \{V_W \in \mathfrak{X}(TM) : \forall p \in M : V(p) \in W(p)\}$ is stable:
$\forall V_1, V_2 \in V_W, \exists X \in V_W : \frac{\partial}{\partial p}\Phi_{V_1}(t, p)(V_2(p))) = X(\Phi_{V_1}(t, p))$ whenever the flow is defined.
iii) The set Aut(W)$\cap V_W$ spans W
iv) There is an involutive family $(V_j)_{j \in J}$ which generates W

2. Formulation using forms :

Theorem 1428 *(Malliavin p.133) A class 2 form $\varpi \in \Lambda_1(M;V)$ on a class 2 finite dimensional manifold M valued in a finite dimensional vector space V such that $\ker \varpi(p)$ has a constant finite dimension on M defines a distribution on M : $W(p) = \ker \varpi(p)$. This distribution is integrable iff : $\forall u, v \in W(p)$: $\varpi(p)u = 0, \varpi(p)v = 0 \Rightarrow d\varpi(u,v) = 0$*

Corollary 1429 *A function $f \in C_2(M;\mathbb{R})$ on a m dimensional manifold M such that $\dim \ker f'(p) = Cte$ defines an integrable distribution, whose folliation is given by $f(p) = Cte$*

Proof. The derivative $f'(p)$ defines a 1-form df on N. Its kernel has dimension m-1 at most.

d(df)=0 thus we have always $d\varpi(u,v) = 0$. ∎

$W(p) = \ker \varpi(p)$ is represented by a system of partial differential equations called a Pfaff system.

15.4.3   Manifold with boundary

In physics usually manifolds enclose a system. The walls are of paramount importance as it is where some conditions determining the evolution of the system are defined. Such manifolds are manifolds with boundary. They are the geometrical objects of the Stokes' theorem and are essential in partial differential equations. We present here a new theorem which gives a stricking definition of these objects.

Hypersurfaces

A hypersurface divides a manifold in two disjoint parts :

Theorem 1430 *(Schwartz IV p.305) For any n-1 dimensional class 1 submanifold M of a n dimensional class 1 real manifold N, every point p of M has a neighborhood n(p) in N such that :*

i) n(p) is homeomorphic to an open ball

ii) $M \cap n(p)$ is closed in n(p) and there are two disjoint connected subsets n_1, n_2 such that :

$n(p) = (M \cap n(p)) \cup n_1 \cup n_2$,

$\forall q \in M \cap n(p) : q \in \overline{n_1} \cap \overline{n_2}$

iii) there is a function $f : N \to \mathbb{R}$ such that : $n(p) = \{q : f(q) = 0\}, n_1 = \{q : f(q) < 0\}, n_2 = \{q : f(q) > 0\}$

Theorem 1431 *Lebesgue (Schwartz IV p.305) :Any closed class 1 hypersurface M of a finite dimensional real affine space E parts E in at least 2 regions, and exactly two if M is connected.*

Definition

There are several ways to define a manifold with boundary, always in finite dimensions. We will use only the following, which is the most general and useful (Schwartz IV p.343) :

Definition 1432 *A **manifold with boundary** is a set M :*

i) which is a subset of a n dimensional real manifold N

ii) identical to the closure of its interior : $M = \overline{\left(\overset{\circ}{M} \right)}$

iii) whose border ∂M called its boundary is a hypersurface in N

Remarks :

i) M inherits the topology of N so the interior $\overset{\circ}{M}$, the border ∂M are well defined (see topology). The condition i) prevents "spikes" or "barbed" areas protuding from M. So M is exactly the disjointed union of its interior and its boundary:

$$M = \overline{M} = \overset{\circ}{M} \cup \partial M = \left(\overline{(\overset{\circ}{M^c})} \right)^c$$

$\overset{\circ}{M} \cap \partial M = \varnothing$

$\partial M = M \cap \overline{(M^c)} = \partial (M^c)$

ii) M is closed in N, so usually *it is not a manifold*

iii) we will always assume that $\partial M \neq \varnothing$

iv) N must be a real manifold as the sign of the coordinates plays a key role

Properties

Theorem 1433 *(Schwartz IV p.343) If M is a manifold with boundary in N, N and ∂M both connected then :*

i) ∂M splits N in two disjoint regions : $\overset{\circ}{M}$ and M^c

ii) If O is an open in N and $M \cap O \neq \varnothing$ then $M \cap O$ is still a manifod with boundary : $\partial M \cap O$

iii) any point p of ∂M is adherent to M, $\overset{\circ}{M}$ and M^c

Theorem 1434 *(Lafontaine p.209) If M is a manifold with boundary in N, then there is an atlas $(O_i, \varphi_i)_{i \in I}$ of N such that :*

$\varphi_i \left(O_i \cap \overset{\circ}{M} \right) = \{x \in \varphi_i(O_i) : x_1 < 0\}$

$\varphi_i (O_i \cap \partial M) = \{x \in \varphi_i(O_i) : x_1 = 0\}$

Theorem 1435 *(Taylor 1 p.97) If M is a compact manifold with boundary in an oriented manifold N then there is no continuous retraction from M to ∂M.*

Transversal vectors

The tangent spaces $T_p \partial M$ to the boundary are hypersurfaces of the tangent space $T_p N$. The vectors of $T_p N$ which are not in $T_p \partial M$ are said to be **transversal**.

If N and ∂M are both connected then any class 1 path c(t) : $c : [a, b] \to N$ such that c(a)$\in \overset{\circ}{M}$ and c(b)$\in M^c$ meets ∂M at a unique point (see topology). For any transversal vector : $u \in T_p N, p \in \partial M$, if there is such a path with $c'(t) = ku, k > 0$ then u is said to be **outward oriented**, and inward oriented if $c'(t) = ku, k < 0$. Notice that we do not need to define an orientation on N.

Equivalently if V is a vector field such that its flow is defined from $p \in \overset{\circ}{M}$ to $q \in M^c$ then V is outward oriented if $\exists t > 0 : q = \Phi_V (t, p)$.

Fundamental theorems

Manifolds with boundary have a unique characteristic : they can be defined by a function : $f : N \to \mathbb{R}$.

It seems that the following theorems are original, so we give a full proof.

Theorem 1436 *Let N be a n dimensional smooth Hausdorff real manifold.*

i) Let $f \in C_1 (N; \mathbb{R})$ and $P = f^{-1}(0) \neq \varnothing$, if f'(p)$\neq 0$ on P then the set $M = \{p \in N : f(p) \leq 0\}$ is a manifold with boundary in N, with boundary $\partial M = P$. And : $\forall p \in \partial M, \forall u \in T_p \partial M : f'(p)u = 0$

ii) Conversely if M is a manifold with boundary in N there is a function : $f \in C_1 (N; \mathbb{R})$ such that :

$\overset{\circ}{M} = \{p \in N : f(p) < 0\}, \partial M = \{p \in N : f(p) = 0\}$

$\forall p \in \partial M : f'(p) \neq 0$ and : $\forall u \in T_p \partial M : f'(p)u = 0$, for any transversal vector v : $f'(p)v \neq 0$

If M and ∂M are connected then for any transversal outward oriented vector v : $f'(p)v > 0$

iii) for any riemannian metric on N the vector gradp defines a vector field outward oriented normal to the boundary

N is smooth finite dimensional Hausdorff, thus paracompact and admits a Riemanian metric

Proof of i)

Proof. f is continuous thus P is closed in N and M'=$\{p \in N : f(p) < 0\}$ is open.

The closure of M' is the set of points which are limit of sequences in M' : $\overline{M'} = \{\lim q_n, q_n \in M'\} = \{p \in N : f(p) \leq 0\} = M$

f has constant rank 1 on P, thus the set P is a closed n-1 submanifold of N and $\forall p \in P : T_p P = \ker f'(p)$ thus $\forall u \in T_p \partial M : f'(p)u = 0$. ∎

Proof of ii)

Proof. 1) there is an atlas $(O_i, \varphi_i)_{i \in I}$ of N such that :

$\varphi_i \left(O_i \cap \overset{\circ}{M} \right) = \{x \in \varphi_i(O_i) : x_1 < 0\}$

$\varphi_i (O_i \cap \partial M) = \{x \in \varphi_i(O_i) : x_1 = 0$

Denote : $\varphi_i^1(p) = x_1$ thus $\forall p \in M : \varphi_i^1(p) \leq 0$

N admits a smooth partition of unity subordinated to O_i :

$\chi_i \in C_\infty(N; \mathbb{R}_+) : \forall p \in O_i^c : \chi_i(p) = 0; \forall p \in N : \sum_i \chi_i(p) = 1$

Define : $f(p) = \sum_i \chi_i(p) \varphi_i^1(p)$

Thus :

$\forall p \in \overset{\circ}{M} : f(p) = \sum_i \chi_i(p) \varphi_i^1(p) < 0$

$\forall p \in \partial M : f(p) = \sum_i \chi_i(p) \varphi_i^1(p) = 0$

Conversely :

$\sum_i \chi_i(p) = 1 \Rightarrow J = \{i \in I : \chi_i(p) \neq 0\} \neq \varnothing$

let be : $L = \{i \in I : p \in O_i\} \neq \varnothing$ so $\forall i \notin L : \chi_i(p) = 0$

Thus $J \cap L \neq \varnothing$ and $f(p) = \sum_{i \in J \cap L} \chi_i(p) \varphi_i^1(p)$

let $p \in N : f(p) < 0$: there is at least one $j \in J \cap L$ such that $\varphi_i^1(p) < 0 \Rightarrow p \in \overset{\circ}{M}$

let $p \in N : f(p) = 0 : \sum_{i \in J \cap L} \chi_i(p) \varphi_i^1(p) = 0, \varphi_i^1(p) \leq 0 \Rightarrow \varphi_i^1(p) = 0$

2) Take a path on the boundary : $c : [a, b] \to \partial M$

$c(t) \in \partial M \Rightarrow \varphi_i^1(c(t)) = 0 \Rightarrow (\varphi_i^1)'(c(t)) c'(t) = 0 \Rightarrow \forall p \in \partial M, \forall u \in T_p \partial M : (\varphi_i^1(p))' u = 0$

$f'(p)u = \sum_i \left(\chi_i'(p) \varphi_i^1(p) u + \chi_i(p) (\varphi_i^1)'(p) u_x \right)$

$p \in \partial M \Rightarrow \varphi_i^1(p) = 0 \Rightarrow f'(p)u = \sum_i \chi_i(p) (\varphi_i^1)'(p) u = 0$

3) Let $p \in \partial M$ and v_1 transversal vector. We can take a basis of $T_p N$ comprised of v_1 and n-1 vectors $(v_\alpha)_{\alpha=2}^n$ of $T_p \partial M$

$\forall u \in T_p N : u = \sum_{\alpha=1}^n u_\alpha v_\alpha$

$f'(p)u = \sum_{\alpha=1}^n u_\alpha f'(p)v_\alpha = u_1 f'(p)v_1$

As $f'(p) \neq 0$ thus for any transversal vector we have $f'(p)u \neq 0$

4) Take a vector field V such that its flow is defined from $p \in \overset{\circ}{M}$ to $q \in M^c$ and $V(p) = v_1$

v_1 is outward oriented if $\exists t > 0 : q = \Phi_V(t, p)$. .Then :

$t \leq 0 \Rightarrow \Phi_V(t, p) \in \overset{\circ}{M} \Rightarrow f(\Phi_V(t, p)) \leq 0$

$t = 0 \Rightarrow f(\Phi_V(t, p)) = 0$

$\frac{d}{dt} \Phi_V(t, p)|_{t=0} = V(p) = v_1$

$\frac{d}{dt} f(\Phi_V(t, p))|_{t=0} = f'(p)v_1 = \lim_{t \to 0^-} \frac{1}{t} f(\Phi_V(t, p)) \geq 0$

5) Let g be a riemannian form on N. So we can associate to the 1-form df a vector field : $V^\alpha = g^{\alpha\beta} \partial_\beta f$ and $f'(p)V = g^{\alpha\beta} \partial_\beta f \partial_\alpha f \geq 0$ is zero only if f'(p)=0. So we can define a vector field outward oriented. ∎

Proof of iii)

Proof. V is normal (for the metric g) to the boundary : $u \in \partial M : g_{\alpha\beta} u^\alpha V^\beta = g_{\alpha\beta} u^\alpha g^{\beta\gamma} \partial_\gamma f = u^\gamma \partial_\gamma f = f'(p)u = 0$ ∎

Theorem 1437 *Let M be a m dimensional smooth Hausdorff real manifold, $f \in C_1(M; \mathbb{R})$ such that f'(p)$\neq$ 0.on M*

i) Then $M_t = \{p \in M : f(p) \leq t\}$, for any $t \in f(M)$ is a family of manifolds with boundary $\partial M_t = \{p \in M : f(p) = t\}$

ii) f defines a folliation of M with leaves ∂M_t

iii) if M is connected compact then f(M)=[a,b] and there is a transversal vector field V whose flow is a diffeomorphism for the boundaries $\partial M_t = \Phi_V(\partial M_a, t)$

Proof. M is smooth finite dimensional Hausdorff, thus paracompact and admits a Riemanian metric

i) f'(p)$\neq$ 0. Thus f'(p) has constant rank m-1.

The theorem of constant rank tells us that for any t in f(M)$\subset \mathbb{R}$ the set $f^{-1}(t)$ is a closed m-1 submanifold of M and

$\forall p \in f^{-1}(t) : T_p f^{-1}(t) = \ker f'(p)$

We have a family of manifolds with boundary : $M_t = \{p \in N : f(p) \leq t\}$ for $t \in f(M)$

ii) Frobenius theorem tells us that f defines a foliation of M, with leaves the boundary $\partial M_t = \{p \in N : f(p) = t\}$

And we have $\forall p \in M_t, \ker f'(p) = T_p \partial M_t$

$\Rightarrow \forall u \in T_p \partial M_t : f'(p)u = 0, \forall u \in T_p M, u \notin T_p \partial M_t : f'(p)u \neq 0$

iii) If M is connected then f(M)=$|a,b|$ an interval in $\mathbb{R}$. If M is compact then f has a maximum and a minimum :

$a \leq f(p) \leq b$

There is a Riemannian structure on M, let be g the bilinear form and define the vector field :

$V = \frac{gradf}{\|gradf\|^2} :: \forall p \in M : V(p) = \frac{1}{\lambda}\left(g(p)^{\alpha\beta} \partial_\beta f(p)\right) \partial_\alpha$ with $\lambda = \sum_{\alpha\beta} g^{\alpha\beta}(\partial_\alpha f)(\partial_\beta f) > 0$

$f'(p)V = \frac{1}{\lambda}g^{\alpha\beta}\partial_\beta f \partial_\alpha f = 1$ So V is a vector field everywhere transversal and outward oriented. Take $p_a \in \partial M_a$

The flow $\Phi_V(p_a, s)$ of V is such that : $\forall s \geq 0 : \exists \theta \in [a,b] : \Phi_V(p_a, s) \in \partial M_\theta$ whenever defined.

Define : $h : \mathbb{R} \to [a,b] : h(s) = f(\Phi_V(p_a, s))$

$\frac{\partial}{\partial s}\Phi_V(p,s)|_{s=\theta} = V(\Phi_V(p,\theta))$

$\frac{d}{ds}h(s)|_{s=\theta} = f'(\Phi_V(p,\theta))V(\Phi_V(p,\theta)) = 1 \Rightarrow h(s) = s$

and we have : $\Phi_V(p_a, s) \in \partial M_s$ ∎

An application of these theorems is the propagation of waves. Let us take $M = \mathbb{R}^4$ endowed with the Lorentz metric, that is the space of special relativity. Take a constant vector field V of components (v_1, v_2, v_3, c) with $\sum_{\alpha=1}^{3}(v_\alpha)^2 = c^2$. This is a field of rays of ligth. Take $f(p) = \langle p, V \rangle = p_1 v_1 + p_2 v_2 + p_3 v_3 - cp_4$

The folliation is the family of hyperplanes orthogonal to V. A wave is represented by a map : $F : M \to E$ with E some vector space, such that : $F(p) = \chi \circ f(p)$ where $\chi : \mathbb{R} \to E$. So the wave has the same value on any point on the front wave, meaning the hyperplanes f(p)=s. f(p) is the phase of the wave.

For any component $F_i(p)$ we have the following derivatives :

$\alpha = 1,2,3 : \frac{\partial}{\partial p_\alpha}F_i = F_i'(-v_\alpha) \to \frac{\partial^2}{\partial p_\alpha^2}F_i = F_i''(v_\alpha^2)$

$\frac{\partial}{\partial p_4}F_i = F_i'(-c) \to \frac{\partial^2}{\partial p_\alpha^2}F_i = F_i''(c^2)$

so : $\frac{\partial^2}{\partial p_1^2}F_i + \frac{\partial^2}{\partial p_2^2}F_i + \frac{\partial^2}{\partial p_3^2}F_i - \frac{\partial^2}{\partial p_4^2}F_i = 0 = \Box F_i$

F follows the wave equation. We have plane waves with wave vector V.

We would have spherical waves with $f(p) = \langle p, p \rangle$

Another example is the surfaces of constant energy in symplectic manifolds.

15.4.4 Homology on manifolds

This is the generalization of the concepts in the Affine Spaces, exposed in the Algebra part. On this subject see Nakahara p.230.

A r-simplex S^r on $\mathbb{R}^n$ is the convex hull of the r dimensional subspaces defined by r+1 independants points $(A_i)_{i1}^k$:

$S^r = \langle A_0, ... A_r \rangle = \{P \in \mathbb{R}^n : P = \sum_{i=0}^r t_i A_i; 0 \le t_i \le 1, \sum_{i=0}^r t_i = 1\}$

A simplex is not a differentiable manifold, but is a topological (class 0) manifold with boundary. It can be oriented.

Definition 1438 *A **r-simplex on a manifold** M modeled on $\mathbb{R}^n$ is the image of a r-simplex S^r on $\mathbb{R}^n$ by a smooth map : $f : \mathbb{R}^n \to M$*

It is denoted : $M^r = \langle p_0, p_1, ...p_r \rangle = \langle f(A_0), ...f(A_r) \rangle$

Definition 1439 *A **r-chain on a manifold** M is the formal sum : $\sum_i k_i M_i^r$ where M_i^r is any r-simplex on M counted positively with its orientation, and $k_i \in \mathbb{R}$*

Notice two differences with the affine case :

i) here the coefficients $k_i \in \mathbb{R}$ (in the linear case the coefficients are in $\mathbb{Z}$).

ii) we do not precise a simplicial complex C : any r simplex on M is suitable

The set of r-chains on M is denoted $G_r(M)$. It is a group with formal addition.

Definition 1440 *The **border of the simplex** $\langle p_0, p_1, ...p_r \rangle$ on the manifold M is the r-1-chain :*

$\partial \langle p_0, p_1, ...p_r \rangle = \sum_{k=0}^r (-1)^k \langle p_0, p_1, ., \widehat{p_k}, ...p_r \rangle$ *where the point p_k has been removed. Conventionnaly : $\partial \langle p_0 \rangle = 0$*

$M^r = f(S^r) \Rightarrow \partial M^r = f(\partial S^r)$

$\partial^2 = 0$

A r-chain such that $\partial M^r = 0$ is a **r-cycle**. The set $Z_r(M) = \ker(\partial)$ is the r-cycle subgroup of $G_r(M)$ and $Z_0(M) = G_0(M)$

Conversely if there is $M^{r+1} \in G_{r+1}(M)$ such that $N = \partial M \in G_r(C)$ then N is called a **r-border**. The set of r-borders is a subgroup $B_r(M)$ of $G_r(M)$ and $B_n(M) = 0$.One has : $B_r(M) \subset Z_r(M) \subset G_r(M)$

The **r-homology group** of M is the quotient set : $H_r(M) = Z_r(M)/B_r(M)$

The rth Betti number of M is $b_r(M) = \dim H_r(M)$

16   TENSORIAL BUNDLE

16.1   Tensor fields

16.1.1   Tensors in the tangent space

1. The tensorial product of copies of the vectorial space tangent and its topological dual at every point of a manifold is well defined as for any other vector space (see Algebra). So contravariant and covariant tensors, and mixed tensors of any type (r,s) are defined in the usual way at ever point of a manifold.

2. All operations valid on tensors apply fully on the tangent space at one point p of a manifold M : $\otimes_s^r T_p M$ is a vector space over the field K (the same as M), product or contraction of tensors are legitimate operations. The space $\otimes T_p M$ of tensors of all types is an algebra over K.

3. With an atlas $\left(E, (O_i, \varphi_i)_{i \in I}\right)$ of the manifold M, at any point p the maps : $\varphi_i'(p) : T_p M \rightarrow E$ are vector space isomorphisms, so there is a unique extension to an isomorphism of algebras in $L(\otimes T_p M; \otimes E)$ which preserves the type of tensors and commutes with contraction (see Tensors). So any chart (O_i, φ_i) can be uniquely extended to a chart $(O_i, \varphi_{i,r,s})$:

$\varphi_{i,r,s}(p) : \otimes_s^r T_p M \rightarrow \otimes_s^r E$

$\forall S_p, T_p \in \otimes_s^r T_p M, k, k' \in K :$

$\varphi_{i,r,s}(p)(k S_p + k' T_p) = k \varphi_{i,r,s}(p) S_p + k' \varphi_{i,r,s}(p)(T_p)$

$\varphi_{i,r,s}(p)(S_p \otimes T_p) = \varphi_{i,r,s}(p)(S_p) \otimes \varphi_{i,r,s}(p)(T_p)$

$\varphi_{i,r,s}(p)(Trace(S_p)) = Trace(\varphi_{i,r,s}(p)((S_p)))$

with the property :

$(\varphi_i'(p) \otimes \varphi_i'(p))(u_p \otimes v_p) = \varphi_i'(p)(u_p) \otimes \varphi_i'(p)(v_p)$

$\varphi_i'(p) \otimes \left(\varphi_i'(p)^t\right)^{-1}(u_p \otimes \mu_p) = \varphi_i'(p)(u_p) \otimes \left(\varphi_i'(p)^t\right)^{-1}(\mu_p), ...$

4. Tensors on $T_p M$ can be expressed locally in any basis of $T_p M$. The natural bases are the bases induced by a chart, with vectors $(\partial x_\alpha)_{\alpha \in A}$ and covectors $(dx^\alpha)_{\alpha \in A}$ with : $\partial x_\alpha = \varphi_i(p)^{-1} e_\alpha, dx^\alpha = \varphi_i(p)^t e^\alpha$ where $(e_\alpha)_{\alpha \in A}$ is a basis of E and $(e^\alpha)_{\alpha \in A}$ a basis of E'.

The components of a tensor T_p in $\otimes_s^r T_p M$ *expressed in a holonomic basis* are :

$T_p = \sum_{\alpha_1 ... \alpha_r} \sum_{\beta_1 .... \beta_s} t_{\beta_1 ... \beta_q}^{\alpha_1 ... \alpha_r} \partial x_{\alpha_1} \otimes .. \otimes \partial x_{\alpha_r} \otimes dx^{\beta_1} \otimes ... \otimes dx^{\beta_s}$

and : $\varphi_{i,r,s}(p)\left(\partial x_{\alpha_1} \otimes .. \otimes \partial x_{\alpha_r} \otimes dx^{\beta_1} \otimes ... \otimes dx^{\beta_s}\right) = e_{\alpha_1} \otimes .. \otimes e_{\alpha_r} \otimes e^{\beta_1} \otimes ... \otimes e^{\beta_s}$

The image of T_p by the previous map $\varphi_{i,r,s}(p)$ is a tensor t in $\otimes_s^r E$ which has the same components in the basis of $\otimes_s^r E$:

$\varphi_{i,r,s}(p) T_p = \sum_{\alpha_1 ... \alpha_r} \sum_{\beta_1 .... \beta_s} t_{\beta_1 ... \beta_q}^{\alpha_1 ... \alpha_r} e_{\alpha_1} \otimes .. \otimes e_{\alpha_r} \otimes e^{\beta_1} \otimes ... \otimes e^{\beta_s}$

16.1.2   Change of charts

1. In a change of basis in the tangent space the usual rules apply (see Algebra). When the change of bases is induced by a change of chart the matrix giving the new basis with respect to the old one is given by the jacobian.

2. If the old chart is (O_i, φ_i) and the new chart : (O_i, ψ_i) (we can assume that the domains are the same, this issue does not matter here).

Coordinates in the old chart : $x = \varphi_i(p)$

Coordinates in the new chart : $y = \psi_i(p)$

Old holonomic basis :

$\partial x_\alpha = \varphi'_i(p)^{-1} e_\alpha,$

$dx^\alpha = \varphi'_i(x)^t e^\alpha$ with $dx^\alpha(\partial x_\beta) = \delta^\alpha_\beta$

New holonomic basis :

$\partial y_\alpha = \psi'_i(p)^{-1} e_\alpha,$

$dy^\alpha = \psi'_i(y)^* e^\alpha$ with $dy^\alpha(\partial y_\beta) = \delta^\alpha_\beta$

In a n-dimensional manifold the new coordinates $\left(y^i\right)_{i=1}^n$ are expressed with respect to the old coordinates by :

$\alpha = 1..n : y^\alpha = F^\alpha\left(x^1, ...x^n\right) \Leftrightarrow \psi_i(p) = F \circ \varphi_i(p) \Leftrightarrow F(x) = \psi_i \circ \varphi_i^{-1}(x)$

The jacobian is : $J = [F'(x)] = \left[J^\alpha_\beta\right] = \left[\dfrac{\partial F^\alpha}{\partial x^\beta}\right]_{n \times n} \simeq \left[\dfrac{\partial y^\alpha}{\partial x^\beta}\right]$

$\partial y_\alpha = \sum_\beta \left[J^{-1}\right]^\beta_\alpha \partial x_\beta \simeq \dfrac{\partial}{\partial y^\alpha} = \sum_\beta \dfrac{\partial x^\beta}{\partial y^\alpha}\dfrac{\partial}{\partial x^\beta} \Leftrightarrow \partial y_\alpha = \psi'^{-1}_i \circ \varphi'_i(p)\partial x_\alpha$

$dy^\alpha = \sum_\beta [J]^\beta_\alpha dx_\beta \simeq dy^\alpha = \sum_\beta \dfrac{\partial y^\alpha}{\partial x^\beta}dx^\beta \Leftrightarrow dy^\alpha = \psi'^*_i \circ \left(\varphi'_i(x)^*\right)^{-1} dx_\alpha$

The components of vectors :

$u_p = \sum_\alpha u^\alpha_p \partial x_\alpha = \sum_\alpha \widehat{u}^\alpha_p \partial y_\alpha$ with $\widehat{u}^\alpha_p = \sum_\beta J^\alpha_\beta u^\beta_p \simeq \sum_\beta \dfrac{\partial y^\alpha}{\partial x^\beta} u^\beta_p$

The components of covectors :

$\mu_p = \sum_\alpha \mu_{p\alpha} dx^\alpha = \sum_\alpha \widehat{\mu}_{p\alpha} dy^\alpha$ with $\widehat{\mu}_{p\alpha} = \sum_\beta \left[J^{-1}\right]^\beta_\alpha \mu_{p\beta} \simeq \sum_\beta \dfrac{\partial x^\beta}{\partial y^\alpha}\mu_{p\beta}$

For a type (r,s) tensor :

$T = \sum_{\alpha_1...\alpha_r}\sum_{\beta_1....\beta_s} t^{\alpha_1...\alpha_r}_{\beta_1...\beta_q}\partial x_{\alpha_1} \otimes .. \otimes \partial x_{\alpha_r} \otimes dx^{\beta_1} \otimes ... \otimes dx^{\beta_s}$

$T = \sum_{\alpha_1...\alpha_r}\sum_{\beta_1....\beta_s} \widehat{t}^{\alpha_1...\alpha_r}_{\beta_1...\beta_q}\partial y_{\alpha_1} \otimes .. \otimes \partial y_{\alpha_r} \otimes dy^{\beta_1} \otimes ... \otimes dy^{\beta_s}$

with :

$\widehat{t}^{\alpha_1...\alpha_r}_{\beta_1...\beta_q} = \sum_{\lambda_1...\lambda_r}\sum_{\mu_1....\mu_s} t^{\lambda_1...\lambda_r}_{\mu_1...\mu_s} [J]^{\alpha_1}_{\lambda_1} .. [J]^{\alpha_r}_{\lambda_r} \left[J^{-1}\right]^{\mu_1}_{\beta_1} .. \left[J^{-1}\right]^{\mu_s}_{\beta_s}$

$\widehat{t}^{\alpha_1...\alpha_r}_{\beta_1...\beta_q}(q) = \sum_{\lambda_1...\lambda_r}\sum_{\mu_1....\mu_s} t^{\lambda_1...\lambda_r}_{\mu_1...\mu_s} \dfrac{\partial y^{\alpha_1}}{\partial x^{\lambda_1}}...\dfrac{\partial y^{\alpha_r}}{\partial x^{\lambda_r}}\dfrac{\partial x^{\mu_1}}{\partial y^{\beta_1}}...\dfrac{\partial x^{\mu_s}}{\partial y^{\beta_s}}$

For a r-form :

$\varpi = \sum_{(\alpha_1...\alpha_r)} \varpi_{\alpha_1...\alpha_r}dx^{\alpha_1} \otimes dx^{\alpha_2} \otimes ... \otimes dx^{\alpha_r} = \sum_{\{\alpha_1...\alpha_r\}} \varpi_{\alpha_1...\alpha_r}dx^{\alpha_1} \wedge dx^{\alpha_2} \wedge ... \wedge dx^{\alpha_r}$

$\varpi = \sum_{(\alpha_1...\alpha_r)} \widehat{\varpi}_{\alpha_1...\alpha_r}dy^{\alpha_1} \otimes dy^{\alpha_2} \otimes ... \otimes dy^{\alpha_r} = \sum_{\{\alpha_1...\alpha_r\}} \widehat{\varpi}_{\alpha_1...\alpha_r}dy^{\alpha_1} \wedge dy^{\alpha_2} \wedge ... \wedge dy^{\alpha_r}$

with $\widehat{\varpi}_{\alpha_1...\alpha_r} = \sum_{\{\beta_1....\beta_r\}} \varpi_{\beta_1...\beta_r} \det\left[J^{-1}\right]^{\beta_1...\beta_r}_{\alpha_1...\alpha_r}$

where $\det\left[J^{-1}\right]^{\beta_1...\beta_r}_{\alpha_1...\alpha_r}$ is the determinant of the matrix $\left[J^{-1}\right]$ with elements row β_k column α_l

16.1.3 Tensor bundle

The tensor bundle is defined in a similar way as the vector bundle.

Definition 1441 *The (r,s) tensor bundle is the set* $\otimes^r_s TM = \cup_{p \in M} \otimes^r_s T_pM$

Theorem 1442 $\otimes_s^r TM$ *has the structure of a class r-1 manifold, with dimension (rs+1)xdimM*

The open cover of $\otimes_s^r TM$ is defined by : $O_i' = \cup_{p \in O_I} \{\otimes_s^r T_p M\}$

The maps : $O_i' \to U_i \times \otimes_s^r E :: (\varphi_i(p), \varphi_{i,r,s}(p) T_p)$ define an atlas of $\otimes_s^r TM$

The dimension of $\otimes_s^r TM$ is (rs+1)xdimM. Indeed we need m coordinates for p and mxrxs components for T_p.

Theorem 1443 $\otimes_s^r TM$ *has the structure of vector bundle over M, modeled on* $\otimes_s^r E$

$\otimes_s^r TM$ is a manifold

Define the projection : $\pi_{r,s} : \otimes_s^r TM \to M :: \pi_{r,s}(T_p) = p$. This is a smooth surjective map and $\pi_{r,s}^{-1}(p) = \otimes_s^r T_p M$

Define the trivialization : $\Phi_{i,r,s} : O_i \times \otimes_s^r E \to \otimes_s^r TM :: \Phi_{i,r,s}(p,t) = \varphi_{i,r,s}^{-1}(\varphi_i(p)) t \in \otimes_s^r T_p M$. This is a class c-1 map if the manifold is of class c.

If $p \in O_i \cap O_j$ then $\varphi_{i,r,s}^{-1} \circ \varphi_{j,r,s}(p) t$ and $\varphi_{j,r,s}^{-1} \circ \varphi_{i,r,s}(p) t$ define the same tensor of $\otimes_s^r T_p M$

Theorem 1444 $\otimes_s^r TM$ *has a structure of a vector space with pointwise operations.*

16.1.4 Tensor fields

Definition

Definition 1445 *A **tensor field** of type (r,s) is a map* $: T : M \to \otimes_s^r TM$ *which associates at each point p of M a tensor T(p)*

A tensor field of type (r,s) over the open $U_i \subset E$ is a map : $t_i : U_i \to \otimes_s^r E$

A tensor field is a collection of maps : $T_i : O_i \times \otimes_s^r E \to \otimes_s^r TM :: T(p) = \Phi_{i,r,s}(p, t_i(\varphi_i(p)))$ with t_i a tensor field on E.

This reads :

$T(p) = \sum_{\alpha_1 \ldots \alpha_r} \sum_{\beta_1 \ldots \beta_s} t_{\beta_1 \ldots \beta_q}^{\alpha_1 \ldots \alpha_r} \partial_{\alpha_1} \otimes .. \otimes \partial_{\alpha_r} \otimes dx^{\beta_1} \otimes ... \otimes dx^{\beta_s}$

$\varphi_{i,r,s}(p)(T(p)) = \sum_{\alpha_1 \ldots \alpha_r} \sum_{\beta_1 \ldots \beta_s} t_{\beta_1 \ldots \beta_q}^{\alpha_1 \ldots \alpha_r}(\varphi_i(p)) e_{\alpha_1} \otimes .. \otimes e_{\alpha_r} \otimes e^{\beta_1} \otimes ... \otimes e^{\beta_s}$

The tensor field if of class c if all the functions $t_i : U_i \to \otimes_s^r E$ are of class c.

Warning! As with vector fields, the components of a given tensor fields vary through the domains of an atlas.

Notation 1446 $\mathfrak{X}_c(\otimes_s^r TM)$ *is the set of fields of class c type (r,s) tensors on the manifold M*

$\mathfrak{X}_c(\Lambda_s TM)$ *is the set of fields of class c antisymmetric type (0,s) tensors on the manifold M*

A vector field can be seen as a (1,0) type contravariant tensor field $\mathfrak{X}\left(\otimes_0^1 TM\right) \simeq \mathfrak{X}\left(TM\right)$

A vector field on the cotangent bundle is a (0,1) type covariant tensor field $\mathfrak{X}\left(\otimes_1^0 TM\right) \simeq \mathfrak{X}\left(TM^*\right)$

Scalars can be seen a (0,0) tensors. Similarly a map : $T : M \to K$ is just a scalar *function*. So the 0-covariant tensor fields are scalar maps: $\mathfrak{X}\left(\otimes_0^0 TM\right) = \mathfrak{X}\left(\wedge_0 TM\right) \simeq C\left(M;K\right)$

Operations on tensor fields

1. All usual operations with tensors are available with tensor fields when they are implemented at the same point of M.

With the tensor product (pointwise) the set of tensor fields over a manifold is an algebra denoted $\mathfrak{X}\left(\otimes TM\right) = \oplus_{r,s}\mathfrak{X}\left(\otimes_s^r TM\right)$.

If the manifold is of class c, $\otimes_s^r TM$ is a class r-1 manifold, the tensor field is of class c-1 if the map : $t : U_i \to \otimes_s^r E$ is of class c-1. So the maps : $t_{\beta_1...\beta_q}^{\alpha_1...\alpha_r} : M \to \mathbb{R}$ giving the components of the tensor field in a holonomic basis are class c-1 scalar functions. And this property does not depend of the choice of an atlas of class c.

2. The trace operator (see the Algebra part) is the unique linear map :
$Tr : \mathfrak{X}\left(\otimes_1^1 TM\right) \to C\left(M;K\right)$ such that $Tr\left(\varpi \otimes V\right) = \varpi\left(V\right)$

From the trace operator one can define the contraction on tensors as a linear map : $\mathfrak{X}\left(\otimes_s^r TM\right) \to \mathfrak{X}\left(\otimes_{s-1}^{r-1} TM\right)$ which depends on the choice of the indices to be contracted.

3. It is common to meet complicated operators over vector fields, including derivatives, and to wonder if they have some tensorial significance. A useful criterium is the following (Kolar p.61):

If the multilinear (with scalars) map on vector fields
$F \in \mathcal{L}^s\left(\mathfrak{X}\left(TM\right)^s ; \mathfrak{X}\left(\otimes^r TM\right)\right)$ is still linear for any function, meaning :
$\forall f_k \in C_\infty\left(M;K\right), \forall \left(V_k\right)_{k=1}^s, F\left(f_1 V_1, ... f_s V_s\right) = f_1 f_2 ... f_s F\left(V_1, ... V_s\right)$
then $\exists T \in \mathfrak{X}\left(\otimes_s^r TM\right) :: \forall \left(V_k\right)_{k=1}^s, F\left(V_1, ... V_s\right) = T\left(V_1, ... V_s\right)$

16.1.5 Pull back, push forward

The push-forward and the pull back of a vector field by a map can be generalized but work differently according to the type of tensors. For some transformations we need only a differentiable map, for others we need a diffeomorphism, and then the two operations - push forward and pull back - are the opposite of the other.

Definitions

1. For any differentiable map f between the manifolds M,N (on the same field):

Push-forward for vector fields :
$f_* : \mathfrak{X}\left(TM\right) \to \mathfrak{X}\left(TN\right) :: f_* V = f'V \Leftrightarrow f_* V\left(f\left(p\right)\right) = f'(p)V\left(p\right)$
Pull-back for 0-forms (functions) :
$f^* : \mathfrak{X}\left(\Lambda_0 TN^*\right) \to \mathfrak{X}\left(\Lambda_0 TM^*\right) :: f^* h = h \circ f \Leftrightarrow f^* h\left(p\right) = h\left(f\left(p\right)\right)$

Pull-back for 1-forms :

$f^* : \mathfrak{X}(\Lambda_1 TN^*) \to \mathfrak{X}(\Lambda_1 TM^*) :: f^*\mu = \mu \circ f' \Leftrightarrow f^*\mu(p) = \mu(f(p)) \circ f'(p)$

Notice that the operations above do not need a diffeormorphism, so M,N do not need to have the same dimension.

2. For any diffeomorphism f between the manifolds M,N (which implies that they must have the same dimension) we have the inverse operations :

Pull-back for vector fields :

$f^* : \mathfrak{X}(TN) \to \mathfrak{X}(TM) :: f^*W = (f^{-1})' V \Leftrightarrow f^*W(p) = (f^{-1})'(f(p))W(f(p))$

Push-forward for 0-forms (functions) :

$f_* : \mathfrak{X}(\Lambda_0 TM^*) \to \mathfrak{X}(\Lambda_0 TN^*) :: f_*g = g \circ f^{-1} \Leftrightarrow f_*g(q) = g(f^{-1}(q))$

Push-forward for 1-forms :

$f_* : \mathfrak{X}(\Lambda_1 TM^*) \to \mathfrak{X}(\Lambda_1 TN^*) :: f_*\lambda = \varpi \circ (f^{-1})' \Leftrightarrow f_*\lambda(q) = \lambda(f^{-1}(q)) \circ (f^{-1})'(q)$

3. For any mix (r,s) type tensor, on finite dimensional manifolds M,N with the same dimension, and any diffeomorphism $f : M \to N$

Push-forward of a tensor :

$f_* : \mathfrak{X}(\otimes_s^r T_p M) \to \mathfrak{X}(\otimes_s^r T_p N) :: (f_* T_p)(f(p)) = f'_{r,s}(p) T_p$

Pull-back of a tensor :

$f^* : \mathfrak{X}(\otimes_s^r T_p M) \to \mathfrak{X}(\otimes_s^r T_p N) :: (f^* S_q)(f^{-1}(q)) = (f'_{r,s})^{-1}(q) S_q$

where $f'_{r,s}(p) : \otimes_s^r T_p M \to \otimes_s^r T_{f(p)} N$ is the extension to the algebras of the isomorphism : $f'(p) : T_p M \to T_{f(p)} N$

Properties

Theorem 1447 *(Kolar p.62) Whenever they are defined, the push forward f_* and pull back f^* of tensors are linear operators (with scalars) :*

$f^* \in \mathcal{L}(\mathfrak{X}(\otimes_s^r T_p M); \mathfrak{X}(\otimes_s^r T_p N))$

$f_* \in \mathcal{L}(\mathfrak{X}(\otimes_s^r T_p M); \mathfrak{X}(\otimes_s^r T_p N))$

which are the inverse map of the other :

$f^* = (f^{-1})_*$

$f_* = (f^{-1})^*$

They preserve the commutator of vector fields:

$[f_* V_1, f_* V_2] = f_* [V_1, V_2]$

$[f^* V_1, f^* V_2] = f^* [V_1, V_2]$

and the exterior product of r-forms :

$f^*(\varpi \wedge \pi) = f^*\varpi \wedge f^*\pi$

$f_*(\varpi \wedge \pi) = f_*\varpi \wedge f_*\pi$

They can be composed with the rules :

$(f \circ g)^* = g^* \circ f^*$

$(f \circ g)_* = f_* \circ g_*$

They commute with the exterior differential (if f is of class 2) :

$d(f^*\varpi) = f^*(d\varpi)$

$d(f_*\varpi) = f_*(d\varpi)$

So for functions :

$h \in C\left(N;K\right) : \left(f^*h\right)'\left(p\right) = h'\left(f\left(p\right)\right) \circ f'\left(p\right)$

$g \in C\left(M;K\right) : \left(f_*g\right)'\left(q\right) = g'\left(f^{-1}\left(q\right)\right) \circ \left(f^{-1}\right)'\left(q\right)$

and for 1-forms and vector fields :

$\mu \in \mathfrak{X}\left(\Lambda_1 TN^*\right), V \in \mathfrak{X}\left(TM\right) : f^*\mu\left(V\right) = \mu\left(f_*V\right)$

$\lambda \in \mathfrak{X}\left(\Lambda_1 TM^*\right), W \in \mathfrak{X}\left(TN\right) : f_*\lambda\left(W\right) = \lambda\left(f^*W\right)$

Components expressions

For a diffeomorphism f between the n dimensional manifolds M with atlas $\left(K^n, \left(O_i, \varphi_i\right)_{i\in I}\right)$ and the manifold N with atlas $\left(K^n, \left(Q_j, \psi_j\right)_{j\in J}\right)$ the formulas are

Push forward : $f_* : \mathfrak{X}\left(\otimes_s^r TM\right) \to \mathfrak{X}\left(\otimes_s^r TN\right)$

$T\left(p\right) = \sum_{\alpha_1...\alpha_r} \sum_{\beta_1....\beta_s} T_{\beta_1...\beta_q}^{\alpha_1...\alpha_r}\left(p\right) \partial x_{\alpha_1} \otimes .. \otimes \partial x_{\alpha_r} \otimes dx^{\beta_1} \otimes ... \otimes dx^{\beta_s}$

$\left(f_*T\right)\left(q\right) = \sum_{\alpha_1...\alpha_r} \sum_{\beta_1....\beta_s} \widehat{T}_{\beta_1...\beta_q}^{\alpha_1...\alpha_r}\left(q\right) \partial y_{\alpha_1} \otimes .. \otimes \partial y_{\alpha_r} \otimes dy^{\beta_1} \otimes ... \otimes dy^{\beta_s}$

with :

$\widehat{T}_{\beta_1...\beta_q}^{\alpha_1...\alpha_r}\left(q\right) = \sum_{\lambda_1...\lambda_r} \sum_{\mu_1....\mu_s} T_{\mu_1...\mu_s}^{\lambda_1...\lambda_r}\left(f^{-1}\left(q\right)\right) \left[J\right]_{\lambda_1}^{\alpha_1} .. \left[J\right]_{\lambda_r}^{\alpha_r} \left[J^{-1}\right]_{\beta_1}^{\mu_1} .. \left[J^{-1}\right]_{\beta_s}^{\mu_s}$

$\widehat{T}_{\beta_1...\beta_q}^{\alpha_1...\alpha_r}\left(q\right) = \sum_{\lambda_1...\lambda_r} \sum_{\mu_1....\mu_s} T_{\mu_1...\mu_s}^{\lambda_1...\lambda_r}\left(f^{-1}\left(q\right)\right) \frac{\partial y^{\alpha_1}}{\partial x^{\lambda_1}}...\frac{\partial y^{\alpha_r}}{\partial x^{\lambda_r}} \frac{\partial x^{\mu_1}}{\partial y^{\beta_1}}...\frac{\partial x^{\mu_s}}{\partial y^{\beta_s}}$

Pull-back : $f^* : \mathfrak{X}\left(\otimes_s^r TN\right) \to \mathfrak{X}\left(\otimes_s^r TM\right)$

$S\left(q\right) = \sum_{\alpha_1...\alpha_r} \sum_{\beta_1....\beta_s} S_{\beta_1...\beta_q}^{\alpha_1...\alpha_r}\left(q\right) \partial y_{\alpha_1} \otimes .. \otimes \partial y_{\alpha_r} \otimes dy^{\beta_1} \otimes ... \otimes dy^{\beta_s}$

$f^*S\left(p\right) = \sum_{\alpha_1...\alpha_r} \sum_{\beta_1....\beta_s} \widehat{S}_{\beta_1...\beta_q}^{\alpha_1...\alpha_r}\left(p\right) \partial x_{\alpha_1} \otimes .. \otimes \partial x_{\alpha_r} \otimes dx^{\beta_1} \otimes ... \otimes dx^{\beta_s}$

with :

$\widehat{S}_{\beta_1...\beta_q}^{\alpha_1...\alpha_r}\left(p\right) = \sum_{\lambda_1...\lambda_r} \sum_{\mu_1....\mu_s} S_{\mu_1...\mu_s}^{\lambda_1...\lambda_r}\left(f\left(p\right)\right) \left[J^{-1}\right]_{\lambda_1}^{\alpha_1} .. \left[J^{-1}\right]_{\lambda_r}^{\alpha_r} \left[J\right]_{\beta_1}^{\mu_1} .. \left[J\right]_{\beta_s}^{\mu_s}$

$\widehat{S}_{\beta_1...\beta_q}^{\alpha_1...\alpha_r}\left(q\right) = \sum_{\lambda_1...\lambda_r} \sum_{\mu_1....\mu_s} S_{\mu_1...\mu_s}^{\lambda_1...\lambda_r}\left(f\left(p\right)\right) \frac{\partial x^{\alpha_1}}{\partial y^{\lambda_1}}...\frac{\partial x^{\alpha_r}}{\partial y^{\lambda_r}} \frac{\partial y^{\mu_1}}{\partial x^{\beta_1}}...\frac{\partial y^{\mu_s}}{\partial x^{\beta_s}}$

where x are the coordinates on M, y the coordinates on N, and J is the jacobian :

$\left[J\right] = \left[F'\left(x\right)\right] = \left[\frac{\partial y^{\alpha}}{\partial x^{\beta}}\right] ; \left[F'\left(x\right)\right]^{-1} = \left[J\right]^{-1} = \left[\frac{\partial x^{\alpha}}{\partial y^{\beta}}\right]$

F is the transition map : $F : \varphi_i\left(O_i\right) \to \psi_j\left(Q_j\right) :: y = \psi_j \circ f \circ \varphi_i^{-1}\left(x\right) = F\left(x\right)$

For a r-form these formulas simplify :

Push forward :

$\varpi = \sum_{\left(\alpha_1...\alpha_r\right)} \varpi_{\alpha_1...\alpha_r} dx^{\alpha_1} \otimes dx^{\alpha_2} \otimes ... \otimes dx^{\alpha_r} = \sum_{\{\alpha_1...\alpha_r\}} \varpi_{\alpha_1...\alpha_r} dx^{\alpha_1} \wedge dx^{\alpha_2} \wedge ... \wedge dx^{\alpha_r}$

$\left(f_*\varpi\right)\left(q\right) = \sum_{\left(\alpha_1...\alpha_r\right)} \widehat{\varpi}_{\alpha_1...\alpha_r}\left(q\right) dy^{\alpha_1} \otimes dy^{\alpha_2} \otimes ... \otimes dy^{\alpha_r} = \sum_{\{\alpha_1...\alpha_r\}} \widehat{\varpi}_{\alpha_1...\alpha_r}\left(q\right) dy^{\alpha_1} \wedge dy^{\alpha_2} \wedge ... \wedge dy^{\alpha_r}$

with :

$\widehat{\varpi}_{\alpha_1...\alpha_r}\left(q\right) = \sum_{\{\beta_1....\beta_r\}} \varpi_{\beta_1...\beta_r}\left(f^{-1}\left(q\right)\right) \det\left[J^{-1}\right]_{\alpha_1...\alpha_r}^{\beta_1...\beta_r}$

$= \sum_{\mu_1....\mu_s} \varpi_{\mu_1...\mu_s}\left(f^{-1}\left(q\right)\right) \left[J^{-1}\right]_{\alpha_1}^{\mu_1} .. \left[J^{-1}\right]_{\alpha_r}^{\mu_r}$

Pull-back :

$\varpi\left(q\right) = \sum_{\left(\alpha_1...\alpha_r\right)} \varpi_{\alpha_1...\alpha_r}\left(q\right) dy^{\alpha_1} \otimes dy^{\alpha_2} \otimes ... \otimes dy^{\alpha_r} = \sum_{\{\alpha_1...\alpha_r\}} \varpi_{\alpha_1...\alpha_r}\left(q\right) dy^{\alpha_1} \wedge dy^{\alpha_2} \wedge ... \wedge dy^{\alpha_r}$

$f^* \varpi (p) = \sum_{(\alpha_1 ... \alpha_r)} \widehat{\varpi}_{\alpha_1 ... \alpha_r} (p) \, dx^{\alpha_1} \otimes dx^{\alpha_2} \otimes ... \otimes dx^{\alpha_r} = \sum_{\{\alpha_1 ... \alpha_r\}} \widehat{\varpi}_{\alpha_1 ... \alpha_r} (p) \, dx^{\alpha_1} \wedge dx^{\alpha_2} \wedge ... \wedge dx^{\alpha_r}$

with :

$\widehat{\varpi}_{\alpha_1 ... \alpha_r} (p) = \sum_{\{\beta_1 ,..., \beta_r\}} \varpi_{\beta_1 ... \beta_r} (f(p)) \det [J]_{\alpha_1 ... \alpha_r}^{\beta_1 ... \beta_r} = \sum_{\mu_1 ,..., \mu_s} \varpi_{\mu_1 ... \mu_s} (f(p)) [J]_{\alpha_1}^{\mu_1} .. [J]_{\alpha_r}^{\mu_r}$

where $\det [J^{-1}]_{\alpha_1 ... \alpha_r}^{\beta_1 ... \beta_r}$ is the determinant of the matrix $[J^{-1}]$ with r column $(\alpha_1, .. \alpha_r)$ comprised each of the components $\{\beta_1 ... \beta_r\}$

Remark :
A change of chart can also be formalized as a push-forward :
$\varphi_i : O_i \rightarrow U_i :: x = \varphi_i (p)$
$\psi_i : O_i \rightarrow V_i :: y = \psi_i (p)$
$\psi_i \circ \varphi_i^{-1} : O_i \rightarrow O_i :: y = \psi_i \circ \varphi_i^{-1} (x)$
The change of coordinates of a tensor is the push forward : $\widehat{t}_i = \left(\psi_i \circ \varphi_i^{-1} \right)_* t_i$.
As the components in the holonomic basis are the same as in E, we have the same relations between T and $\widehat{T}$

16.2 Lie derivative

16.2.1 Invariance, transport and derivation

As this is a problem frequently met in physics it is useful to understand how the mathematics work.

Equivariance
1. Let be two observers doing some experiments about the same phenomenon. They use models which are described in the tensor bundle of the same manifold M modelled on a Banach E, but using different charts.
   Observer 1 : charts $(O_i, \varphi_i)_{i \in I}$, $\varphi_i (O_i) = U_i \subset E$ with coordinates x
   Observer 2 : charts $(O_i, \psi_i)_{i \in I}$, $\psi_i (O_i) = V_i \subset E$ with coordinates y
   We assume that the cover O_i is the same (it does not matter here).
   The physical phenomenon is represented in the models by a tensor $T \in \otimes_s^r TM$. This is a geometrical quantity : it does not depend on the charts used. The measures are done at the same point p.
   Observer 1 mesures the components of $T : T (p) = \sum_{\alpha_1 ... \alpha_r} \sum_{\beta_1 ,..., \beta_s} t_{\beta_1 ... \beta_q}^{\alpha_1 ... \alpha_r} (p) \, \partial x_{\alpha_1} \otimes .. \otimes \partial x_{\alpha_r} \otimes dx^{\beta_1} \otimes ... \otimes dx^{\beta_s}$
   Observer 2 mesures the components of $T : T (p) = \sum_{\alpha_1 ... \alpha_r} \sum_{\beta_1 ,..., \beta_s} s_{\beta_1 ... \beta_q}^{\alpha_1 ... \alpha_r} (q) \, \partial y_{\alpha_1} \otimes .. \otimes \partial y_{\alpha_r} \otimes dy^{\beta_1} \otimes ... \otimes dy^{\beta_s}$
   So in their respective charts the measures are :
   $t = (\varphi_i)_* T$
   $s = (\psi_i)_* T$
   Passing from one set of measures to the other is a change of charts :
   $s = \left(\psi_i \circ \varphi_i^{-1} \right)_* t = (\psi_i)_* \circ \left(\varphi_i^{-1} \right)_* t$
   So the measures are related : they are **equivariant**. They change according to the rules of the charts.

2. This is just the same rule as in affine space : when we use different frames, we need to adjust the mesures according to the proper rules in order to be able to make any sensible comparison. The big difference here is that these rules should apply for any point p, and any set of transition maps $\psi_i \circ \varphi_i^{-1}$. So we have stronger conditions for the specification of the functions $t_{\beta_1...\beta_q}^{\alpha_1...\alpha_r}(p)$.

Invariance

1. If both observers find the *same* numerical results the tensor is indeed special : $t = (\psi_i)_* \circ (\varphi_i^{-1})_* t$. It is **invariant** by some specific diffeomorphism $(\psi_i \circ \varphi_i^{-1})$ and the physical phenomenon has a **symmetry** which is usually described by the action of a group. Among these groups the one parameter groups of diffeomorphisms have a special interest because they are easily related to physical systems and can be characterized by an infinitesimal generator which is a vector field (they are the axes of the symmetry).

2. Invariance can also occur when with one single operator doing measurements of the same phenomenon at two different points. If he uses the same chart $(O_i, \varphi_i)_{i \in I}$ with coordinates x as above :

Observation 1 at point p : $T(p) = \sum_{\alpha_1...\alpha_r} \sum_{\beta_1....\beta_s} t_{\beta_1...\beta_q}^{\alpha_1...\alpha_r}(p) \partial x_{\alpha_1} \otimes .. \otimes \partial x_{\alpha_r} \otimes dx^{\beta_1} \otimes ... \otimes dx^{\beta_s}$

Observation 2 at point q : $T(q) = \sum_{\alpha_1...\alpha_r} \sum_{\beta_1....\beta_s} t_{\beta_1...\beta_q}^{\alpha_1...\alpha_r}(q) \partial x_{\alpha_1} \otimes .. \otimes \partial x_{\alpha_r} \otimes dx^{\beta_1} \otimes ... \otimes dx^{\beta_s}$

Here we have a big difference with affine spaces, where we can always use a common basis $(e_\alpha)_{\alpha \in A}$. Even if the chart is the same, the tangent spaces are not the same, and we cannot tell much without some tool to compare the holonomic bases at p and q. Let us assume that we have such a tool. So we can "transport" T(p) at q and express it in the holonomic frame at q. If we find the same figures we can say that T is invariant when we go from p to q. More generally if we have such a procedure we can give a precise meaning to the variation of the tensor field between p and q.

In differential geometry we have several tools to transport tensors on tensor bundles : the "push-forward", which is quite general, and derivations.

Transport by push forward

If there is a diffeomorphism : $f : M \to M$ then with the push-forward $\widehat{T} = f_* T$ reads :
$\widehat{T}(f(p)) = f^* T_j(p) = \Phi_{i,r,s}(p, t_j(\varphi_j \circ f(p))) = \Phi_{i,r,s}(p, t_j(\varphi_i(p)))$
The components of the tensor $\widehat{T}$, expressed in the holonomic basis are :
$\widehat{T}_{\beta_1...\beta_q}^{\alpha_1...\alpha_r}(f(p)) = \sum_{\lambda_1...\lambda_r} \sum_{\mu_1....\mu_s} T_{\mu_1...\mu_s}^{\lambda_1...\lambda_r}(p) [J]_{\lambda_1}^{\alpha_1} .. [J]_{\lambda_r}^{\alpha_r} [J^{-1}]_{\beta_1}^{\mu_1} .. [J^{-1}]_{\beta_s}^{\mu_s}$
where $[J] = \left[\frac{\partial y^\alpha}{\partial x^\beta}\right]$ is the matrix of f'(p)
So they are a linear (possibly complicated) combination of the components of T.

Definition 1448 *A tensor T is said to be **invariant** by a diffeomorphism on the manifold M $f : M \to M$ if : $T = f^* T \Leftrightarrow T = f_* T$*

If T is invariant then the components of the tensor at p and f(p) must be linearly dependent.

If there is a one parameter group of diffeomorphisms, it has an infinitesimal generator which is a vector field V. If a tensor T is invariant by such a one parameter group the Lie derivative $\pounds_V T = 0$.

Derivation

1. Not all physical phenomenons are invariant, and of course we want some tool to measure how a tensor changes when we go from p to q. This is just what we do with the derivative : $T(a+h) = T(a) + T'(a)h + \epsilon(h)h$ .So we need a derivative for tensor fields. Manifolds are not isotropic : all directions on the tangent spaces are not equivalent. Thus it is clear that a derivation depends on the direction u along which we differentiate, meaning something like the derivative D_u along a vector, and the direction u will vary at each point. There are two ways to do it : either u is the tangent c'(t) to some curve p=c(t), or u=V(p) with V a vector field. For now on let us assume that u is given by some vector field V (we would have the same results with c'(t)).

So we shall look for a map : $D_V : \mathfrak{X}(\otimes_s^r TM) \rightarrow \mathfrak{X}(\otimes_s^r TM)$ with $V \in \mathfrak{X}(TM)$ which preserves the type of the tensor field.

2. We wish also that this derivation D has some nice useful properties, as classical derivatives :

i) it should be linear in V : $\forall V, V' \in VM, k, k' \in K : D_{kV+k'V'}T = kD_V T + k'D_{V'}T$ so that we can compute easily the derivative along the vectors of a basis. This condition, joined with that $D_V T$ should be a tensor of the same type as T leads to say that :

$D : \mathfrak{X}(\otimes_s^r TM) \rightarrow \mathfrak{X}(\otimes_{s+1}^r TM)$

For a (0,0) type tensor, meaning a function on M, the result is a 1-form.

ii) D should be a linear operator on the tensor fields :

$\forall S, T \in \mathfrak{X}(\otimes_s^r TM), k, k' \in K : D(kS + k'T) = kDS + k'DT$

iii) D should obey the Leibnitz rule with respect to the tensorial product :

$D(S \otimes T) = (DS) \otimes T + S \otimes (DT)$

The tensor fields have a structure of algebra $\mathfrak{X}(\otimes TM)$ with the tensor product. These conditions make D a derivation on $\mathfrak{X}(\otimes TM)$ (see Tensors in the Algebra part).

iv) In addition we wish some kind of relation between the operation on TM and TM*. Without a bilinear form the only general relation which is available is the trace operator, well defined and is the unique linear map : $Tr : \mathfrak{X}(\otimes_1^1 TM) \rightarrow C(M; K)$ such that $\forall \varpi \in \mathfrak{X}(\otimes_1^0 TM), V \in \mathfrak{X}(\otimes_0^1 TM)$: $Tr(\varpi \otimes V) = \varpi(V)$

So we impose that D commutes with the trace operator. Then it commutes with the contraction of tensors.

3. There is a general theorem (Kobayashi p.30) which tells that any derivation can be written as a linear combination of a Lie derivative and a covariant derivative, which are seen in the next subsections.

4. The parallel transport of a tensor T by a derivation D along a vector field is done by defining the "transported tensor" $\widehat{T}$ as the solution of a differential equation $D_V \widehat{T} = 0$ and the initial condition $\widehat{T}(p) = T(p)$. Similarly a tensor is invariant if $D_V T = 0$.

5. Conversely with a derivative we can look for the curves such that a given tensor is invariant. We can see these curves as integral curves for both the transport and the tensor. Of special interest are the curves such that their tangent are themselves invariant by parallel transport. They are the geodesics. If the covariant derivative comes from a metric these curves are integral curves of the length.

16.2.2   Lie derivative

The idea is to use the flow of a vector field to transport a tensor : at each point along a curve we use the diffeomorphism to push forward the tensor along the curve and we compute a derivative at this point. It is clear that the result depends on the vector field : in some way the Lie derivative is a generalization of the derivative along a vector. This is a very general tool, in that it does not require any other ingredient than the vector field V. The covariant derivative is richer, but requires the definition of specific maps.

Definition

Let T be a tensor field $T \in \mathfrak{X}(\otimes_s^r TM)$ and V a vector field $V \in \mathfrak{X}(TM)$. The flow Φ_V is defined in a domain which is an open neighborhood of 0xM and in this domain it is a diffeomorphism $M \to M$. For t small the tensor at $\Phi_V(-t, p)$ is pushed forward at p by $\Phi_V(t, .)$:

$(\Phi_V(t, .)_* T)(p) = (\Phi_V(t, .))_{r,s}(p) T(\Phi_V(-t, p))$

The two tensors are now in the same tangent space at p, and it is possible to compute for any p in M :

$\mathcal{L}_V T(p) = \lim_{t \to 0} \frac{1}{t}((\Phi_V(t, .)_* T)(p) - T(p)) = \lim_{t \to 0} \frac{1}{t}((\Phi_V(t, .)_* T)(p) - (\Phi_V(0, .)_* T)(p))$

The limit exists as the components and J are differentiable and :

Definition 1449 *The **Lie derivative of a tensor** field $T \in \mathfrak{X}(\otimes_s^r TM)$ along the vector field $V \in \mathfrak{X}(TM)$ is : $\mathcal{L}_V T(p) = \frac{d}{dt}((\Phi_V(t, .)_* T)(p))|_{t=0}$*

In components :
$(\Phi_V(t, .)_* T)_{\beta_1...\beta_q}^{\alpha_1...\alpha_r}(p)$
$= \sum_{\lambda_1...\lambda_r} \sum_{\mu_1,...,\mu_s} T_{\mu_1...\mu_s}^{\lambda_1...\lambda_r}(\Phi_V(-t, p)) [J]_{\lambda_1}^{\alpha_1} .. [J]_{\lambda_r}^{\alpha_r} [J^{-1}]_{\beta_1}^{\mu_1} .. [J^{-1}]_{\beta_s}^{\mu_s}$
with : $F : U_i \to U_i :: y = \varphi_i \circ \Phi_V(t, .) \circ \varphi_i^{-1}(x) = F(x)$
$[F'(x)] = [J] = \left[\frac{\partial y^\alpha}{\partial x^\beta}\right]$
so the derivatives of $\Phi_V(t, p)$ with respect to p are involved

Properties of the Lie derivative

Theorem 1450 *(Kolar p.63) The Lie derivative along a vector field $V \in \mathfrak{X}(TM)$ on a manifold M is a derivation on the algebra $\mathfrak{X}(\otimes TM)$:*
i) it is a linear operator : $\pounds_V \in \mathcal{L}(\mathfrak{X}(\otimes^r_s TM); \mathfrak{X}(\otimes^r_s TM))$
ii) it is linear with respect to the vector field V
iii) it follows the Leibnitz rule with respect to the tensorial product
Moreover:
iv) it commutes with any contraction between tensors
v) antisymmetric tensors go to antisymmetric tensors

So $\forall V, W \in \mathfrak{X}(TM), \forall k, k' \in K, \forall S, T \in \mathfrak{X}(\otimes TM)$
$\pounds_{V+W} = \pounds_V + \pounds_W$
$\pounds_V(kS + k'T) = k\pounds_V S + k'\pounds_V T$
$\pounds_V(S \otimes T) = (\pounds_V S) \otimes T + S \otimes (\pounds_V T)$
which gives with $f \in C(M; K) : \pounds_V(f \times T) = (\pounds_V f) \times T + f \times (\pounds_V T)$
(pointwise multiplication)

Theorem 1451 *(Kobayashi I p.32) For any vector field $V \in \mathfrak{X}(TM)$ and tensor field $T \in \mathfrak{X}(\otimes TM)$:*
$\Phi_V(-t, .)^* \pounds_V T = -\frac{d}{dt}(\Phi_V(-t, .)^* T)|_{t=0}$

Theorem 1452 *The Lie derivative of a vector field is the commutator of the vectors fields :*

$\forall V, W \in \mathfrak{X}(TM) : \pounds_V W = -\pounds_W V = [V, W]$
$f \in C(M; K) : \pounds_V f = i_V f = V(f) = \sum_\alpha V^\alpha \partial_\alpha f = f'(V)$
Remark : $V(f)$ is the differential operator associated to V acting on the function f

Theorem 1453 *Exterior product:*
$\forall \lambda, \mu \in \mathfrak{X}(\Lambda TM^*) : \pounds_V(\lambda \wedge \mu) = (\pounds_V \lambda) \wedge \mu + \lambda \wedge (\pounds_V \mu)$

Theorem 1454 *Action of a form on a vector:*
$\forall \lambda \in \mathfrak{X}(\Lambda_1 TM^*), W \in \mathfrak{X}(TM) : \pounds_V(\lambda(W)) = (\pounds_V \lambda)(W) + \lambda(\pounds_V W)$
$\forall \lambda \in \mathfrak{X}(\Lambda_r TM^*), W_1, ..W_r \in \mathfrak{X}(TM) :$
$(\pounds_V \lambda)(W_1, ...W_r) = V(\lambda(W_1, ...W_r)) - \sum_{k=1}^r \lambda(W_1, ... [V, W_k] ..W_r)$

Remark : $V(\lambda(W_1, ...W_r))$ is the differential operator associated to V acting on the function $\lambda(W_1, ...W_r)$

Theorem 1455 *Interior product of a r form and a vector field :*
$\forall \lambda \in \mathfrak{X}(\Lambda_r TM^*), V, W \in \mathfrak{X}(TM) : \pounds_V(i_W \lambda) = i_{\pounds_V W}(\lambda) + i_W(\pounds_V \lambda)$

Remind that : $(i_W \lambda)(W_1, ...W_{r-1}) = \lambda(W, W_1, ...W_{r-1})$

Theorem 1456 *The* **bracket of the Lie derivative operators** $\pounds_V, \pounds_W$ *for the vector fields V,W is :* $[\pounds_V, \pounds_W] = \pounds_V \circ \pounds_W - \pounds_W \circ \pounds_V$ *and we have* $:[\pounds_V, \pounds_W] = \pounds_{[V,W]}$

Parallel transport   The Lie derivative along a curve is defined only if this is the integral curve of a tensor field V. The transport is then equivalent to the push forward by the flow of the vector field.

Theorem 1457 *(Kobayashi I p.33) A tensor field T is invariant by the flow of a vector field V iff $£_V T = 0$*

This result stands for any one parameter group of diffeomorphism, with V = its infinitesimal generator.

In the next subsections are studied several one parameter group of diffeomorphisms which preserve some tensor T (the metric of a pseudo riemannian manifold, the 2 form of a symplectic manifold). These groups have an infinitesimal generator V and $£_V T = 0$.

16.3   Exterior algebra

16.3.1   Definitions

For any manifold M a r-form in $T_p M^*$ is an antisymmetric r covariant tensor in the tangent space at p. A field of r-form is a field of antisymmetric r covariant tensor in the tangent bundle TM. All the operations on the exterior algebra of $T_p M$ are available, and similarly for the fields of r-forms, whenever they are implemented pointwise (for a fixed p). So the exterior product of two r forms fields can be computed.

Notation 1458 $\mathfrak{X}(\Lambda T M^*) = \oplus_{r=0}^{\dim M} \mathfrak{X}(\Lambda_r T M^*)$ *is the **exterior algebra** of the manifold M.*

This is an algebra over the same field K as M with pointwise operations.
In a holonomic basis a field of r forms reads :
$\varpi(p) = \sum_{(\alpha_1...\alpha_r)} \varpi_{\alpha_1...\alpha_r}(p) \, dx^{\alpha_1} \otimes dx^{\alpha_2} \otimes ... \otimes dx^{\alpha_r}$
$= \sum_{\{\alpha_1...\alpha_r\}} \varpi_{\alpha_1...\alpha_r}(p) \, dx^{\alpha_1} \wedge dx^{\alpha_2} \wedge ... \wedge dx^{\alpha_r}$
with $(\alpha_1...\alpha_r)$ any r indexes in A, $\{\alpha_1...\alpha_r\}$ any ordered set of r indexes in A

$\sigma \in \mathfrak{S}_r : \varpi_{\sigma(\alpha_1...\alpha_r)} = \epsilon(\sigma) \varpi_{\alpha_1...\alpha_r}$
$\varpi_{\alpha_1...\alpha_r} : M \to K$ the form is of class c if the functions are of class c.

To each r form is associated a r multilinear antisymmetric map, valued in the field K :
$\forall \varpi \in \mathfrak{X}(\wedge_r T M^*), V_1, ..., V_r \in \mathfrak{X}(TM) :$
$\varpi(V_1, ..., V_r) = \sum_{(\alpha_1...\alpha_r)} \varpi_{\alpha_1...\alpha_r} v_1^{\alpha_1} v_2^{\alpha_2} ... v_r^{\alpha_r}$
Similarly a r-form on M can be valued in a *fixed* Banach vector space F. It reads :
$\varpi = \sum_{\{\alpha_1...\alpha_r\}} \sum_{i=1}^q \varpi_{\alpha_1...\alpha_r}^i f_i \otimes dx^{\alpha_1} \wedge dx^{\alpha_2} \wedge ... \wedge dx^{\alpha_r}$
where $(f_i)_{i=1}^q$ is a basis of F.
All the results for r-forms valued in K can be extended to these forms.

Notation 1459 $\Lambda_r(M; F)$ *is the space of fields of r-forms on the manifold M valued in the fixed vector space F*

So $\mathfrak{X}(\Lambda_r TM^*) = \Lambda_r(M; K)$.

Definition 1460 *The **canonical form** on the manifold M modeled on E is the field of 1 form valued in E :* $\Theta = \sum_{\alpha \in A} dx^\alpha \otimes e_\alpha$

So : $\Theta(p)(u_p) = \sum_{\alpha \in A} u_p^\alpha e_\alpha \in E$

It is also possible to consider r-forms valued in TM. They read :
$\varpi = \sum_{\beta\{\alpha_1...\alpha_r\}} \sum_\beta \varpi_{\alpha_1...\alpha_r}^\beta \partial x_\beta \otimes dx^{\alpha_1} \wedge dx^{\alpha_2} \wedge ... \wedge dx^{\alpha_r} \in \mathfrak{X}(\Lambda_r TM^* \otimes TM)$
So this is a field of mixed tensors $\otimes_r^1 TM$ which is antisymmetric in the lower indices. To keep it short we use the :

Notation 1461 $\Lambda_r(M; TM)$ *is the space of fields of r-forms on the manifold M valued in the tangent bundle*

Their theory involves the derivatives on graded algebras and leads to the Frölicher-Nijenhuis bracket (see Kolar p.67). We will see more about them in the Fiber bundle part.

16.3.2 Interior product

The interior product $i_V \varpi$ of a r-form ϖ and a vector V is an operation which, when implemented pointwise, can be extended to fields of r forms and vectors on a manifold M, with the same properties. In a holonomic basis of M:
$\forall \varpi \in \mathfrak{X}(\wedge_r TM^*), \pi \in \mathfrak{X}(\wedge_s TM^*), V, W \in \mathfrak{X}(TM), f \in C(M; K), k \in K :$
$i_V \varpi = \sum_{k=1}^r (-1)^{k-1} \sum_{\{\alpha_1...\alpha_r\}} V^{\alpha_k} \varpi_{\alpha_1...\alpha_r} dx^{\alpha_1} \wedge ... \wedge \widehat{dx^{\alpha_k}} ... dx^{\alpha_r}$
where $\hat{}$ is for a variable that shall be omitted.
$i_V(\varpi \wedge \pi) = (i_V \varpi) \wedge \pi + (-1)^{\deg \varpi} \varpi \wedge (i_V \pi)$
$i_V \circ i_V = 0$
$i_{fV} = f i_V$
$i_{[V,W]} \varpi = (i_W \varpi) V - (i_V \varpi) W$
$i_V \varpi(kV) = 0$
$\varpi \in \mathfrak{X}(\wedge_2 TM^*) : (i_V \varpi) W = \varpi(V, W) = -\varpi(W, V) = -(i_W \varpi) V$

16.3.3 Exterior differential

The exterior differential is an operation which is specific both to differential geometry and r-forms. But, as functions are 0 forms, it extends to functions on a manifold.

Definition 1462 *On a m dimensional manifold M the* **exterior differential** *is the operator :* $d : \mathfrak{X}_1 (\wedge_r TM^*) \to \mathfrak{X}_0 (\wedge_{r+1} TM^*)$ *defined in a holonomic basis by :*

$$d \left(\sum_{\{\alpha_1...\alpha_r\}} \varpi_{\alpha_1...\alpha_r} dx^{\alpha_1} \wedge dx^{\alpha_2} \wedge ... \wedge dx^{\alpha_r} \right)$$
$$= \sum_{\{\alpha_1...\alpha_r\}} \sum_{\beta=1}^m \partial_\beta \varpi_{\alpha_1...\alpha_r} dx^\beta \wedge dx^{\alpha_1} \wedge dx^{\alpha_2} \wedge ... \wedge dx^{\alpha_r}$$

Even if this definition is based on components one can show that d is the unique "natural" operator $\Lambda_r TM^* \to \Lambda_{r+1} TM^*$. So the result does not depend on the choice of a chart.

For :
$f \in C_2 (M; K) : df = \sum_{\alpha \in A} (\partial_\alpha f) dx^\alpha$ so df(p) = f'(p)$\in \mathcal{L} (T_p M; K)$
$\varpi \in \Lambda_1 TM^*$:
$d(\sum_{\alpha \in A} \varpi_\alpha dx^\alpha) = \sum_{\alpha < \beta}(\partial_\beta \varpi_\alpha - \partial_\alpha \varpi_\beta)(dx^\beta \wedge dx^\alpha) = \sum_{\alpha < \beta}(\partial_\beta \varpi_\alpha)(dx^\beta \otimes dx^\alpha - dx^\alpha \otimes dx^\beta)$
$\varpi \in \Lambda_r TM^*$:
$d\varpi = \sum_{\{\alpha_1...\alpha_{r+1}\}} \left(\sum_{k=1}^{r+1}(-1)^{k-1}\partial_{\alpha_k} \varpi_{\alpha_1.\widehat{\alpha_k}...\alpha_{r+1}} \right) dx^{\alpha_1} \wedge dx^{\alpha_2} \wedge ... \wedge dx^{\alpha_{r+1}}$

Theorem 1463 *(Kolar p.65) On a m dimensional manifold M the exterior differential is a linear operator :* $d \in \mathcal{L} (\mathfrak{X}_1 (\wedge_r TM^*) ; \mathfrak{X}_0 (\wedge_{r+1} TM^*))$ *which has the following properties :*
i) it is nilpotent : $d^2 = 0$
ii) it commutes with the push forward by any differential map
iii) it commutes with the Lie derivative $\pounds_V$ *for any vector field V*

So :
$\forall \lambda, \mu \in \mathfrak{X} (\Lambda_r TM^*), \pi \in \mathfrak{X} (\Lambda_s TM^*), \forall k; k' \in K, V \in \mathfrak{X} (TM), f \in C_2 (M; K)$
$d (k\lambda + k'\mu) = kd\lambda + k'd\mu$
$d (d\varpi) = 0$
$f^* \circ d = d \circ f^*$
$\pounds_V \circ d = d \circ \pounds_V$
$\forall \varpi \in \mathfrak{X} (\Lambda_m TM^*) : d\varpi = 0$

Theorem 1464 *On a m dimensional manifold M the exterior differential d, the Lie derivative along a vector field V and the interior product are linked in the formula :* $\forall \varpi \in \mathfrak{X} (\Lambda_r TM^*), V \in \mathfrak{X} (TM) : \pounds_V \varpi = i_V d\varpi + d \circ i_V \varpi$

This is an alternate definition of the exterior differential.

Theorem 1465 $\forall \lambda \in \mathfrak{X} (\Lambda_r TM^*), \mu \in \mathfrak{X} (\Lambda_s TM^*) : d (\lambda \wedge \mu) = (d\lambda) \wedge \mu + (-1)^{\deg \lambda} \lambda \wedge (d\mu)$

so for $f \in C_2 (M; K) : d (f\varpi) = (df) \wedge \varpi + f d\varpi$

Theorem 1466 *Value for vector fields :*
$\forall \varpi \in \mathfrak{X} (\Lambda_r TM^*), V_1, ..., V_{r+1} \in \mathfrak{X} (TM) :$
$d\varpi(V_1, V_2, ...V_{r+1})$
$= \sum_{i=1}^{r+1}(-1)^i V_i \left(\varpi(V_1, ...\widehat{V_i}...V_{r+1}) \right) + \sum_{\{i,j\}}(-1)^{i+j} \varpi([V_i, V_j], V_1, ...\widehat{V_i}, ...\widehat{V_j}...V_{r+1})$

359

here V_i is the differential operator linked to V_i acting on the function $\varpi(V_1, ... \widehat{V_i}...V_{r+1})$

Which gives : $d\varpi(V,W) = (i_W \varpi) V - (i_V \varpi) W - i_{[V,W]} \varpi$

and if $\varpi \in X(\Lambda_1 TM^*) : d\varpi(V,W) = \pounds_V (i_V \varpi) - \pounds_W (i_V \varpi) - i_{[V,W]} \varpi$

If ϖ is a r-form valued in a fixed vector space, the exterior differential is computed by :

$\varpi = \sum_{\{\alpha_1...\alpha_r\}} \sum_i \varpi^i_{\alpha_1...\alpha_r} e_i \otimes dx^{\alpha_1} \wedge dx^{\alpha_2} \wedge ... \wedge dx^{\alpha_r}$

$\rightarrow d\varpi = \sum_{\{\alpha_1...\alpha_r\}} \sum_{\beta=1}^m \sum_i \partial_\beta \varpi^i_{\alpha_1...\alpha_r} e_i \otimes dx^\beta \wedge dx^{\alpha_1} \wedge dx^{\alpha_2} \wedge ... \wedge dx^{\alpha_r}$

16.3.4 Poincaré's lemna

Definition 1467 *On a manifold M :*

a closed form is a field of r-form $\varpi \in X(\Lambda_r TM^*)$ such that $d\varpi = 0$

an exact form is a field of r-form $\varpi \in X(\Lambda_r TM^*)$ such that there is $\lambda \in X(\Lambda_{r-1} TM^*)$ with $\varpi = d\lambda$

An exact form is closed, the lemna of Poincaré gives a converse.

Theorem 1468 *Poincaré's lemna : A closed differential form is locally exact.*

Which means that : If $\varpi \in X(\Lambda_r TM^*)$ such that $d\varpi = 0$ then, for any $p \in M$, there is a neighborhood n(p) and $\lambda \in X(\Lambda_{r-1} TM^*)$ such that $\varpi = d\lambda$ in n(p).

The solution is not unique : $\lambda + d\mu$ is still a solution, whatever μ. The study of the subsets of closed forms which differ only by an exact form is the main topic of cohomology (see below).

If M is an open simply connected subset of a real finite dimensional affine space, $\varpi \in \Lambda_1 TM^*$ of class q, such that $d\varpi = 0$, then there is a function $f \in C_{q+1}(M;\mathbb{R})$ such that $df = \varpi$

If $M = \mathbb{R}^n : \varpi = \sum_{\alpha=1}^n a_\alpha(x) dx^\alpha, d\varpi = 0 \Rightarrow \lambda(x) = \sum_{\alpha=1}^n x^\alpha \int_0^1 a_\alpha(tx) dt$: $d\lambda = \varpi$

16.4 Covariant derivative

The general theory of connections is seen in the Fiber bundle part. We will limit here to the theory of covariant derivation, which is part of the story, but simpler and very useful for many practical purposes.

In this section the manifold M is a m dimensional smooth real manifold with atlas $(O_i, \varphi_i)_{i \in I}$

The theory of affine connection and covariant derivative can be extended to Banach manifolds of infinite dimension (see Lewis).

16.4.1 Covariant derivative

A covariant derivative is a derivative for tensor fields, which meets the requirements for the transportation of tensors (see Lie Derivatives). On the tensor bundle of a manifold it is identical to an affine connection, which is a more general breed of connections (see Fiber Bundles).

Definition

Definition 1469 *A **covariant derivative** on a manifold M is a linear operator $\nabla \in \mathcal{L}\left(\mathfrak{X}\left(TM\right);D\right)$ from the space of vector fields to the space D of derivations on the tensorial bundle of M, such that for every $V \in \mathfrak{X}\left(TM\right)$:*
i) $\nabla_V \in \mathcal{L}\left(\mathfrak{X}\left(\otimes_s^r TM\right);\mathfrak{X}\left(\otimes_{s+1}^r TM\right)\right)$
ii) ∇_V follows the Leibnitz rule with respect to the tensorial product
iii) ∇_V commutes with the trace operator

Definition 1470 *An **affine connection** on a manifold M over the field K is a bilinear operator $\nabla \in \mathcal{L}^2\left(\mathfrak{X}\left(TM\right);\mathfrak{X}\left(TM\right)\right)$ such that :*
$\forall f \in C_1\left(M;K\right) : \nabla_{fX}Y = f\nabla_X Y$
$\nabla_X(fY) = f\nabla_X Y + (i_X df)Y$

In a holonomic basis of M the coefficients of ∇ are the Christoffel symbols of the connection : $\Gamma_{\beta\eta}^\alpha\left(p\right)$

Theorem 1471 *An affine connection defines uniquely a covariant derivative and conversely a covariant derivative defines an affine connection.*

Proof. i) According to the rules above, a covariant derivative is defined if we have the derivatives of $\partial_\alpha, dx^\alpha$ which are tensor fields. So let us denote :
$\left(\nabla\partial_\alpha\right)\left(p\right) = \sum_{\gamma,\eta=1}^m X_{\eta\alpha}^\gamma\left(p\right)dx^\eta \otimes \partial_\gamma$
$\nabla dx^\alpha = \sum_{\gamma,\eta=1}^m Y_{\eta\gamma}^\alpha\left(p\right)dx^\eta \otimes dx^\gamma$
By definition : $\left(dx^\alpha\left(\partial_\beta\right)\right) = \delta_\beta^\alpha \Rightarrow \nabla\left(Tr\left(dx^\alpha\left(\partial_\beta\right)\right)\right) = Tr\left(\left(\nabla dx^\alpha\right)\otimes\partial_\beta\right) + Tr\left(dx^\alpha\otimes\nabla\partial_\beta\right) = 0$
$Tr\left(\sum_{\eta,\gamma=1}^m Y_{\eta\gamma}^\alpha dx^\eta\otimes dx^\gamma\otimes\partial_\beta\right) = -Tr\left(dx^\alpha\otimes\sum_{\eta,\gamma=1}^m X_{\eta\beta}^\gamma dx^\eta\otimes\partial_\gamma\right)$
$\sum_{\eta,\gamma=1}^m Y_{\eta\gamma}^\alpha dx^\eta\left(dx^\gamma\left(\partial_\beta\right)\right) = -\sum_{\eta,\gamma=1}^m X_{\eta\beta}^\gamma dx^\eta\left(dx^\alpha\left(\partial_\gamma\right)\right)$
$\sum_{\eta,\gamma=1}^m Y_{\eta\beta}^\alpha dx^\eta = -\sum_{\eta,\gamma=1}^m X_{\eta\beta}^\alpha dx^\eta \Leftrightarrow Y_{\eta\beta}^\alpha = -X_{\eta\beta}^\alpha$
So the derivation is fully defined by the value of the Christofell coefficients $\Gamma_{\beta\eta}^\alpha\left(p\right)$ scalar functions for a holonomic basis and we have:
$\nabla\partial_\alpha = \sum_{\beta,\gamma=1}^m \Gamma_{\beta\alpha}^\gamma dx^\beta\otimes\partial_\gamma$
$\nabla dx^\alpha = -\sum_{\beta,\gamma=1}^m \Gamma_{\beta\gamma}^\alpha dx^\beta\otimes dx^\gamma$
ii) Conversely an affine connection with Christofell coefficients $\Gamma_{\beta\eta}^\alpha\left(p\right)$ defines a unique covariant connection (Kobayashi I p.143). ∎

Christoffel symbols in a change of charts
A covariant derivative is not unique : it depends on the coefficients Γ which have been computed in a chart. However a given covariant derivative ∇ is a geometric object, which is independant of a the choice of a basis. In a change of charts *the Christoffel coefficients are not tensors*, and change according to specific rules.

Theorem 1472 *The Christoffel symbols in the new basis are :*
$\widehat{\Gamma}_{\beta\gamma}^\alpha = \sum_{\lambda\mu}\left[J^{-1}\right]_\beta^\mu\left[J^{-1}\right]_\gamma^\lambda\left(\Gamma_{\mu\lambda}^\nu\left[J\right]_\nu^\alpha - \partial_\mu\left[J\right]_\lambda^\alpha\right)$ with the Jacobian $J = \left[\frac{\partial y^\alpha}{\partial x^\beta}\right]$

Proof. Coordinates in the old chart : $x = \varphi_i(p)$

Coordinates in the new chart : $y = \psi_i(p)$

Old holonomic basis :

$\partial x_\alpha = \varphi_i'(p)^{-1} e_\alpha,$

$dx^\alpha = \varphi_i'(x)^t e^\alpha$ with $dx^\alpha(\partial x_\beta) = \delta_\beta^\alpha$

New holonomic basis :

$\partial y_\alpha = \psi_i'(p)^{-1} e_\alpha = \sum_\beta \left[J^{-1}\right]_\alpha^\beta \partial x_\beta$

$dy^\alpha = \psi_i'(y)^* e^\alpha = \sum_\beta [J]_\alpha^\beta dx_\beta$ with $dy^\alpha(\partial y_\beta) = \delta_\beta^\alpha$

Transition map:

$\alpha = 1..n : y^\alpha = F^\alpha(x^1,...x^n) \Leftrightarrow F(x) = \psi_i \circ \varphi_i^{-1}(x)$

Jacobian : $J = [F'(x)] = \left[J_\beta^\alpha\right] = \left[\frac{\partial F^\alpha}{\partial x^\beta}\right]_{n \times n} \simeq \left[\frac{\partial y^\alpha}{\partial x^\beta}\right]$

$V = \sum_\alpha V^\alpha \partial x_\alpha = \sum_\alpha \widehat{V}^\alpha \partial y_\alpha$ with $\widehat{V}^\alpha = \sum_\beta J_\beta^\alpha V^\beta \simeq \sum_\beta \frac{\partial y^\alpha}{\partial x^\beta} V^\beta$

If we want the same derivative with both charts, we need for any vector field :

$\nabla V = \sum_{\alpha,\beta=1}^m \left(\frac{\partial}{\partial x^\beta} V^\alpha + \Gamma_{\beta\gamma}^\alpha V^\gamma\right) dx^\beta \otimes \partial x_\alpha = \sum_{\alpha,\beta=1}^m \left(\frac{\partial}{\partial y^\beta} \widehat{V}^\alpha + \widehat{\Gamma}_{\beta\gamma}^\alpha \widehat{V}^\gamma\right) dy^\beta \otimes \partial y_\alpha$

∇V is a (1,1) tensor, whose components change according to :

$T = \sum_{\alpha\beta} T_\beta^\alpha \partial x_\alpha \otimes dx^\beta = \sum_{\alpha\beta} \widehat{T}_\beta^\alpha \partial y_\alpha \otimes dy^\beta$ with $\widehat{T}_\beta^\alpha = \sum_{\lambda\mu} T_\mu^\lambda [J]_\lambda^\alpha \left[J^{-1}\right]_\beta^\mu$

So : $\frac{\partial}{\partial y^\beta} \widehat{V}^\alpha + \widehat{\Gamma}_{\beta\gamma}^\alpha \widehat{V}^\gamma = \sum_{\lambda\mu} \left(\frac{\partial}{\partial x^\mu} V^\lambda + \Gamma_{\mu\nu}^\lambda V^\nu\right) [J]_\lambda^\alpha \left[J^{-1}\right]_\beta^\mu$

which gives : $\widehat{\Gamma}_{\beta\gamma}^\alpha = \left[J^{-1}\right]_\beta^\mu \left[J^{-1}\right]_\gamma^\lambda \left(\Gamma_{\mu\lambda}^\nu [J]_\nu^\alpha - \partial_\mu [J]_\lambda^\alpha\right)$ $\blacksquare$

Properties

$\forall V, W \in \mathfrak{X}(TM), \forall S, T \in \mathfrak{X}(\otimes_s^r TM), k, k' \in K, \forall f \in C_1(M;K)$

$\nabla_V \in \mathcal{L}\left(\mathfrak{X}(\otimes_s^r TM); \mathfrak{X}(\otimes_{s+1}^r TM)\right)$

$\nabla_V (kS + k'T) = k\nabla_V S + k'\nabla_V T$

$\nabla_V (S \otimes T) = (\nabla_V S) \otimes T + S \otimes (\nabla_V T)$

$\nabla_{fV} W = f\nabla_V W$

$\nabla_V (fW) = f\nabla_V W + (i_V df)W$

$\nabla f = df \in \mathfrak{X}(\otimes_1^0 TM)$

$\nabla_V (Tr(T)) = Tr(\nabla_V T)$

Coordinate expressions in a holonomic basis:

for a vector field : $V = \sum_{\alpha=1}^m V^\alpha \partial_\alpha$:

$\nabla V = \sum_{\alpha,\beta=1}^m \left(\partial_\beta V^\alpha + \Gamma_{\beta\gamma}^\alpha V^\gamma\right) dx^\beta \otimes \partial_\alpha$

for a 1-form : $\varpi = \sum_{\alpha=1}^m \varpi_\alpha dx^\alpha$:

$\nabla \varpi = \sum_{\alpha,\beta=1}^m \left(\partial_\beta \varpi_\alpha - \Gamma_{\beta\alpha}^\gamma \varpi_\gamma\right) dx^\beta \otimes dx^\alpha$

for a mix tensor :

$T(p) = \sum_{\alpha_1...\alpha_r} \sum_{\beta_1....\beta_s} T_{\beta_1...\beta_s}^{\alpha_1...\alpha_r}(p) \partial x_{\alpha_1} \otimes .. \otimes \partial x_{\alpha_r} \otimes dx^{\beta_1} \otimes ... \otimes dx^{\beta_s}$

$\nabla T(p) = \sum_{\alpha_1...\alpha_r} \sum_{\beta_1....\beta_s} \sum_\gamma \widehat{T}_{\gamma\beta_1...\beta_s}^{\alpha_1...\alpha_r} \partial x_{\alpha_1} \otimes .. \otimes \partial x_{\alpha_r} \otimes dx^\gamma \otimes dx^{\beta_1} \otimes ... \otimes dx^{\beta_s}$

$\widehat{T}_{\gamma\beta_1...\beta_s}^{\alpha_1...\alpha_r} = \partial_\gamma T_{\beta_1...\beta_s}^{\alpha_1...\alpha_r} + \sum_{k=1}^r \Gamma_{\gamma\eta}^{\alpha_k} T_{\beta_1...\beta_s}^{\alpha_1 . \alpha_{k-1}\eta\alpha_{k+1}..\alpha_r} - \sum_{k=1}^s \Gamma_{\gamma\beta_k}^\eta T_{\beta_1..\beta_{k-1}\eta\beta_{k+1}..\beta_s}^{\alpha_1...\alpha_r}$

362

16.4.2 Exterior covariant derivative

The covariant derivative of a r-form is not an antisymmetric tensor. In order to get an operator working on r-forms, one defines the exterior covariant derivative which applies to r-forms on M, *valued in the tangent bundle*.

Definition 1473 *The **exterior covariant derivative** associated to the covariant derivative ∇, is the linear map :*
$\nabla_e \in \mathcal{L}\left(\Lambda_r\left(M;TM\right);\Lambda_{r+1}\left(M;TM\right)\right)$
with the condition : $\forall X_0, X_1, ..X_r \in \mathfrak{X}\left(TM\right), \varpi \in \mathfrak{X}\left(\Lambda_r TM^*\right)$
$\left(\nabla_e \varpi\right)\left(X_0, X_1, ...X_r\right)$
$= \sum_{i=0}^{r}(-1)^i \nabla_{X_i}\varpi(X_0, X_1, ...\widehat{X_i}...X_r) + \sum_{\{i,j\}}(-1)^{i+j}\varpi([X_i, X_j], X_0, X_1, ...\widehat{X_i}, ...\widehat{X_j}...X_r)$

This formula is similar to the one for the exterior differential (∇ replacing £).

Which leads to the formula :
$\nabla_e \varpi = \sum_\beta \left(d\varpi^\beta + \left(\sum_\gamma \left(\sum_\alpha \Gamma_{\alpha\gamma}^\beta dx^\alpha\right)\wedge\varpi^\gamma\right)\right)\partial x_\beta$

Proof. Such a form reads : ∎
$\varpi = \sum_{\beta\{\alpha_1...\alpha_r\}}\sum_\beta \varpi_{\alpha_1...\alpha_r}^\beta dx^{\alpha_1}\wedge dx^{\alpha_2}\wedge ... \wedge dx^{\alpha_r}\otimes\partial x_\beta \in \Lambda_r\left(M;TM\right)$
Proof. Let us denote : $\sum_\beta \varpi^\beta(X_0, ...\widehat{X_i}...X_r)\partial x_\beta = \sum_\beta \Omega_i^\beta \partial x_\beta$
From the exterior differential formulas , β fixed:
$\left(d\varpi^\beta\right)\left(X_0, X_1, ...X_r\right)$
$= \sum_{i=0}^{r}(-1)^i X_i^\alpha \partial_\alpha \left(\varpi^\beta(X_0, ...\widehat{X_i}...X_r)\right) + \sum_{\{i,j\}}(-1)^{i+j}\varpi^\beta([X_i, X_j], X_0, ...\widehat{X_i}, ...\widehat{X_j}...X_r)$
So :
$\nabla_e \varpi(X_0, X_1, ...X_r)$
$= \sum_\beta \left(d\varpi^\beta\right)\left(X_0, X_1, ...X_r\right)\partial x_\beta + \sum_{i=0}^{r}(-1)^i\left(\nabla_{X_i}\left(\sum_\beta \Omega_i^\beta \partial x_\beta\right) - \sum_{\alpha\beta}X_i^\alpha \partial_\alpha\left(\Omega_i^\beta\right)\partial x_\beta\right)$
$= \sum_\beta \left(d\varpi^\beta\right)\left(X_0, X_1, ...X_r\right)\partial x_\beta + \sum_{i=0}^{r}(-1)^i\left(\sum_{\alpha\beta\gamma}\left(\partial_\alpha\Omega_i^\beta + \Gamma_{\alpha\gamma}^\beta \Omega_i^\gamma\right)X_i^\alpha\partial x_\beta - \sum_{\alpha\beta}X_i^\alpha \partial_\alpha\left(\Omega_i^\beta\right)\partial x_\beta\right)$
$= \sum_\beta \left(d\varpi^\beta\right)\left(X_0, X_1, ...X_r\right)\partial x_\beta + \sum_{i=0}^{r}(-1)^i\left(\sum_{\alpha\beta\gamma}\Gamma_{\alpha\gamma}^\beta \Omega_i^\gamma X_i^\alpha\partial x_\beta\right)$
$\Omega_i^\gamma = \sum_{\{\lambda_0...\widehat{\lambda_i}..\lambda_{r-1}\}}\varpi_{\{\lambda_0...\widehat{\lambda_i}..\lambda_{r-1}\}}^\gamma X_0^{\lambda_0}X_1^{\lambda_1}...\widehat{X_i^{\lambda_i}}...X_r^{\lambda_r}$
$\sum_{\alpha\gamma}\Gamma_{\alpha\gamma}^\beta \sum_{i=0}^{r}(-1)^i\Omega_i^\gamma X_i^\alpha = \sum_\gamma \sum_{i=0}^{r}\sum_{\{\lambda_0....\lambda_{r-1}\}}\Gamma_{\lambda_i\gamma}^\beta \varpi_{\{\lambda_0....\lambda_{r-1}\}}^\gamma X_0^{\lambda_0}X_1^{\lambda_1}...X_i^{\lambda_i}...X_r^{\lambda_r}$
$= \left(\sum_\gamma\left(\sum_\alpha \Gamma_{\alpha\gamma}^\beta dx^\alpha\right)\wedge\varpi^\gamma\right)\left(X_0, ...X_r\right)$
$\nabla_e \varpi(X_0, X_1, ...X_r) = \sum_\beta \left(d\varpi^\beta + \left(\sum_\gamma\left(\sum_\alpha \Gamma_{\alpha\gamma}^\beta dx^\alpha\right)\wedge\varpi^\gamma\right)\right)\left(X_0, ...X_r\right)\partial x_\beta$
∎

A vector field can be considered as a 0-form valued in TM, and $\forall X \in \mathfrak{X}\left(TM\right) : \nabla_e X = \nabla X$ (we have the usual covariant derivative of a vector field on M)

Theorem 1474 *Exterior product:*
$\forall \varpi_r \in \Lambda_r\left(M;TM\right), \varpi_s \in \Lambda_s\left(M;TM\right) : \nabla_e\left(\varpi_r \wedge \varpi_s\right) = \left(\nabla_e\varpi_r\right)\wedge\varpi_s + (-1)^r \varpi_r \wedge \nabla_e\varpi_s$

So the formula is the same as for the exterior differential d.

Theorem 1475 *Pull-back,push forward (Kolar p.112) The exterior covariant derivative commutes with the pull back of forms :*

$$\forall f \in C_2\left(N;M\right), \varpi \in \mathfrak{X}\left(\Lambda_r T N^*\right) : \nabla_e\left(f^*\varpi\right) = f^*\left(\nabla_e\varpi\right)$$

16.4.3   Curvature

Definition 1476 *The **Riemann curvature** of a covariant connection ∇ is the multilinear map :*

$$R : \left(\mathfrak{X}\left(TM\right)\right)^3 \to \mathfrak{X}\left(M\right) :: R(X,Y,Z) = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z$$

It is also called the Riemann tensor or curvature tensor. As there are many objects called curvature we opt for Riemann curvature.

The name curvature comes from the following : for a vector field V :

$$R(\partial_\alpha, \partial_\beta, V) = \nabla_{\partial_\alpha}\nabla_{\partial_\beta}V - \nabla_{\partial_\beta}\nabla_{\partial_\alpha}V - \nabla_{[\partial_\alpha,\partial_\beta]}V = \left(\nabla_{\partial_\alpha}\nabla_{\partial_\beta} - \nabla_{\partial_\beta}\nabla_{\partial_\alpha}\right)V$$

because $[\partial_\alpha, \partial_\beta] = 0$

So R is a measure of the obstruction of the covariant derivative to be commutative : $\nabla_{\partial_\alpha}\nabla_{\partial_\beta} - \nabla_{\partial_\beta}\nabla_{\partial_\alpha} \neq 0$

Theorem 1477 *The Riemann curvature is a tensor valued in the tangent bundle :* $R \in \mathfrak{X}\left(\Lambda_2 T M^* \overset{1}{\underset{1}{\otimes}} TM\right)$

$$R = \sum_{\{\gamma\eta\}}\sum_{\alpha\beta} R^\alpha_{\gamma\eta\beta}dx^\gamma \wedge dx^\eta \otimes dx^\beta \otimes \partial x_\alpha \text{ with}$$
$$R^\varepsilon_{\alpha\beta\gamma} = \partial_\alpha\Gamma^\varepsilon_{\beta\gamma} - \partial_\beta\Gamma^\varepsilon_{\alpha\gamma} + \Gamma^\varepsilon_{\alpha\eta}\Gamma^\eta_{\beta\gamma} - \Gamma^\varepsilon_{\beta\eta}\Gamma^\eta_{\alpha\gamma}$$

Proof. $R(X,Y,Z) = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z$

$$= \nabla_X\left(\left(\partial_\alpha Z^\varepsilon + \Gamma^\varepsilon_{\alpha\gamma}Z^\gamma\right)Y^\alpha\partial_\varepsilon\right) - \nabla_Y\left(\left(\partial_\alpha Z^\varepsilon + \Gamma^\varepsilon_{\alpha\gamma}Z^\gamma\right)X^\alpha\partial_\varepsilon\right)$$
$$- \left(\partial_\alpha Z^\varepsilon + \Gamma^\varepsilon_{\alpha\gamma}Z^\gamma\right)\left(\left(X^\eta\partial_\eta Y^\alpha - Y^\eta\partial_\eta X^\alpha\right)\right)$$
$$= \left(\partial_\beta\left(\left(\partial_\alpha Z^\varepsilon + \Gamma^\varepsilon_{\alpha\gamma}Z^\gamma\right)Y^\alpha\right) + \Gamma^\varepsilon_{\beta\eta}\left(\left(\partial_\alpha Z^\eta + \Gamma^\eta_{\alpha\gamma}Z^\gamma\right)Y^\alpha\right)\right)X^\beta\partial_\varepsilon$$
$$- \left(\partial_\beta\left(\left(\partial_\alpha Z^\varepsilon + \Gamma^\varepsilon_{\alpha\gamma}Z^\gamma\right)X^\alpha\right) - \Gamma^\varepsilon_{\beta\eta}\nabla_Y\left(\left(\partial_\alpha Z^\eta + \Gamma^\eta_{\alpha\gamma}Z^\gamma\right)X^\alpha\right)\right)Y^\beta\partial_\varepsilon$$
$$- \left(\left(\partial_\alpha Z^\varepsilon\right)\left(\left(X^\eta\partial_\eta Y^\alpha - Y^\eta\partial_\eta X^\alpha\right)\right) + \Gamma^\varepsilon_{\alpha\gamma}Z^\gamma\left(\left(X^\eta\partial_\eta Y^\alpha - Y^\eta\partial_\eta X^\alpha\right)\right)\right)\partial_\varepsilon$$

The component of ∂_ε is:

$$= \left(\partial_\beta\left(\partial_\alpha Z^\varepsilon + \Gamma^\varepsilon_{\alpha\gamma}Z^\gamma\right)\right)X^\beta Y^\alpha + \left(\partial_\alpha Z^\varepsilon + \Gamma^\varepsilon_{\alpha\gamma}Z^\gamma\right)X^\beta\partial_\beta Y^\alpha + \Gamma^\varepsilon_{\beta\eta}\left(\partial_\alpha Z^\eta\right)Y^\alpha X^\beta +$$
$$\Gamma^\varepsilon_{\beta\eta}\Gamma^\eta_{\alpha\gamma}Z^\gamma X^\beta Y^\alpha$$
$$- \left(\partial_\beta\left(\partial_\alpha Z^\varepsilon + \Gamma^\varepsilon_{\alpha\gamma}Z^\gamma\right)\right)X^\alpha Y^\beta - \left(\partial_\alpha Z^\varepsilon + \Gamma^\varepsilon_{\alpha\gamma}Z^\gamma\right)Y^\beta\partial_\beta X^\alpha - \Gamma^\varepsilon_{\beta\eta}\partial_\alpha Z^\eta X^\alpha Y^\beta -$$
$$\Gamma^\varepsilon_{\beta\eta}\Gamma^\eta_{\alpha\gamma}Z^\gamma X^\alpha Y^\beta$$
$$- \left(\partial_\alpha Z^\varepsilon\right)X^\eta\left(\partial_\eta Y^\alpha\right) + \left(\partial_\alpha Z^\varepsilon\right)Y^\eta\left(\partial_\eta X^\alpha\right) - \Gamma^\varepsilon_{\alpha\gamma}Z^\gamma X^\eta\left(\partial_\eta Y^\alpha\right) + \Gamma^\varepsilon_{\alpha\gamma}Z^\gamma Y^\eta\left(\partial_\eta X^\alpha\right)$$
$$= \left(\partial_\beta\partial_\alpha Z^\varepsilon\right)X^\beta Y^\alpha + \left(\partial_\beta\Gamma^\varepsilon_{\alpha\gamma}\right)X^\beta Z^\gamma Y^\alpha + \Gamma^\varepsilon_{\alpha\gamma}\left(\partial_\beta Z^\gamma\right)X^\beta Y^\alpha + \left(\partial_\alpha Z^\varepsilon\right)\left(\partial_\beta Y^\alpha\right)X^\beta$$
$$+ \Gamma^\varepsilon_{\alpha\gamma}Z^\gamma\left(\partial_\beta Y^\alpha\right)X^\beta + \Gamma^\varepsilon_{\beta\eta}\left(\partial_\alpha Z^\eta\right)Y^\alpha X^\beta + \Gamma^\varepsilon_{\beta\eta}\Gamma^\eta_{\alpha\gamma}Z^\gamma Y^\alpha X^\beta$$
$$- \left(\partial_\beta\partial_\alpha Z^\varepsilon\right)X^\alpha Y^\beta - \left(\partial_\beta\Gamma^\varepsilon_{\alpha\gamma}\right)Z^\gamma X^\alpha Y^\beta - \Gamma^\varepsilon_{\alpha\gamma}\left(\partial_\beta Z^\gamma\right)X^\alpha Y^\beta - \left(\partial_\alpha Z^\varepsilon\right)\left(\partial_\beta X^\alpha\right)Y^\beta$$
$$+ \Gamma^\varepsilon_{\alpha\gamma}Z^\gamma\left(\partial_\beta X^\alpha\right)Y^\beta - \Gamma^\varepsilon_{\beta\eta}\left(\partial_\alpha Z^\eta\right)X^\alpha Y^\beta - \Gamma^\varepsilon_{\beta\eta}\Gamma^\eta_{\alpha\gamma}Z^\gamma X^\alpha Y^\beta$$
$$- \left(\partial_\alpha Z^\varepsilon\right)X^\eta\partial_\eta Y^\alpha + \left(\partial_\alpha Z^\varepsilon\right)Y^\eta\left(\partial_\eta X^\alpha\right) - \Gamma^\varepsilon_{\alpha\gamma}Z^\gamma X^\eta\left(\partial_\eta Y^\alpha\right) + \Gamma^\varepsilon_{\alpha\gamma}Z^\gamma Y^\eta\left(\partial_\eta X^\alpha\right)$$
$$= \left(\partial_\beta\partial_\alpha Z^\varepsilon\right)X^\beta Y^\alpha - \left(\partial_\beta\partial_\alpha Z^\varepsilon\right)X^\alpha Y^\beta$$

$+\left(\partial_\alpha Z^\varepsilon\right)\left(\partial_\beta Y^\alpha\right)X^\beta - \left(\partial_\alpha Z^\varepsilon\right)\left(\partial_\beta X^\alpha\right)Y^\beta - \left(\partial_\alpha Z^\varepsilon\right)X^\eta\left(\partial_\eta Y^\alpha\right)+\left(\partial_\alpha Z^\varepsilon\right)Y^\eta\left(\partial_\eta X^\alpha\right)$

$+\Gamma^\varepsilon_{\alpha\gamma}\left(\partial_\beta Z^\gamma\right)X^\beta Y^\alpha+\Gamma^\varepsilon_{\beta\eta}\left(\partial_\alpha Z^\eta\right)Y^\alpha X^\beta-\Gamma^\varepsilon_{\beta\eta}\left(\partial_\alpha Z^\eta\right)X^\alpha Y^\beta-\Gamma^\varepsilon_{\alpha\gamma}\left(\partial_\beta Z^\gamma\right)X^\alpha Y^\beta$

$+\Gamma^\varepsilon_{\alpha\gamma}Z^\gamma\left(\partial_\beta Y^\alpha\right)X^\beta-\Gamma^\varepsilon_{\alpha\gamma}Z^\gamma X^\eta\left(\partial_\eta Y^\alpha\right)+\Gamma^\varepsilon_{\alpha\gamma}Z^\gamma\left(\partial_\beta X^\alpha\right)Y^\beta+\Gamma^\varepsilon_{\alpha\gamma}Z^\gamma Y^\eta\left(\partial_\eta X^\alpha\right)$

$+\left(\partial_\beta\Gamma^\varepsilon_{\alpha\gamma}\right)X^\beta Z^\gamma Y^\alpha+\Gamma^\varepsilon_{\beta\eta}\Gamma^\eta_{\alpha\gamma}Z^\gamma Y^\alpha X^\beta-\left(\partial_\beta\Gamma^\varepsilon_{\alpha\gamma}\right)Z^\gamma X^\alpha Y^\beta-\Gamma^\varepsilon_{\beta\eta}\Gamma^\eta_{\alpha\gamma}Z^\gamma X^\alpha Y^\beta$

$=\left(\partial_\alpha\partial_\beta Z^\varepsilon\right)X^\alpha Y^\beta-\left(\partial_\beta\partial_\alpha Z^\varepsilon\right)X^\alpha Y^\beta+\left(\partial_\alpha Z^\varepsilon\right)\left(\left(\partial_\beta Y^\alpha\right)X^\beta-X^\beta\left(\partial_\beta Y^\alpha\right)+Y^\beta\left(\partial_\beta X^\alpha\right)-\left(\partial_\beta X^\alpha\right)Y^\beta\right)$

$+\Gamma^\varepsilon_{\beta\eta}X^\alpha Y^\beta\left(\partial_\alpha Z^\eta\right)-\Gamma^\varepsilon_{\beta\eta}X^\alpha Y^\beta\left(\partial_\alpha Z^\eta\right)+\Gamma^\varepsilon_{\alpha\eta}X^\alpha Y^\beta\left(\partial_\beta Z^\eta\right)-\Gamma^\varepsilon_{\alpha\eta}X^\alpha Y^\beta\left(\partial_\beta Z^\eta\right)$

$+\Gamma^\varepsilon_{\alpha\gamma}Z^\gamma\left(\left(\partial_\beta Y^\alpha\right)X^\beta-X^\beta\left(\partial_\beta Y^\alpha\right)\right)+\Gamma^\varepsilon_{\alpha\gamma}Z^\gamma\left(\left(\partial_\beta X^\alpha\right)Y^\beta+Y^\beta\left(\partial_\beta X^\alpha\right)\right)$

$+\left(\partial_\beta\Gamma^\varepsilon_{\alpha\gamma}\right)X^\beta Z^\gamma Y^\alpha-\left(\partial_\beta\Gamma^\varepsilon_{\alpha\gamma}\right)Z^\gamma X^\alpha Y^\beta+\Gamma^\varepsilon_{\beta\eta}\Gamma^\eta_{\alpha\gamma}Z^\gamma Y^\alpha X^\beta-\Gamma^\varepsilon_{\beta\eta}\Gamma^\eta_{\alpha\gamma}Z^\gamma X^\alpha Y^\beta$

$=\left(\partial_\beta\Gamma^\varepsilon_{\alpha\gamma}\right)X^\beta Y^\alpha Z^\gamma-\left(\partial_\beta\Gamma^\varepsilon_{\alpha\gamma}\right)X^\alpha Y^\beta Z^\gamma+\Gamma^\varepsilon_{\beta\eta}\Gamma^\eta_{\alpha\gamma}X^\beta Y^\alpha Z^\gamma-\Gamma^\varepsilon_{\beta\eta}\Gamma^\eta_{\alpha\gamma}X^\alpha Y^\beta Z^\gamma$

$R(X,Y,Z)=\left(\left(\partial_\alpha\Gamma^\varepsilon_{\beta\gamma}\right)X^\alpha Y^\beta Z^\gamma-\left(\partial_\beta\Gamma^\varepsilon_{\alpha\gamma}\right)X^\alpha Y^\beta Z^\gamma+\Gamma^\varepsilon_{\alpha\eta}\Gamma^\eta_{\beta\gamma}X^\alpha Y^\beta Z^\gamma-\Gamma^\varepsilon_{\beta\eta}\Gamma^\eta_{\alpha\gamma}X^\alpha Y^\beta Z^\gamma\right)\partial_\varepsilon$

$R(X,Y,Z)=R^\varepsilon_{\alpha\beta\gamma}X^\alpha Y^\beta Z^\gamma\partial_\varepsilon$

With : $R^\varepsilon_{\alpha\beta\gamma}=\partial_\alpha\Gamma^\varepsilon_{\beta\gamma}-\partial_\beta\Gamma^\varepsilon_{\alpha\gamma}+\Gamma^\varepsilon_{\alpha\eta}\Gamma^\eta_{\beta\gamma}-\Gamma^\varepsilon_{\beta\eta}\Gamma^\eta_{\alpha\gamma}$

Clearly : $R^\varepsilon_{\alpha\beta\gamma}=-R^\varepsilon_{\beta\alpha\gamma}$ so : $R=\sum_{\{\alpha\beta\}\gamma\varepsilon}R^\varepsilon_{\alpha\beta\gamma}dx^\alpha\wedge dx^\beta\otimes dx^\gamma\otimes\partial_\varepsilon$ ∎

Theorem 1478 *For any covariant derivative ∇ and its exterior covariant derivative ∇_e :*

$\forall\varpi\in\Lambda_r\left(M;TM\right):\nabla_e\left(\nabla_e\varpi\right)=R\Lambda\varpi$ *where R is the Riemann curvature of ∇*

More precisely in a holonomic basis :

$\nabla_e\left(\nabla_e\varpi\right)=\sum_{\alpha\beta}\left(\sum_{\{\gamma\eta\}}R^\alpha_{\gamma\eta\beta}dx^\gamma\wedge dx^\eta\right)\wedge\varpi^\beta\otimes\partial x_\alpha$

Proof. $\nabla_e\varpi=\sum_\alpha\left(d\varpi^\alpha+\left(\sum_\beta\left(\sum_\alpha\Gamma^\alpha_{\gamma\beta}dx^\gamma\right)\wedge\varpi^\beta\right)\right)\partial x_\alpha=\sum_\alpha\left(d\varpi^\alpha+\sum_\beta\Omega^\alpha_\beta\wedge\varpi^\beta\right)\otimes\partial x_\alpha$

with $\Omega^\alpha_\beta=\sum_\gamma\Gamma^\alpha_{\gamma\beta}dx^\gamma$

$\nabla_e\left(\nabla_e\varpi\right)=\sum_\alpha\left(d\left(\nabla_e\varpi\right)^\alpha+\sum_\beta\Omega^\alpha_\beta\wedge\left(\nabla_e\varpi\right)^\beta\right)\otimes\partial x_\alpha$

$=\sum_\alpha\left(d\left(d\varpi^\alpha+\sum_\beta\Omega^\alpha_\beta\wedge\varpi^\beta\right)+\sum_\beta\Omega^\alpha_\beta\wedge\left(d\varpi^\beta+\sum_\gamma\Omega^\beta_\gamma\wedge\varpi^\gamma\right)\right)\otimes\partial x_\alpha$

$=\sum_{\alpha\beta}\left(d\Omega^\alpha_\beta\wedge\varpi^\beta-\Omega^\alpha_\beta\wedge d\varpi^\beta+\Omega^\alpha_\beta\wedge d\varpi^\beta+\Omega^\alpha_\beta\wedge\sum_\gamma\Omega^\beta_\gamma\wedge\varpi^\gamma\right)\otimes\partial x_\alpha$

$=\sum_{\alpha\beta}\left(d\Omega^\alpha_\beta\wedge\varpi^\beta+\sum_\gamma\Omega^\alpha_\gamma\wedge\Omega^\gamma_\beta\wedge\varpi^\beta\right)\otimes\partial x_\alpha$

$\nabla_e\left(\nabla_e\varpi\right)=\sum_{\alpha\beta}\left(d\Omega^\alpha_\beta+\sum_\gamma\Omega^\alpha_\gamma\wedge\Omega^\gamma_\beta\right)\wedge\varpi^\beta\otimes\partial x_\alpha$

$d\Omega^\alpha_\beta+\sum_\gamma\Omega^\alpha_\gamma\wedge\Omega^\gamma_\beta=d\left(\sum_\eta\Gamma^\alpha_{\eta\beta}dx^\eta\right)+\sum_\gamma\left(\sum_\varepsilon\Gamma^\alpha_{\varepsilon\gamma}dx^\varepsilon\right)\wedge\left(\sum_\eta\Gamma^\gamma_{\eta\beta}dx^\eta\right)$

$=\sum_{\eta\gamma}\partial_\gamma\Gamma^\alpha_{\eta\beta}dx^\gamma\wedge dx^\eta+\sum_{\eta\varepsilon\gamma}\Gamma^\alpha_{\varepsilon\gamma}\Gamma^\gamma_{\eta\beta}dx^\varepsilon\wedge dx^\eta$

$=\sum_{\eta\gamma}\left(\partial_\gamma\Gamma^\alpha_{\eta\beta}+\sum_\varepsilon\Gamma^\beta_{\gamma\varepsilon}\Gamma^\gamma_{\eta\beta}\right)dx^\gamma\wedge dx^\eta$ ∎

Definition 1479 *The **Ricci tensor** is the contraction of R with respect to the indexes ε,β :*

$Ric=\sum_{\alpha\gamma}Ric_{\alpha\gamma}dx^\alpha\otimes dx^\gamma$ *with* $Ric_{\alpha\gamma}=\sum_\beta R^\beta_{\alpha\beta\gamma}$

It is a symmetric tensor if R comes from the Levi-Civita connection.

Remarks :

i) The curvature tensor can be defined for any covariant derivative : there is no need of a Riemannian metric or a symmetric connection.

ii) The formula above is written in many ways in the litterature, depending on the convention used to write $\Gamma^\alpha_{\beta\gamma}$. This is why I found useful to give the complete calculations.

iii) R is always antisymmetric in the indexes α, β

16.4.4   Torsion

Definition 1480 *The **torsion** of an affine connection ∇ is the map:*
$$T : \mathfrak{X}(TM) \times \mathfrak{X}(TM) \to \mathfrak{X}(TM) :: T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]$$

It is a tensor field : $T = \sum_{\alpha,\beta,\gamma} T^\gamma_{\alpha\beta} dx^\alpha \otimes dx^\beta \otimes \partial x_\gamma \in \mathfrak{X}(\otimes^1_2 TM)$ with $T^\gamma_{\alpha\beta} = -T^\gamma_{\alpha\beta} = \Gamma^\gamma_{\alpha\beta} - \Gamma^\gamma_{\beta\alpha}$ so this is a 2 form valued in the tangent bundle : $T = \sum_{\{\alpha,\beta\}} \sum_\gamma \left(\Gamma^\gamma_{\alpha\beta} - \Gamma^\gamma_{\beta\alpha} \right) dx^\alpha \Lambda dx^\beta \otimes \partial x_\gamma \in \Lambda_2(M;TM)$

Definition 1481 *An affine connection is **torsion free** if its torsion vanishes.*

Theorem 1482 *An affine connection is torsion free iff the covariant derivative is **symmetric** : $T = 0 \Leftrightarrow \Gamma^\gamma_{\alpha\beta} = \Gamma^\gamma_{\beta\alpha}$*

Theorem 1483 *(Kobayashi I p.149) If the covariant connection ∇ is torsion free then :*
$$\forall \varpi \in \mathfrak{X}(\Lambda_r TM^*) : d\varpi = \frac{1}{r!} \sum_{\sigma \in \mathfrak{S}(r)} \epsilon(\sigma) \nabla \varpi$$

Definition 1484 *A covariant connection on a manifold whose curvature and torsion vanish is said to be **flat** (or locally affine).*

16.4.5   Parallel transport by a covariant connection

Parallel transport of a tensor

Definition 1485 *A tensor field $T \in \mathfrak{X}(\otimes^r_s TM)$ on a manifold is invariant by a covariant connection along a path $c : [a,b] \to M$ on M if its covariant derivative along the tangent, evaluated at each point of the path, is null : $\nabla_{c'(t)} T(c(t)) = 0$*

Definition 1486 *The transported tensor $\widehat{T}$ of a tensor field $T \in \mathfrak{X}(\otimes^r_s TM)$ along a path $c : [a,b] \to M$ on the manifold M is defined as a solution of the differential equation : $\nabla_{c'(t)} \widehat{T}(c(t)) = 0$ with initial condition : $\widehat{T}(c(a)) = T(c(a))$*

If T in a holonomic basis reads :

$T(p)$
$= \sum_{\alpha_1...\alpha_r} \sum_{\beta_1....\beta_s} \sum_\gamma T(p)^{\alpha_1...\alpha_r}_{\beta_1...\beta_s} \partial x_{\alpha_1} \otimes .. \otimes \partial x_{\alpha_r} \otimes dx^\gamma \otimes dx^{\beta_1} \otimes ... \otimes dx^{\beta_s}$

$\nabla_{c'(t)} \widehat{T}(t) = 0 \Leftrightarrow \sum_\gamma \widehat{T}^{\alpha_1...\alpha_r}_{\gamma\beta_1...\beta_s} v^\gamma = 0$ with c'(t)$=\sum_\gamma v^\gamma \partial x_\gamma$

The tensor field $\widehat{T}$ is defined by the first order linear differential equations :

$\sum_\gamma v^\gamma \partial_\gamma \widehat{T}^{\alpha_1...\alpha_r}_{\beta_1...\beta_s}$
$= -\sum_{k=1}^r v^\gamma \widehat{\Gamma}^{\alpha_k}_{\gamma\eta} T^{\alpha_1.\alpha_{k-1}\eta\alpha_{k+1}..\alpha_r}_{\beta_1...\beta_s} + \sum_{k=1}^s v^\gamma \widehat{\Gamma}^\eta_{\gamma\beta_k} T^{\alpha_1...\alpha_r}_{\beta_1..\beta_{k-1}\eta\beta_{k+1}..\beta_s}$

$\widehat{T}(c(a)) = T(c(a))$

where Γ, c(t) and v are assumed to be known.

They define a map : $Pt_c : [a,b] \times \mathfrak{X}(\otimes^r_s TM) \to \mathfrak{X}(\otimes^r_s TM)$

If $S, T \in \mathfrak{X}(\otimes^r_s TM), k, k' \in K$ then : $Pt_c(t, kS + k'T) = kPt_c(t, S) + k'Pt_c(t, T)$ but the components of $\widehat{T}$ do not depend linearly of the components of T.

The map : $Pt_c(., T) : [a, b] \to \mathfrak{X}(\otimes^r_s TM) :: Pt_c(t, T)$ is a path in the tensor bundle. So it is common to say that one "lifts" the curve c on M to a curve in the tensor bundle.

Given a vector field V, a point p in M, the set of vectors $u_p \in T_pM$ such that $\nabla_V u_p = 0 \Leftrightarrow \sum_\alpha \left(\partial_\alpha V^\gamma + \Gamma^\gamma_{\alpha\beta} V^\beta \right) u_p^\alpha = 0$ is a vector subspace of T_pM, called the horizontal vector subspace at p (depending on V). So parallel transported vectors are horizontal vectors.

Notice the difference with the transports previously studied :

i) transport by "push-forward" : it can be done everywhere, but the components of the transported tensor depend linearly of the components of T_0

ii) transport by the Lie derivative : it is the transport by push forward with the flow of a vector field, with similar constraints

Holonomy

If the path c is a loop : $c : [a, b] \to M :: c(a) = c(b) = p$ the parallel transport goes back to the same tangent space at p. In the vector space T_pM, which is isomorphic to K^n, the parallel transport for a given loop is a linear map on T_pM, which has an inverse (take the opposite loop with the reversed path) and the set of all such linear maps at p has a structure group : this is the **holonomy group** $H(M, p)$ at p .If the loops re restricted to loops which are homotopic to a point this is the restricted holonomy group $H_0(M, p)$. The holonomy group is a finite dimensional Lie group (Kobayashi I p.72).

Geodesic

1. Definitions:

Definition 1487 *A path* $c \in C_1([a, b]; M)$ *in a manifold M endowed with a covariant derivative* ∇ *describes a* **geodesic** *$c([a, b])$ if the tangent to the curve $c([a, b])$ is parallel transported.*

367

So c describes a geodesic if : $\nabla_{c'(t)}c'(t) = 0 \Leftrightarrow \sum \left(\frac{dV^\beta}{dt} + \Gamma^\beta_{\alpha\gamma}\left(c\left(t\right)\right) V^\alpha V^\gamma \right) = 0$ with $V(t) = c'(t)$

A curve C, that is a 1 dimensional submanifold in M, can be described by different paths. If C is a geodesic for some parameter t, then it is still a geodesic for a parameter $\tau = h(t)$ iff $\tau = kt + k'$ meaning iff h is an affine map.

For a given curve C, which is a geodesic, any path $c \in C_1\left(\left[a,b\right]; M\right)$ such that $c\left(\left[a,b\right]\right) = C$ and for which $\nabla_{c'(t)}c'(t) = 0$ is called an **affine parameter**. They are all linked to each other by an affine map.

If a geodesic is a class 1 path and the covariant derivative is smooth (the coefficients Γ are smooth maps), then c is smooth.

If we define $\widehat{\Gamma}^\alpha_{\beta\gamma} = k\Gamma^\alpha_{\beta\gamma} + (1-k)\Gamma^\alpha_{\beta\gamma}$ with a fixed scalar k, we still have a covariant derivative, which has the same geodesics. In particular with k=1/2 this covariant derivative is torsion free.

2. Fundamental theorem:

Theorem 1488 *(Kobayashi I p.139, 147) For any point p and any vector v in T_pM of a finite dimensional real manifold M endowed with a covariant connection, there is a unique geodesic $c \in C_1\left(J_p; M\right)$ such that $c(0) = p, c'(0) = v$ where J_p is an open interval in $\mathbb{R}$, including 0, depending both of p and v_p.*

For each p there is a neighborhood N(p) of $\left(p, \overrightarrow{0}\right)$x0 in TMxR in which the **exponential map** *: $\exp : TM \times \mathbb{R} \to M :: \exp tv_p = c(t)$ is defined. The point c(t) is the point on the geodesic located at the affine parameter t from p. This map is differentiable and smooth if the covariant derivative is smooth. It is a diffeomorphism from N(p) to a neighborhood n(p) of p in M.*

Warning ! this map exp is not the flow of a vector field, even if its construct is similar. $\frac{d}{dt}\left(\exp tv_p\right)|_{t=\theta}$ is the vector v_p parallel transported along the geodesic.

Theorem 1489 *In a finite dimensional real manifold M endowed with a covariant connection, if there is a geodesic passing through $p \neq q$ in M, it is unique. A geodesic is never a loop.*

This is a direct consequence of the previous theorem.

3. Normal coordinates:

Definition 1490 *In a m dimensional real manifold M endowed with a covariant connection, a system of **normal coordinates** is a local chart defined in a neighborhood n(p) of a point p, with m independant vectors $(\varepsilon_i)_{i=1}^m$ in T_pM, by which to a point $q \in n(p)$ is associated the coordinates $(y_1, ..., y_m)$ such that : $q = \exp v$ with $v = \sum_{i=1}^m y^i \varepsilon_i$.*

In this coordinate system the geodesics are expressed as straigth lines : $c_i(t) \simeq tv$ and the Christofell coefficients are such that *at p* : $\forall i, j, k : \widehat{\Gamma}^i_{jk}(p) + \widehat{\Gamma}^i_{kj}(p) = 0$ so they vanish if the connection is torsion free. Then the covariant derivative of any tensor coincides with the derivative of its components.

Theorem 1491 *(Kobayashi I p.149) Any point p of a finite dimensional real manifold M endowed with a covariant connection has is a convex neighborhood n(p) : two points in n(p) can be joigned by a geodesic which lies in n(p). So there is a system of normal coordinates centered at any point.*

n(p) is defined by a ball centered at p with radius given in a normal coordinate system.

Affine transformation

Definition 1492 *A map $f \in C_1(M;N)$ between the manifolds M,N endowed with the covariant derivatives $\nabla, \widehat{\nabla}$, is an **affine transformation** if it maps a parallel transported vector along a curve c in M into a parallel transported vector along the curve f(c) in N.*

Theorem 1493 *(Kobayashi I p.225) An affine transformation f between the manifolds M,N endowed with the covariant derivatives $\nabla, \widehat{\nabla}$, and the corresponding torsions and curvature tensors $T, \widehat{T}, R, \widehat{R}$*

i) maps geodesics into geodesics

ii) commutes with the exponential : $\exp t\,(f'(p)v_p) = f(\exp t v_p)$

iii) for $X, Y, Z \in \mathfrak{X}(TM)$:

$f^*(\nabla_X Y) = \widehat{\nabla}_{f^*X} f^*Y$

$f^*(T(X,Y)) = \widehat{T}(f^*X, f^*Y)$

$f^*R(X,Y,Z) = \widehat{R}(f^*X, f^*Y, f^*Z)$

iv) is smooth if the connections have smooth Christofell symbols

Definition 1494 *A vector field V on a manifold M endowed with the covariant derivatives ∇ is an infinitesimal generator of affine transformations if $f_t : M \rightarrow M :: f_t(p) = \exp Vt(p)$ is an affine transformation on M.*

V$\in \mathfrak{X}(TM)$ an infinitesimal generator of affine transformationsis on M iff :

$\forall X \in \mathfrak{X}(TM) : \nabla_X(\pounds_V - \nabla_V) = R(V,X)$

The set of affine transformations on a manifold M is a group. If M has a finite number of connected components it is a Lie group with the open compact topology. The set of vector fields which are infinitesimal generators of affine transformations is a Lie subalgebra of $\mathfrak{X}(TM)$, with dimension at most $m^2 + m$. If its dimension is $m^2 + m$ then the torsion and the riemann tensors vanish.

Jacobi field

Definition 1495 *Let a family of geodesics in a manifold M endowed with a covariant derivatives ∇ be defined by a smooth map : $C : [0,1] \times [-a,+a] \rightarrow M$,$a \in \mathbb{R}$ such that $\forall s \in [-a,+a] : C(.,s) \rightarrow M$ is a geodesic on M. The **deviation vector** of the family of geodesics is defined as : $J_t = \frac{\partial C}{\partial s}|_{s=0} \in T_{C(t,0)}M$*

It measures the variation of the family of geodesics along a transversal vector J_t

Theorem 1496 *(Kobayashi II p.63) The deviation vector J of a family of geodesics satisfies the equation :*

$\nabla_{v_t}^2 J_t + \nabla_{v_t}(T(J_t, v_t)) + R(J_t, v_t, v_t) = 0$ *with* $v_t = \frac{\partial C}{\partial t}|_{s=0}$

It is fully defined by the values of $J_t, \nabla_{v_t} J_t$ *at a point t.*

Conversely a vector field $J \in \mathfrak{X}(TM)$ *is said to be a **Jacobi field** if there is a geodesic c(t) in M such that :*

$\forall t : \nabla_{v_t}^2 J(c(t)) + \nabla_{v_t}(T(J(c(t)), v_t)) + R(J(c(t)), v_t, v_t) = 0$ *with* $v_t = \frac{dc}{dt}$

It is then the deviation vector for a family of geodesics built from c(t).

Jacobi fields are the infinitesimal generators of affine transformations.

Definition 1497 *Two points p,q on a geodesic are said to be conjugate if there is a Jacobi field which vanishes both at p and q.*

16.4.6 Submanifolds

If M is a submanifold in N, a covariant derivative ∇ defined on N does not necessarily induce a covariant derivative $\widehat{\nabla}$ on M : indeed even if X,Y are in $\mathfrak{X}(TM)$, $\nabla_X Y$ is not always in $\mathfrak{X}(TM)$.

Definition 1498 *A submanifold M of a manifold N endowed with a covariant derivatives ∇ is **autoparallel** if for each curve in M, the parallel transport of a vector $v_p \in T_p M$ stays in M, or equivalently if $\forall X, Y \in \mathfrak{X}(TM), \nabla_X Y \in \mathfrak{X}(TM)$.*

Theorem 1499 *(Kobayashi II p.54) If a submanifold M of a manifold N endowed with a covariant derivatives ∇ is **autoparallel** then ∇ induces a covariant derivative $\widehat{\nabla}$ on M and :$\forall X, Y \in \mathfrak{X}(TM) : \nabla_X Y = \widehat{\nabla}_X Y$.*

Moreover the curvature and the torsion are related by :

$\forall X, Y, Z \in \mathfrak{X}(TM) : R(X, Y, Z) = \widehat{R}(X, Y, Z), T(X, Y) = \widehat{T}(X, Y)$

M is said to be **totally geodesic** at p if $\forall v_p \in T_p M$ the geodesic of N defined by (p,v_p) lies in M for small values of the parameter t. A submanifold is totally geodesic if it is totally geodesic at each of its point.

An autoparallel submanifold is totally geodesic. But the converse is true only if the covariant derivative on N is torsion free.

17 INTEGRAL

Orientation of a manifold and therefore integral are meaningful only for finite dimensional manifolds. So in this subsection we will limit ourselves to this case.

17.1 Orientation of a manifold

17.1.1 Orientation function

Definition 1500 *Let M be a class 1 finite dimensional manifold with atlas $\left(E, (O_i, \varphi_i)_{i \in I}\right)$, where an orientation has been chosen on E. An **orientation function** is the map : $\theta_i : O_i \to \{+1, -1\}$ with $\theta(p) = +1$ if the holonomic basis defined by φ_i at p has the same orientation as the basis of E and $\theta(p) = -1$ if not.*

If there is an atas of M such that it is possible to define a continuous orientation function over M then it is possible to define continuously an orientation in the tangent bundle.

This leads to the definition :

17.1.2 Orientable manifolds

Definition 1501 *A manifold M is **orientable** if there is a continuous system of orientation functions. It is then oriented if an orientation function has been chosen.*

Theorem 1502 *A class 1 finite dimensional real manifold M is **orientable** iff there is an atlas atlas $\left(E, (O_i, \varphi_i)_{i \in I}\right)$ such that $\forall i, j \in I : \det\left(\varphi_j \circ \varphi_i^{-1}\right)' > 0$*

Proof. We endow the set $\Theta = \{+1, -1\}$ with the discrete topology : $\{+1\}$ and $\{-1\}$ are both open and closed subsets, so we can define continuity for θ_i. If θ_i is continuous on O_i then the subsets $\theta_i^{-1}(+1) = O_i^+, \theta_i^{-1}(-1) = O_i^-$ are both open and closed in O_i . If O_i is connected then we have either $O_i^+ = O_i$, or $O_i^- = O_i$. More generally θ_i has the same value over each of the connected components of O_i.

Let be another chart j such that $p \in O_i \cap O_j$. We have now two maps : $\theta_k : O_k \to \{+1, -1\}$ for k=i,j. We go from one holonomic basis to the other by the transition map :

$e_\alpha = \varphi_i'(p) \partial x_\alpha = \varphi_j'(p) \partial y_\alpha \Rightarrow \partial y_\alpha = \varphi_j'(p)^{-1} \circ \varphi_i'(p) \partial x_\alpha$

The bases $\partial x_\alpha, \partial y_\alpha$ have the same orientation iff $\det \varphi_j'(p)^{-1} \circ \varphi_i'(p) > 0$. As the maps are class 1 diffeomorphisms, the determinant does not vanish and thus keep a constant sign in the neighborhood of p. So in the neighborhood of each point p the functions θ_i, θ_j will keep the same value (which can be different), and so all over the connected components of O_i, O_j. ∎

There are manifolds which are not orientable. The most well known examples are the Möbius strip and the Klein bottle.

Notice that if M is disconnected it can be orientable but the orientation is in fact distinct on each connected component.

By convention a set of disconnected points $M = \cup_{i \in M} \{p_i\}$ is a 0 dimensional orientable manifold and its orientation is given by a function $\theta(p_i) = \pm 1$.

Theorem 1503 *A finite dimensional complex manifold is orientable*

Proof. At any point p there is a canonical orientation of the tangent space, which does not depend of the choice of a real basis or a chart. ∎

Theorem 1504 *An open subset of an orientable manifold is orientable.*

Proof. Its atlas is a restriction of the atlas of the manifold. ∎

An open subset of $\mathbb{R}^m$ is an orientable m dimensional manifold.

A curve on a manifold M defined by a path : $c : J \to M :: c(t)$ is a submanifold if c'(t) is never zero. Then it is orientable (take as direct basis the vectors such that c'(t)u>0).

If $(V_i)_{i=1}^m$ are m linearly independant continuous vector fields over M then the orientation of the basis given by them is continuous in a neighborhood of each point. But it does not usually defines an orientation on M, because if M is not parallelizable there is not such vector fields.

A diffeomorphism $f : M \to N$ between two finite dimensional real manifolds preserves (resp.reverses) the orientation if in two atlas: $\det \left(\psi_j \circ f \circ \varphi_i^{-1} \right)' > 0$ (resp.<0).

As $\det \left(\psi_j \circ f \circ \varphi_i^{-1} \right)'$ is never zero and continuous it has a constant sign : If two manifolds M,N are diffeomorphic, if M is orientable then N is orientable.

Notice that M,N must have the same dimension.

17.1.3 Volume form

Definition 1505 *A **volume form** on a m dimensional manifold M is a m-form $\varpi \in \mathfrak{X}(\Lambda_m TM^*)$ which is never zero on M.*

Any m form μ on M can then be written $\mu = f\varpi$ with $f \in C(M; \mathbb{R})$.

Warning ! the symbol "$dx^1 \wedge ... \wedge dx^m$" is not a volume form, except if M is an open of $\mathbb{R}^m$. Indeed it is the coordinate expression of a m form in some chart φ_i :$\varpi_i(p) = 1 \, \forall p \in O_i$. At a transition $p \in O_i \cap O_j$ we have, for the same form : $\varpi_j = \det \left[J^{-1} \right] \neq 0$ so we still have a volume form, but it is defined only on the part of O_j which intersects O_i. We cannot say anything outside O_i. And of course put $\varpi_j(q) = 1$ would not define the same form. More generally $f(p) dx^1 \wedge ... \wedge dx^m$ were f is a function on M, meaning that its value is the same in any chart, does not define a volume form, not even a m form. In a pseudo-riemannian manifold the volume form is $\sqrt{|\det g|} dx^1 \wedge ... \wedge dx^m$ where the value of $|\det g|$ is well defined at any point, but changes according to the usual rules in a change of basis.

Theorem 1506 *(Lafontaine p.201) A class 1 m dimensional manifold M which is the union of countably many compact sets is orientable iff there is a volume form.*

As a consequence a m dimensional submanifold of M is itself orientable (take the restriction of the volume form). It is not true for a n<m submanifold.

A riemannian, pseudo-riemannian or symplectic manifold has such a form, thus is orientable if it is the union of countably many compact sets.

17.1.4   Orientation of an hypersurface

Definition 1507 *Let M be a hypersurface of a class 1 n dimensional manifold N. A vector $u_p \in T_pN, p \in M$ is **transversal** if $u_p \notin T_pM$*

At any point we can have a basis comprised of $(u_p, \varepsilon_2, ...\varepsilon_n)$ where $(\varepsilon_\beta)_{\beta=2}^{n}$ is a local basis of T_pM . Thus we can define a transversal orientation function by the orientation of this basis : say that $\theta(u_p) = +1$ if $(u_p, \varepsilon_2, ...\varepsilon_n)$ is direct and $\theta(u_p) = -1$ if not.

M is transversally orientable if there is a continuous map θ.

Theorem 1508 *The boundary of a manifold with boundary is transversally orientable*

See manifold with boundary. It does not require N to be orientable.

Theorem 1509 *A manifold M with boundary ∂M in an orientable class 1 manifold N is orientable.*

Proof. The interior of M is an open subset of N, so is orientable. There is an outward going vector field n on ∂M , so we can define a direct basis (e_α) on ∂M as a basis such that $(n, e_1, ..., e_{m-1})$ is direct in N and ∂M is an orientable manifold ∎

17.2   Integral

In the Analysis part measures and integral are defined on any set . A m dimensional real manifold M is locally homeomorphic to $\mathbb{R}^m$, thus it implies some constraints on the Borel measures on M, whose absolutely continuous part must be related to the Lebesgue measure. Conversely any m form on a m dimensional manifold defines an absolutely continuous measure, called a Lebesgue measure on the manifold, and we can define the integral of a m form.

17.2.1   Definitions

Principle

1. Let M be a Hausdorff, m dimensional real manifold with atlas
$\left(\mathbb{R}^m, (O_i, \varphi_i)_{i \in I}\right), U_i = \varphi_i(O_i)$ and μ a positive, locally finite Borel measure on M. It is also a Radon measure.

i) On $\mathbb{R}^m$ is defined the Lebesgue measure $d\xi$ which can be seen as the tensorial product of the measures $d\xi^k, k = 1...m$ and reads : $d\xi = d\xi^1 \otimes ... \otimes d\xi^n$ or more simply : $d\xi = d\xi^1 .. d\xi^m$

ii) The charts define push forward positive Radon measures $\nu_i = \varphi_{i*}\mu$ on $U_i \subset \mathbb{R}^m$ such that $\forall B \subset U_i : \varphi_{i*}\mu(B) = \mu(\varphi_i^{-1}(B))$

Each of the measures ν_i can be uniquely decomposed in a singular part λ_i and an absolute part $\widehat{\nu}_i$, which itself can be written as the integral of some positive function $g_i \in C(U_i; \mathbb{R})$ with respect to the Lebesgue measure on $\mathbb{R}^m$

Thus for each chart there is a couple (g_i, λ_i) such that : $\nu_i = \varphi_{i*}\mu = \widehat{\nu}_i + \lambda_i$, $\widehat{\nu}_i = g_i(\xi) d\xi$

If a measurable subset A belongs to the intersection of the domains $O_i \cap O_j$ and for any i,j :

$\varphi_{i*}\mu(\varphi_i(A)) = \mu(A) = \varphi_{j*}\mu(\varphi_j(A))$

Thus there is a unique Radon measure ν on $U = \cup_i U_i \subset \mathbb{R}^m$ such that : $\nu = \nu_i$ on each U_i. ν can be seen as the push forward on $\mathbb{R}^m$ of the measure μ on M by the atlas. This measure can be decomposed as above :

$\nu = \widehat{\nu} + \lambda$, $\widehat{\nu} = g(\xi) d\xi$

iii) Conversely the pull back $\varphi_i^* \nu$ of ν by each chart on each open O_i gives a Radon measure μ_i on O_i and μ is the unique Radon measure on M such that $\mu|_{O_i} = \varphi_i^* \nu$ on each O_i.

iv) Pull back and push forward are linear operators, they apply to the singular and the absolutely continuous parts of the measures. So the absolutely continuous part of μ denoted $\widehat{\mu}$ is the pull back of the product of g with the Lebesgue measure :

$\widehat{\mu}|_{O_i} = \varphi_i^*(\widehat{\nu}|_{U_i}) = \varphi_i^* \widehat{\nu}_i = \varphi_i^*(g_i(\xi) d\xi)$

$\widehat{\nu}|_{U_i} = \varphi_{i*}(\widehat{\mu}|_{O_i}) = \varphi_{i*}\widehat{\mu}_i = g_i(\xi) d\xi$

2. On the intersections $U_i \cap U_j$ the maps : $\varphi_{ij} = \varphi_j \circ \varphi_i^{-1} : U_i \to U_j$ are class r diffeomorphisms, the push forward of $\nu_i = \varphi_{i*}\mu$ by φ_{ij} is :$(\varphi_{ij})_* \varphi_{i*}\mu = (\varphi_j \circ \varphi_i^{-1})_* \varphi_{i*}\mu = \varphi_{j*}\mu$

$\widehat{\nu}_j = \varphi_{j*}\widehat{\mu}$ being the image of $\widehat{\nu}_i = \varphi_{i*}\widehat{\mu}$ by the diffeomorphism φ_{ij} reads :

$\widehat{\nu}_j = (\varphi_{ij})_* \widehat{\nu}_i = |\det[\varphi'_{ij}]| \widehat{\nu}_i$

which resumes to : $g_j = |\det[\varphi'_{ij}]| g_i$

So, even if there is a function g such that ν is the Radon integral of g, g itself is defined as a family $(g_i)_{i \in I}$ of functions changing according to the above formula through the open cover of M.

3. On the other hand a m form on M reads $\varpi = \varpi(p) dx^1 \wedge dx^2 ... \wedge dx^m$ in the holonomic basis. Its components are a family $(\varpi_i)_{i \in I}$ of functions $\varpi_i : O_i \to \mathbb{R}$ such that : $\varpi_j = \det[\varphi'_{ij}]^{-1} \varpi_i$ on the intersection $O_i \cap O_j$.

The push forward of ϖ by a chart gives a m form on $\mathbb{R}^m$:

$(\varphi_{i*}\varpi_i)(\xi) = \varpi_i(\varphi_i^{-1}(\xi)) e^1 \wedge ... \Lambda e^m$ in the corresponding basis $(e^k)_{k=1}^m$ of $(\mathbb{R}^m)^*$

and on $O_i \cap O_j$:

$(\varphi_{j*}\varpi_j) = (\varphi_{ij})_* \varphi_{i*}\varpi_i = \det[\varphi'_{ij}]^{-1} \varpi_i(\varphi_i^{-1}(\xi)) e^1 \wedge ... \Lambda e^m$

So the rules for the transformations of the component of a m-form, and the functions g_i are similar (but not identical). Which leads to the following

definitions.

Integral of a m form on a manifold

Theorem 1510 *On a m dimensional oriented Hausdorff class 1 real manifold M, any continuous m form ϖ defines a unique, absolutely continuous, Radon measure on M, called the **Lebesgue measure** associated to ϖ.*

Proof. Let $\left(\mathbb{R}^m, (O_i, \varphi_i)_{i \in I}\right), U_i = \varphi_i(O_i)$ be an atlas of M as above. As M is oriented the atlas can be chosen such that $\det\left[\varphi'_{ij}\right] > 0$. Take a continuous m form ϖ on M .On each open $U_i = \varphi_i(O_i)$ we define the Radon measure : $\nu_i = \varphi_{i*}(\varpi_i) \, d\xi$. It is locally finite and finite if $\int_{U_i} |(\varphi_{i*}\varpi_i)| \, d\xi < \infty$ .Then on the subsets $U_i \cap U_j \neq \varnothing$: $\nu_i = \nu_j$ .Thus the family $(\nu_i)_{i \in I}$ defines a unique Radon measure, absolutely continuous, on $U = \cup_i U_i \subset \mathbb{R}^m$ .The pull back, on each chart, of the ν_i give a family $(\mu_i)_{i \in I}$ of Radon measures on each O_i and from there a locally compact, absolutely continuous, Radon measure on M.

It can be shown (Schwartz IV p.319) that the measure does not depend on the atlas with the same orientation on M. ∎

Definition 1511 *The **Lebesgue integral** of a m form ϖ on M is $\int_M \mu_\varpi$ where μ_ϖ is the Lebesgue measure on M which is defined by ϖ.*

It is denoted $\int_M \varpi$

An open subset Ω of an orientable manifold is an orientable manifold of the same dimension, so the integral of a m-form on any open of M is given by restriction of the measure μ : $\int_\Omega \varpi$

Remaks

i) the measure is linked to the Lebesgue measure but, from the definition, whenever we have an absolutely continuous Radon measure μ on M, there is a m form such that μ is the Lebesgue measure for some form. However there are singular measures on M which are not linked to the Lebesgue measure.

ii) without orientation on each domain there are two measures, different by the sign, but there is no guarantee that one can define a unique measure on the whole of M. Such "measures" are called densities.

iii) On $\mathbb{R}^m$ we have the canonical volume form : $dx = dx^1 \wedge ... \wedge dx^m$, which naturally induces the Lebesgue measure, also denoted dx=$dx^1 \otimes ... \otimes dx^m = dx^1 dx^2 ... dx^m$

iv) The product of the Lebesgue form ϖ_μ by a function $f : M \to \mathbb{R}$ gives another measure and : $f\varpi_\mu = \varpi_{f\mu}$.Thus, given a m form ϖ, the integral of any continuous function on M can be defined, but its value depends on the choice of ϖ.

If there is a volume form ϖ_0, then for any function $f : M \to \mathbb{R}$ the linear functional $f \to \int_M f\varpi_0$ can be defined.

Warning ! the quantity $\int_M f dx^1 \wedge ... \wedge dx^m$ where f is a function is not defined (except if M is an open in $\mathbb{R}^m$) because $f dx^1 \wedge ... \wedge dx^m$ is not a m form.

v) If M is a set of a finite number of points $M = \{p_i\}_{i \in I}$ then this is a 0-dimensional manifold, a 0-form on M is just a map : $f : M \to \mathbb{R}$ and the integral is defined as : $\int_M f = \sum_{i \in I} f(p_i)$

vi) For m manifolds M with compact boundary in $\mathbb{R}^m$ the integral $\int_M dx$ is proportionnal to the usual euclidean "volume" delimited by M.

Integrals on a r-simplex

It is useful for practical purposes to be able to compute integrals on subsets of a manifold M which are not submanifolds, for instance subsets delimited regularly by a finite number of points of M. The r-simplices on a manifold meet this purpose (see Homology on manifolds).

Definition 1512 *The integral of a r form $\varpi \in \mathfrak{X}(\Lambda_r T M^*)$ on a r-simplex $M^r = f(S^r)$ of a m dimensional oriented Hausdorff class 1 real manifold M is given by :* $\int_{M^r} = \int_{S^r} f^* \varpi dx$

f$\in C_\infty (\mathbb{R}^m; M)$ and
$S^r = S^r = \langle A_0, ... A_r \rangle = \{P \in \mathbb{R}^m : P = \sum_{i=0}^r t_i A_i ; 0 \le t_i \le 1, \sum_{i=0}^r t_i = 1\}$
is a r-simplex on $\mathbb{R}^m$.

$f^* \varpi \in \mathfrak{X}(\Lambda_r \mathbb{R}^m)$ and the integral $\int_{S^r} f^* \varpi dx$ is computed in the classical way. Indeed $f^* \varpi = \sum \pi_{\alpha_1 ... \alpha_r} dx^{\alpha_1} \wedge ... \wedge dx^{\alpha_r}$ so the integrals are of the kind : $\int_{s^r} \pi_{\alpha_1 ... \alpha_r} dx^{\alpha_1} ... dx^{\alpha_r}$ on domains s^r which are the convex hull of the r dimensional subspaces generated by r+1 points, there are r variables and a r dimensional domain of integration.

Notice that here a m form (meaning a form of the same order as the dimension of the manifold) is not needed. But the condition is to have a r-simplex and a r form.

For a r-chain $C^r = \sum_i k_i M_i^r$ on M then : $\int_{C^r} \varpi = \sum_i k_i \int_{M^r} \varpi = \sum_i k_i \int_{S_i^r} f^* \varpi dx$.
and : $\int_{C^r + D^r} \varpi = \int_{C^r} \varpi + \int_{D^r} \varpi$

17.2.2   Properties of the integral

Theorem 1513 $\int_M$ *is a linear operator :* $\mathfrak{X}(\Lambda_m T M^*) \to \mathbb{R}$

$\forall k, k' \in \mathbb{R}, \varpi, \pi \in \Lambda_m T M^* : \int_M (k\varpi + k'\pi) \mu = k \int_M \varpi \mu + k' \int_M \pi \mu$

Theorem 1514 *If the orientation on M is reversed,* $\int_M \varpi \mu \to - \int_M \varpi \mu$

Theorem 1515 *If a manifold is endowed with a continuous volume form ϖ_0 the induced Lebesgue measure μ_0 on M can be chosen such that it is positive, locally compact, and M is $\sigma-$additive with respect to μ_0.*

Proof. If the component of ϖ_0 is never null and continuous it keeps its sign over M and we can choose ϖ_0 such it is positive. The rest comes from the measure theory. ∎

Theorem 1516 *(Schwartz IV p.332) If f$\in C_1(M; N)$ is a diffeomorphism be-tween two oriented manifolds, which preserves the orientation, then : $\forall \varpi \in \mathfrak{X}_1(\Lambda_m TM^*)$:$\int_M \varpi = \int_N (f_* \varpi)$*

This result is not surprising : the integrals can be seen as the same integral computed in different charts.

Conversely :

Theorem 1517 *Moser's theorem (Lang p.406) Let M be a compact, real, finite dimensional manifold with volume forms ϖ, π such that : $\int_M \varpi = \int_M \pi$ then there is a diffeomorphism $f : M \to M$ such that $\pi = f^* \varpi$*

If M is a m dimensional submanifold of the n>m manifold N, both oriented, f an embedding of M into N, then the integral on M of a m form in N can be defined by :

$\forall \varpi \in \mathfrak{X}_1(\Lambda_m TN^*) : \int_M \varpi = \int_{f(M)} (f_* \varpi)$

because f is a diffeormophism of M to f(M) and f(M) an open subset of N.

Example : a curve $c : J \to N :: c(t)$ on the manifold N is a orientable submanifold if c'(t)$\neq$ 0. For any 1-form over N : $\varpi(p) = \sum_\alpha \varpi_\alpha(p) dx^\alpha$. So $c_* \varpi = \varpi(c(t)) c'(t) dt$ and $\int_c \varpi = \int_J \varpi(c(t)) c'(t) dt$

17.2.3 Stokes theorem

1. For the physicists it is the most important theorem of differential geometry. It can be written :

Theorem 1518 *Stokes theorem ; For any manifold with boundary M in a n dimensional real orientable manifold N and any n–1 form $\varpi \in \mathfrak{X}_1(\wedge_{n-1} TN^*)$:* $\int_M d\varpi = \int_{\partial M} \varpi$

This theorem requires some comments and conditions .

2. Comments :
i) the exterior differential $d\varpi$ is a n-form, so its integral in N makes sense, and the integration over M, which is a closed subset of N, must be read as : $\int_{\overset{\circ}{M}} d\varpi$,

meaning the integral over the open subset $\overset{\circ}{M}$ of N (which is a n-dimensional submanifold of N).

ii) the boundary is a n-1 orientable submanifold in N, so the integral of a the n-1 form ϖ makes sense. Notice that the Lebesgue measures are not the same : on M is is induced by $d\varpi$, on ∂M it is induced by the restriction $\varpi|_{\partial M}$ of ϖ on ∂M

iii) the n-1 form ϖ does not need to be defined over the whole of N : only the domain included in M (with boundary) matters, but as we have not defined forms over manifold with boundary it is simpler to look at it this way. And of course it must be at least of class 1 to compute its exterior derivative.

3. Conditions :

There are several alternate conditions. The theorem stands if one of the following condition is met:

   i) the simplest : M is compact

   ii) ϖ is compactly supported : the support $\text{Sup}(\varpi)$ is the closure of the set
: $\{p \in M : \varpi(p) \neq 0\}$

   iii) $\text{Sup}(\varpi) \cap M$ is compact

   Others more complicated conditions exist.

4. There is another useful version of the theorem. If C is a r-chain on M, then both the integral $\int_C \varpi$ and the border ∂C ot the r chain are defined. And the equivalent of the Stokes theorem reads :

   If C is a r-chain on M, $\varpi \in \mathfrak{X}_1 (\wedge_{r-1} TM^*)$ then $\int_C d\varpi = \int_{\partial C} \varpi$

Theorem 1519 *Integral on a curve (Schwartz IV p.339) Let E be a finite dimensional real normed affine space. A continuous curve C generated by a path c :* $[a, b] \rightarrow E$ *on E is rectifiable if* $\ell(c) < \infty$ *with* $\ell(c) = \sup \sum_{k=1}^{n} d(p(t_{k+1})), p(t_k))$ *for any increasing sequence* $(t_n)_{n \in \mathbb{N}}$ *in [a,b] and d the metric induced by the norm. The curve is oriented in the natural way (t increasing).*

For any function $f \in C_1(E; \mathbb{R})$: $\int_C df = f(c(b)) - f(c(a))$

17.2.4 Divergence

Definition

Theorem 1520 *For any vector field* $V \in \mathfrak{X}(TM)$ *on a manifold endowed with a volume form* ϖ_0 *there is a function div(V) on M, called the **divergence** of the vector field, such that* $\pounds_V \varpi_0 = (divV) \varpi_0$

Proof. If M is m dimensional, ϖ_0, $\pounds_V \varpi_0 \in \mathfrak{X}(\wedge_m TM^*)$. All m forms are proportional on M and ϖ_0 is never null, then $\forall p \in M, \exists k \in K : \pounds_V \varpi_0(p) = k\varpi_0(p)$ ∎

Expression in a holonomic basis

   $\forall V \in \mathfrak{X}(TM) : \pounds_V \varpi_0 = i_V d\varpi_0 + d \circ i_V \varpi_0$ and $d\varpi_0 = 0$ so $\pounds_V \varpi_0 = d(i_V \varpi_0)$

   $\varpi_0 = \varpi_0(p) dx^1 \Lambda ... \Lambda dx^m : \pounds_V \omega_0 = d\left(\sum_\alpha V^\alpha (-1)^{\alpha-1} \varpi_0 dx^1 \Lambda .. \widehat{dx^\alpha} . \Lambda dx^m\right)$

   $= \sum_\beta \partial_\beta \left(V^\alpha (-1)^{\alpha-1} \varpi_0\right) dx^\beta \wedge dx^1 \Lambda .. \widehat{dx^\alpha} . \Lambda dx^m = \left(\sum_\alpha \partial_\alpha (V^\alpha \varpi_0)\right) dx^1 \Lambda .. \Lambda dx^m$

   So : $divV = \frac{1}{\varpi_0} \sum_\alpha \partial_\alpha (V^\alpha \varpi_0)$

Properties

   For any $f \in C_1(M; \mathbb{R})$, $V \in \mathfrak{X}(M) : fV \in \mathfrak{X}(M)$ and

   $div(fV) \varpi_0 = d(i_{fV} \varpi_0) = d(f i_V \varpi_0) = df \wedge i_V \varpi_0 + f d(i_V \varpi_0) = df \wedge i_V \varpi_0 + f div(V) \varpi_0$

$$df \wedge i_V \varpi_0 = \left(\sum_\alpha \partial_\alpha f dx^\alpha\right) \wedge \left(\sum_\beta (-1)^\beta V^\beta \varpi_0 dx^1 \wedge ...\widehat{dx^\beta}... \wedge dx^m\right) = \left(\sum_\alpha V^\alpha \partial_\alpha f\right) \varpi_0 = f'(V) \varpi_0$$

So : $div(fV) = f'(V) + fdiv(V)$

Divergence theorem

Theorem 1521 *For any vector field $V \in \mathfrak{X}_1(TM)$ on a manifold N endowed with a volume form ϖ_0, and manifold with boundary M in N: $\int_M (divV) \varpi_0 = \int_{\partial M} i_V \varpi_0$*

Proof. $\pounds_V \varpi_0 = (divV) \varpi_0 = d(i_V \varpi_0)$

In conditions where the Stockes theorem holds :

$\int_M d(i_V \varpi_0) = \int_M (divV) \varpi_0 = \int_{\partial M} i_V \varpi_0$ ∎

ϖ_0 defines a volume form on N, and the interior of M (which is an open subset of N). So any class 1 vector field on N defines a Lebesgue measure on ∂M by $i_V \varpi_0$.

If M is endowed with a Riemannian metric there is an outgoing unitary vector n on ∂M (see next section) which defines a measure ϖ_1 on ∂M and : $i_V \varpi_0 = \langle V, n \rangle \varpi_1 = i_V \varpi_0$ so $\int_M (divV) \varpi_0 = \int_{\partial M} \langle V, n \rangle \varpi_1$

17.2.5   Integral on domains depending on a parameter

Anticipating on the next section.

Layer integral:

Let (M,g) be a class 1 m dimensional real riemannian manifold with the volume form ϖ_0, f a class 1 function : $f : M \to \mathbb{R}$. We want to compute : $\int_M f \varpi_0$. Any function $p \in C_1(M; \mathbb{R})$ such that p'(x)≠ 0 defines a family of manifolds with boundary N(t)=$\{x \in M : p(x) \le t\}$ in M, which are diffeomorphic by the flow of the gradiant grad(p). Using an atlas of M there is a folliation in $\mathbb{R}^m$ and using the Fubini theorem the integral can be computed by summing first over the hypersurface defined by N(t) then by taking the integral over t.

Theorem 1522 *(Schwartz 4 p.99) Let M be a m dimensional class 1 real riemannian manifold with the volume form ϖ_0, then for any function $f \in C_1(M; \mathbb{R})$ ϖ_0−integrable on M such that f'(x)≠ 0 on M, for almost every value of t, the function $g(x) = \frac{f(x)}{\|gradf\|}$ is integrable on the hypersurface $N(t) = \{x \in M : f(x) = t\}$ and we have :*

$\int_M f \varpi_0 = \int_0^\infty \left(\int_{N(t)} \frac{f(x)}{\|gradf\|} \sigma(t)\right) dt$ where $\sigma(t)$ is the volume form induced on N(t) by ϖ_0

(the Schwartz's demontration for an affine space is easily extended to a real manifold)

$\sigma(t) = i_{n(t)} \varpi_0$ where $n = \frac{gradf}{\|gradf\|}$ (see Pseudo riemannian manifolds) so

$\int_M f \varpi_0 = \int_0^\infty \int_{N(t)} \frac{f(x)}{\|gradf\|^2} i_{gradf} \varpi_0$

Remark : the previous theorem does not use the fact that M is riemannian, and the formula is valid whenever g is not degenerate on N(t), but we need both f'$\neq 0$, $\|gradf\| \neq 0$ which cannot be guarantee without a positive definite metric.

Integral on a domain depending on a parameter :

Theorem 1523 *Reynold's theorem: Let (M,g) be a class 1 m dimensional real riemannian manifold with the volume form ϖ_0, f a function $f \in C_1(\mathbb{R} \times M; \mathbb{R})$, N(t) a family of manifolds with boundary in M, then :*

$\frac{d}{dt} \int_{N(t)} f(t,x) \varpi_0(x) = \int_{N(t)} \frac{\partial f}{\partial t}(t,x) \varpi_0(x) + \int_{\partial N(t)} f(x,t) \langle v,n \rangle \sigma(t)$ *where* $v(q(t)) = \frac{dq}{dt}$ *for q(t)$\in N(t)$*

This assumes that there is some map : $\phi : \mathbb{R} \times M \to M :: \phi(t, q(s)) = q(t+s) \in N(t+s)$

If N(t) is defined by a function p : $N(t) = \{x \in M : p(x) \leq t\}$ then :

$\frac{d}{dt} \int_{N(t)} f(t,x) \varpi_0(x) = \int_{N(t)} \frac{\partial f}{\partial t}(t,x) \varpi_0(x) + \int_{\partial N(t)} \frac{f(x,t)}{\|gradp\|} \sigma(t)$

Proof. the boundaries are diffeomorphic by the flow of the vector field (see Manifolds with boundary) :

$V = \frac{gradp}{\|gradp\|^2} :: \forall q_t \in \partial N(t) : \Phi_V(q_t, s) \in \partial N_{t+s}$

So : $v(q(t)) = \frac{\partial}{\partial t} \Phi_V(q_t, s)|_{t=s} = V(q(t)) = \frac{gradp}{\|gradp\|^2}|_{q(t)}$

On the other hand : $n = \frac{gradp}{\|gradp\|}$

$\langle v,n \rangle = \frac{\|gradp\|^2}{\|gradp\|^3} = \frac{1}{\|gradp\|}$ ∎

Formula which is consistent with the previous one if f does not depend on t.

m forms depending on a parameter:

μ is a family $\mu(t)$ of m form on M such that : $\mu : \mathbb{R} \to \mathfrak{X}(\Lambda_m TM^*)$ is a class 1 map and one considers the integral : $\int_{N(t)} \mu$ where N(t) is a manifold with boundary defined by $N(t) = \{x \in M : p(x) \leq t\}$

M is extended to $\mathbb{R} \times M$ with the riemannian metric $G = dt \otimes dt + \sum g_{\alpha\beta} dx^\alpha \otimes dx^\beta$

With $\lambda = dt \wedge \mu(t) : D\mu = \frac{\partial \mu}{\partial t} dt \wedge \mu + d\mu \wedge \mu = \frac{\partial \mu}{\partial t} dt \wedge \mu$

With the previous theorem : $\int_{M \times I(t)} D\mu = \int_{N(t)} \mu \varpi_0$ where $I(t) = [0, t]$

$\frac{d}{dt} \int_{N(t)} \mu = \int_{N(t)} i_v(d_x \varpi) + \int_{N(t)} \frac{\partial \mu}{\partial t} + \int_{\partial N(t)} i_V \mu$

where $d_x \varpi$ is the usual exterior derivative with respect to x, and $v = gradp$

17.3 Cohomology

Also called de Rahm cohomology (there are other concepts of cohomology). It is a branch of algebraic topology adapted to manifolds, which gives a classification of manifolds and is related to the homology on manifolds.

17.3.1 Spaces of cohomology

Definition

Let M be a real smooth manifold modelled over the Banach E.

1. The **de Rahm complex** is the sequence :

$$0 \to \mathfrak{X}\left(\Lambda_0 TM^*\right) \overset{d}{\to} \mathfrak{X}\left(\Lambda_1 TM^*\right) \overset{d}{\to} \mathfrak{X}\left(\Lambda_2 TM^*\right) \overset{d}{\to} ...$$

In the categories parlance this is a sequence because the image of the operator d is just the kernel for the next operation :

if $\varpi \in \mathfrak{X}\left(\Lambda_r TM^*\right)$ then $d\varpi \in \mathfrak{X}\left(\Lambda_{r+1} TM^*\right)$ and $d^2 \varpi = 0$

An exact form is a closed form, the Poincaré lemna tells that the converse is locally true, and cohomology studies this fact.

2.Denote the set of closed r-forms : $F^r\left(M\right) = \left\{\varpi \in \mathfrak{X}\left(\Lambda_r TM^*\right) : d\varpi = 0\right\}$ with $F^0\left(M\right)$ the set of locally constant functions. $F^r\left(M\right)$ is sometimes called the set of cocycles.

Denote the set of exact r–1 forms :

$G^{r-1}\left(M\right) = \left\{\varpi \in \mathfrak{X}\left(\Lambda_r TM^*\right) : \exists \pi \in \mathfrak{X}\left(\Lambda_{r-1} TM^*\right) : \varpi = d\pi\right\}$ sometimes called the set of coboundary.

Definition 1524 *The rth* **space of cohomology** *of a manifold M is the quotient space :* $H^r\left(M\right) = F^r\left(M\right)/G^{r-1}\left(M\right)$

The definition makes sense : $F^r\left(M\right), G^{r-1}\left(M\right)$ are vector spaces over K and $G^{r-1}\left(M\right)$ is a vector subspace of $F^r\left(M\right).$

Two closed forms in one class of equivalence denoted [] differ by an exact form :

$\varpi_1 \sim \varpi_2 \Leftrightarrow \exists \pi \in \mathfrak{X}\left(\Lambda_{r-1} TM^*\right) : \varpi_2 = \varpi_1 + d\pi$

The exterior product extends to $H^r\left(M\right)$

$[\varpi] \in H^p(V), [\pi] \in H^q(V) : [\varpi] \wedge [\pi] = [\varpi \wedge \pi] \in H^{p+q}(V)$

So : $\oplus_{r=0}^{\dim M} H^r\left(M\right) = H^*\left(M\right)$ has the structure of an algebra over the field K

Properties

Definition 1525 *The r* **Betti number** *of the manifold M is the dimension of* $H^r\left(M\right)$

The **Euler characteristic** *of the manifold M is :* $\chi\left(M\right) = \sum_{r=1}^{\dim M} \left(-1\right)^r b_r\left(M\right)$

They are topological invariant : two diffeomorphic manifolds have the same Betti numbers and Euler characteristic.

Betti numbers count the number of "holes" of dimension r in the manifold.

$\chi\left(M\right) = 0$ if dimM is odd.

Definition 1526 *The* **Poincaré polynomial** *on the field K is :* $P\left(M\right) : K \to K : P\left(M\right)\left(z\right) = \sum_r b_r\left(M\right) z^r$

For two manifolds M,N : $P(M \times N) = P(M) \times P(N)$

The Poincaré polynomials can be computed for Lie groups (see Wikipedia, Betti numbers).

If M has n connected components then : $H^0(M) \simeq \mathbb{R}^n$. This follows from the fact that any smooth function on M with zero derivative (i.e. locally constant) is constant on each of the connected components of M. So $b_0(M)$ is the number of connected components of M,

If M is a simply connected manifold then $H^1(M)$ is trivial (it has a unique class of equivalence which is [0]) and $b_1(M) = 0$.

Theorem 1527 *If M,N are two real smooth manifolds and $f : M \to N$ then :*

i) the pull back $f^*\varpi$ of closed (resp.exact) forms ϖ is a closed (resp.exact) form so :

$f^*[\varpi] = [f^*\varpi] \in H^r(M)$

ii) if $f, g \in C_\infty(M; N)$ are homotopic then $\forall \varpi \in H^r(N) : f^*[\varpi] = g^*[\varpi]$

Theorem 1528 ***Künneth formula*** *: Let M_1, M_2 smooth finite dimensional real manifolds :*

$H^r(M_1 \times M_2) = \underset{p+q=r}{\oplus} [H^p(M_1) \otimes H^q(M_2)]$

$H^*(M_1 \times M_2) = H^*(M_1) \times H^*(M_2)$

$b_r(M_1 \times M_2) = \sum_{q+p=r} b_p(M_1) b_q(M_2)$

$\chi(M_1 \times M_2) = \chi(M_1)\chi(M_2)$

17.3.2 de Rahm theorem

Let M be a real smooth manifold

The sets $C^r(M)$ of r-chains on M and $\mathfrak{X}(\Lambda_r TM^*)$ of r-forms on M are real vector spaces. The map :

$\langle\rangle : C^r(M) \times \mathfrak{X}(\Lambda_r TM^*) \to \mathbb{R} :: \langle C, \varpi\rangle = \int_C \varpi$

is bilinear. And the Stokes theorem reads : $\langle C, d\varpi\rangle = \langle\partial C, \varpi\rangle$

This map stands with the quotient spaces $H^r(M)$ of homologous r-chains and $H_r(M)$ of cohomologous r-forms:

$\langle\rangle : H^r(M) \times H_r(M) \to \mathbb{R} :: \langle[C], [\varpi]\rangle = \int_{[C]} [\varpi]$

In some manner these two vector spaces can be seen as "dual" from each other.

Theorem 1529 *de Rahm : If M is a real, m dimensional, compact manifold, then :*

i) the vector spaces $H^r(M), H_r(M)$ have the same finite dimension equal to the rth Betti number $b_r(M)$

$b_r(M) = 0$ *if $r > dimM$, $b_r(M) = b_{m-r}(M)$*

ii) the map $\langle\rangle : H^r(M) \times H_r(M) \to \mathbb{R}$ is non degenerate

iii) $H^r(M) = H_r(M)^*$

iv) $H^r(M) \simeq H^{m-r}(M)$

v) Let $M_i^r \in C^r(M), i = 1...b_r(M) : \forall i \neq j : [M_i^r] \neq [M_j^r]$ then :

a closed r-form ϖ is exact iff $\forall i = 1...b_r : \int_{M_i^r} \varpi = 0$

$\forall k_i \in \mathbb{R}, i = 1...b_r, \exists \varpi \in \Lambda_r TM^* : d\varpi = 0, \int_{M_i^r} \varpi = k_i$

Theorem 1530 *(Lafontaine p.233) Let M be a smooth real m dimensional, compact, connected manifold, then:*
 i) a m form ϖ is exact iff $\int_M \varpi = 0$
 ii) $H^m(M)$ is isomorphic to $\mathbb{R}$

Notice that they stand for *compact* manifolds.

17.3.3 Degree of a map

Theorem 1531 *(Lafontaine p.235) Let M,N be smooth real m dimensional, compact, oriented manifolds, $f \in C_\infty(M;N)$ then there is a signed integer k(f) called the **degree of the map** such that :$\exists k(f) \in \mathbb{Z} : \forall \varpi \in \Lambda_m TM^* : \int_M f^* \varpi = k(f) \int_N \varpi$*

If f is not surjective then k(f)=0
If f,g are homotopic then k(f)=k(g)

Theorem 1532 *(Taylor 1 p.101) Let M be a compact manifold with boundary, N a smooth compact oriented real manifold, f$\in C_1(M;N)$ then : $Deg(f|_{\partial M}) = 0$*

18   COMPLEX MANIFOLDS

Everything which has been said before for manifolds stand for complex manifolds, if not stated otherwise. However complex manifolds have specific properties linked on one hand to the properties of holomorphic maps and on the other hand to the relationship between real and complex structures.

It is useful to refer to the Algebra part about complex vector spaces.

The key point is that the main constructs involve only the tangent bundle, not the manifold structure itself.

18.1   Complex manifolds

Complex manifolds, meaning manifolds whose manifold structure is complex are a different story.

18.1.1   General properties

1. Complex manifolds are manifolds modelled on a Banach vector space E over $\mathbb{C}$. The transition maps : $\varphi_j \circ \varphi_i^{-1}$ are C-differentiable maps between Banach vector spaces, so they are holomorphic maps, and smooth. Thus a differentiable complex manifold is smooth.

2. The tangent vector spaces are complex vector spaces : their introduction above does not require the field to be $\mathbb{R}$. So *on the tangent space* real structures can be defined (see below).

3. A map $f \in C_r\left(M; N\right)$ between complex manifolds M,N modeled on E,G is $\mathbb{R}$-differentiable iff the map $F = \psi_j \circ f \circ \varphi_i^{-1} : E \to G$ is $\mathbb{R}$- differentiable. If F is 1-$\mathbb{C}$-differentiable, it is holomorphic, thus smooth and f itself is said to be holomorphic.

F is $\mathbb{C}$-differentiable iff it is R-differentiable and meets the Cauchy-Riemann conditions on partial derivatives $F_y' = iF_x'$ where y,x refer to any real structure on E.

4. A complex manifold of (complex) dimension 1 is called a Riemann manifold. The compactified (as in topology) of $\mathbb{C}$ is the **Riemann sphere**. Important properties of holomorphic functions stand only when the domain is an open of $\mathbb{C}$. So many of these results (but not all of them) are still valid for maps (such as functions or forms) defined on Riemann manifolds, but not on general complex manifolds. We will not review them as they are in fact very specific (see Schwartz).

18.1.2   Maps on complex manifolds

In the previous parts or sections several theorems address specifically complex vector spaces and holomorphic maps. We give their obvious extensions on manifolds.

Theorem 1533 *A holomorphic map $f : M \to F$ from a finite dimensional connected complex manifold M to a normed vector space F is constant if one of the following conditions is met:*

i) f is constant in an open of M
ii) if $F=\mathbb{C}$ and $\operatorname{Re} f$ or $\operatorname{Im} f$ has an extremum or $|f|$ has a maximum
iii) if M is compact

Theorem 1534 *If M,N are finite dimensional connected complex manifolds, f a holomorphic map $f : M \to N$,if f is constant in an open of M then it is constant in M*

A compact connected finite dimensional complex manifold cannot be an affine submanifold of $\mathbb{C}^n$ because its charts would be constant.

18.1.3 Real structure

A real structure on a complex manifold involves only the tangent bundle, not the manifold itself, which keeps its genuine complex manifold structure.

Theorem 1535 *Any tangent bundle of a manifold modeled on a complex space E admits real structures, defined by a real continuous real stucture on E.*

Proof. If M is a complex manifold with atlas $\left(E, (O_k, \varphi_k)_{k \in K}\right)$ it is always possible to define real structures on E : antilinear maps $\sigma : E \to E$ such that $\sigma^2 = Id_E$ and then define a real kernel $E_\mathbb{R}$ and split any vector u of E in a real and an imaginary part both belonging to the kernel, such that : $u = \operatorname{Re} u + i \operatorname{Im} u$. If E is infinite dimensional we will require σ to be continuous

At any point p$\in O_k$ of M the real structure S_k(p) on $T_p M$ is defined by :
$S_k(p) = \varphi_k'^{-1} \circ \sigma \circ \varphi_k' : T_p M \to T_p M$
This is an antilinear map and $S^2(p) = Id_{T_p M}$.
The real kernel of $T_p M$ is $(T_p M)_\mathbb{R} = \{u_p \in T_p M : S_k(p) u_p = u_p\} = \varphi_k'^{-1}(x)(E_\mathbb{R})$
Indeed : $u \in E_\mathbb{R} : \sigma(u) = u \to u_p = \varphi_k'^{-1}(x) u$
$S_k(p) u_p = \varphi_k'^{-1} \circ \sigma \circ \varphi_k' \circ \varphi_k'^{-1}(x)(u) = \varphi_k'^{-1}(x)(u) = u_p$
At the transitions :
$\sigma = \varphi_k' \circ S_k(p) \circ \varphi_k'^{-1} = \varphi_j' \circ S_j(p) \circ \varphi_j'^{-1}$
$S_j(p) = \left(\varphi_k'^{-1} \circ \varphi_j'\right)^{-1} \circ S_k(p) \circ \left(\varphi_k'^{-1} \circ \varphi_j'\right)$
From the definition of the tangent space $S_j(p), S_k(p)$ give the same map so this definition is intrinsic and we have a map :$S : M \to C(TM; TM)$ such that S(p) is a real structure on $T_p M$. ∎
The tangent bundle splits in a real and an imaginary part : $TM = \operatorname{Re} TM \oplus i \operatorname{Im} TM$
We can define tensors on the product of vector spaces $\left((T_p M)_\mathbb{R} \times (T_p M)_\mathbb{R}\right)^r \otimes \left((T_p M)_\mathbb{R} \times (T_p M)_\mathbb{R}\right)^{*s}$
We can always choose a basis $(e_a)_{a \in A}$ of E such that : $\sigma(e_a) = e_a, \sigma(ie_a) = -e_a$ so that the holonomic basis of the real vector space $E_\sigma = E_\mathbb{R} \oplus E_\mathbb{R}$ reads $(e_a, ie_a)_{a \in A}$.

18.2 Complex structures on real manifolds

There are two ways to build a complex structure on the tangent bundle of a real manifold : the easy way by complexification, and the complicated way by a special map.

18.2.1 Complexified tangent bundle

This is the implementation, in the manifold context, of the general procedure for vector spaces (see Algebra). The tangent vector space at each point p of a real manifold M can be complexified : $T_pM_\mathbb{C} = T_pM \oplus iT_pM$. If M is modeled on the real Banach E, then $T_pM_\mathbb{C}$ is isomorphic ot the complexified of E, by taking the complexified of the derivatives of the charts. This procedure does not change anything to the manifold structure of M, it is similar to the tensorial product : the complexified tangent bundle is $TM_\mathbb{C} = TM \otimes \mathbb{C}$.

A holonomic basis of M is still a holonomic basis in $TM_\mathbb{C}$, the vectors may have complex components.

On $TM_\mathbb{C}$ we can define a complexified tangent bundle, and r forms valued in $\mathbb{C} : \mathfrak{X}(\wedge_r TM_\mathbb{C}^*) = \wedge_r (M; \mathbb{C})$.

All the operations in complex vector space are available at each point p of M. The complexified structure is fully dependent on the tangent bundle, so there is no specific rule for a change of charts. This construct is strictly independant of the manifold structure itself.

However there is another way to define a complex structure on a real vector space, by using a complex structure.

18.2.2 Almost complex structures

Definition 1536 *An **almost complex structure** on a real manifold M is a tensor field $J \in \mathfrak{X}(\otimes_1^1 TM)$ such that $\forall u \in T_pM : J^2(p)(u) = -u$*

Theorem 1537 *A complex structure on a real manifold M defines a structure of complex vector space on each tangent space, and on the tangent bundle. A necessary condition for the existence of a complex structure on a manifold M is that the dimension of M is infinite or even.*

A complex structure defines in each tangent space a map : $J(p) \in \mathcal{L}(T_pM; T_pM)$ such that $J^2(p)(u) = -u$. Such a map is a complex structure on T_pM, it cannot exist if M is finite dimensional with an odd dimension, and otherwise defines, continuously, a structure of complex vector space on each tangent space by : $iu = J(u)$ (see Algebra).

A complex vector space has a canonical orientation. So a manifold endowed with a complex structure is orientable, and one can deduce that there are obstructions to the existence of almost complex structures on a manifold.

A complex manifold has an almost complex structure : J(u)=iu but a real manifold endowed with an almost complex structure does not necessarily admits the structure of a complex manifold. There are several criteria for this purpose.

18.2.3   Kähler manifolds

Definition 1538 *An **almost Kähler manifold** is a real manifold M endowed with a non degenerate bilinear symmetric form g, an almost complex structure J, and such its fundamental 2-form is closed. If M is also a complex manifold then it is a **Kähler manifold**.*

i) It is always possible to assume that J preserves g by defining : $\widehat{g}(p)(u_p, v_p) = \frac{1}{2}(g(p)(u_p, v_p) + g(p)(Ju_p, Jv_p))$ and so assume that : $g(p)(u_p, v_p) = g(p)(Ju_p, Jv_p)$

ii) The **fundamental 2-form** is then defined as :

$$\varpi(p)(u_p, v_p) = g(p)(u_p, Jv_p)$$

This is a 2-form, which is invariant by J and non degenerate if g is non degenerate. It defines a structure of symplectic manifold over M.

19 PSEUDO-RIEMANNIAN MANIFOLDS

So far we have not defined a metric on manifolds. The way to define a metric on the topological space M is to define a differentiable norm on the tangent bundle. If M is a real manifold and the norm comes from a bilinear positive definite form we have a Riemanian manifold, which is the equivalent of an euclidean vector space (indeed M is then modelled on an euclidean vector space). Riemannian manifolds have been the topic of many studies and in the litterature most of the results are given in this context. Unfortunately for the physicists the Universe of General Relativity is not riemannian but modelled on a Minkovski space. Most, but not all, the results stand if there is a non degenerate, but non positive definite, metric on M. So we will strive to stay in this more general context.

19.1 General properties

19.1.1 Definitions

Definition 1539 *A **pseudo-riemannian manifold** (M,g) is a real finite dimensional manifold M endowed with a (0,2) symmetric tensor which induces a bilinear symmetric non degenerate form g on TM.*

Thus g has a signature (+p,-q) with p+q=dimM, and we will say that M is a pseudo-riemannian manifold of signature (p,q).

Definition 1540 *A **riemannian manifold** (M,g) is a real finite dimensional manifold M endowed with a (0,2) symmetric tensor which induces a bilinear symmetric definite positive form g on TM.*

Thus a riemannian manifold is a pseudo riemannian manifold of signature (m,0).

The manifold and g will be assumed to be at least of class 1. In the following if not otherwise specified M is a pseudo-riemannian manifold. It will be specified when a theorem stands for riemannian manifold only.

Any real finite dimensional Hausdorff manifold which is either paracompact or second countable admits a riemannian metric.

Any open subset M of a pseudo-riemannian manifold (N,g) is a a pseudo-riemannian manifold $(M,g|_M)$.

The bilinear form is called a **scalar product**, and an **inner product** if it is definite positive. It is also usually called the **metric** (even if it is not a metric in the topological meaning)

The coordinate expressions are in holonomic bases:

$g \in \otimes_2^0 TM : g(p) = \sum_{\alpha\beta} g_{\alpha\beta}(p) \, dx^\alpha \otimes dx^\beta$

$g_{\alpha\beta} = g_{\beta\alpha}$

$u_p \in T_pM : \forall v_p \in T_pM : g(p)(u_p, v_p) = 0 \Rightarrow u_p = 0 \Leftrightarrow \det[g(p)] \neq 0$

Isomorphism between the tangent and the cotangent bundle

Theorem 1541 *A scalar product g on a finite dimensional real manifold M defines an isomorphim between the tangent space T_pM and the cotangent space T_pM^* at any point, and then an isomorphism $\jmath$ between the tangent bundle TM and the cotangent bundle TM*.*

$j : TM \to TM^* :: u_p \in T_pM \to \mu_p = \jmath(u_p) \in T_pM^* ::$
$\forall v_p \in T_pM : g(p)(u_p, v_p) = \mu_p(v_p)$
g induces a scalar product g* on the cotangent bundle,
$g^*(p)(\mu_p, \lambda_p) = \sum_{\alpha\beta} \mu_{p\alpha} \lambda_{p\beta} g^{\alpha\beta}(p)$
which is a defined by the (2,0) symmetric tensor on M:
$g^* = \sum_{\alpha\beta} g^{\alpha\beta}(p) \partial x_\alpha \otimes \partial x_\beta$
with : $\sum_\beta g^{\alpha\beta}(p) g_{\beta\gamma}(p) = \delta^\alpha_\gamma$ so the matrices of g and g* are inverse from each other : $[g^*] = [g]^{-1}$

For any vector : $u_p = \sum_\alpha u_p^\alpha \partial x_\alpha \in T_pM : \mu_p = \jmath(u_p) = \sum_{\alpha\beta} g_{\alpha\beta} u_p^\beta dx^\alpha$

and conversely : $\mu_p = \sum_\alpha \mu_{p\alpha} dx^\alpha \in T_pM^* \to \jmath^{-1}(\mu_p) = u_p = \sum_{\alpha\beta} g^{\alpha\beta} \mu_\beta \partial x_\alpha$

The operation can be done with any mix tensor. Say that one "lifts" or "lowers" the indices with g.

If $f \in C_1(M; \mathbb{R})$ the gradient of f is the vector field grad(f) such that :
$\forall u \in VM : g(p)(gradf, u) = f'(p)u \Leftrightarrow (gradf)^\alpha = \sum_\beta g^{\alpha\beta} \partial_\beta f$

Orthonormal basis

Theorem 1542 *A pseudo-riemannian manifold admits an* **orthonormal basis** *at each point :*

$\forall p : \exists (e_i)_{i=1}^m, e_i \in T_pM : g(p)(e_i, e_j) = \eta_{ij} = \pm\delta_{ij}$

The coefficients η_{ij} define the signature of the metric, they do not depend on the choice of the orthonormal basis or p. We will denote by $[\eta]$ the matrix η_{ij} so that for any orthonormal basis : $[E] = [e_i^\alpha] :: [E]^t [g] [E] = [\eta]$

Warning ! Even if one can find an orthonomal basis at each point, usually there is no chart such that the holonomic basis is orthonormal at each point. And there is no distribution of m vector fields which are orthonormal at each point if M is not parallelizable.

Volume form

At each point p a volume form is a m-form ϖ_p such that $\varpi_p(e_1, ..., e_m) = +1$ for any orthonormal basis (cf.Algebra). Such a form is given by :
$\varpi_0(p) = \sqrt{|\det g(p)|} dx^1 \wedge dx^2 ... \wedge dx^m$

As it never vanishes, this is a **volume form** (with the meaning used for integrals) on M, and a pseudo-riemanian manifold is orientable if it is the union of countably many compact sets.

Divergence

Theorem 1543 *The **divergence** of a vector field V is the function $div(V) \in C(M; \mathbb{R})$ such that : $\pounds_V \varpi_0 = (divV)\varpi_0$ and*

$$divV = \sum_\alpha \left(\partial_\alpha V^\alpha + V^\alpha \tfrac{1}{2} \sum_{\beta\gamma} g^{\gamma\beta} (\partial_\alpha g_{\beta\gamma}) \right)$$

Proof. $divV = \frac{1}{\varpi_0} \sum_\alpha \partial_\alpha (V^\alpha \varpi_0)$ (see Integral)

So : $divV = \frac{1}{\sqrt{|\det g(p)|}} \sum_\alpha \partial_\alpha \left(V^\alpha \sqrt{|\det g(p)|} \right) = \sum_\alpha \partial_\alpha V^\alpha + V^\alpha \frac{1}{\sqrt{|\det g(p)|}} \partial_\alpha \sqrt{|\det g(p)|}$

$= \sum_\alpha \partial_\alpha V^\alpha + V^\alpha \frac{1}{\sqrt{|\det g(p)|}} \frac{1}{2} \frac{1}{\sqrt{|\det g(p)|}} (-1)^p (\det g) \, Tr \left(\left(\frac{\partial}{\partial x^\alpha} [g] \right) [g]^{-1} \right)$

$= \sum_\alpha \partial_\alpha V^\alpha + V^\alpha \frac{1}{2} Tr \left(\left[\sum_\gamma (\partial_\alpha g_{\beta\gamma}) g^{\gamma\eta} \right] \right) = \sum_\alpha \left(\partial_\alpha V^\alpha + V^\alpha \frac{1}{2} \sum_{\beta\gamma} g^{\gamma\beta} (\partial_\alpha g_{\beta\gamma}) \right)$

∎

Complexification

It is always possible to define a complex structure on the tangent bundle of a real manifold by complexification. The structure of the manifold stays the same, only the tangent bundle is involved.

If (M,g) is pseudo-riemannian then, point wise, g(p) can be extended to a hermitian, sequilinear, non degenerate form $\gamma(p)$:

$\forall u, v \in T_p M : \gamma(p)(u, v) = g(p)(u, v) ; \gamma(p)(iu, v) = -ig(p)(u, v) ; \gamma(p)(u, iv) = ig(p)(u, v)$ (see Algebra).

γ defines a tensor field on the complexified tangent bundle $\mathfrak{X} (\otimes_2 TM^*_\mathbb{C})$. The holonomic basis stays the same (with complex components) and γ has same components as g.

Most of the operations on the complex bundle can be extended, as long as they do not involve the manifold structure itself (such as derivation). We will use it in this section only for the Hodge duality, because the properties will be useful in Functional Analysis. Of course if M is also a complex manifold the extension is straightforward.

On the other hand the extension to infinite dimensional manifolds does not seem promising. One of the key point of pseudo-riemanian manifolds is the isomorphism with the dual, which requires finite dimensional manifolds. In fact Hilbert structures (quite normal for infinite dimensional manifolds) is the best extension.

19.1.2 Hodge duality

Here we use the extension of a symmetric bilinear form g to a hermitian, sequilinear, non degenerate form that we still denote g. The field K is $\mathbb{R}$ or $\mathbb{C}$.

Scalar product of r-forms

This is the direct application of the definitions and results of the Algebra part.

Theorem 1544 *On a finite dimensional manifold (M,g) endowed with a scalar product the map :*

$G_r : \mathfrak{X}(\Lambda_r TM^*) \times \mathfrak{X}(\Lambda_r TM^*) \to K ::$

$G_r(\lambda, \mu) = \sum_{\{\alpha_1..\alpha_r\}\{\beta_1..\beta_r\}} \overline{\lambda}_{\alpha_1..\alpha_r} \mu_{\beta_1...\beta_r} \det\left[g^{-1}\right]^{\{\alpha_1..\alpha_r\},\{\beta_1..\beta_r\}}$

is a non degenerate hermitian form and defines a scalar product which does not depend of the basis.

It is definite positive if g is definite positive

In the matrix $\left[g^{-1}\right]$ one takes the elements $g^{\alpha_k \beta_l}$ with $\alpha_k \in \{\alpha_1..\alpha_r\}, \beta_l \in \{\beta_1..\beta_r\}$

$G_r(\lambda, \mu) = \sum_{\{\alpha_1...\alpha_r\}} \overline{\lambda}_{\{\alpha_1...\alpha_r\}} \sum_{\beta_1...\beta_r} g^{\alpha_1\beta_1}...g^{\alpha_r\beta_r} \mu_{\beta_1...\beta_r} = \sum_{\{\alpha_1...\alpha_r\}} \overline{\lambda}_{\{\alpha_1...\alpha_r\}} \mu^{\{\beta_1\beta_2...\beta_r\}}$

where the indexes are lifted and lowered with g.

The result does not depend on the basis.

Proof. In a change of charts for a r-form :

$\lambda = \sum_{\{\alpha_1...\alpha_r\}} \lambda_{\alpha_1...\alpha_r} dx^{\alpha_1} \wedge dx^{\alpha_2} \wedge ... \wedge dx^{\alpha_r} = \sum_{\{\alpha_1...\alpha_r\}} \widehat{\lambda}_{\alpha_1...\alpha_r} dy^{\alpha_1} \wedge dy^{\alpha_2} \wedge ... \wedge dy^{\alpha_r}$

with $\widehat{\lambda}_{\alpha_1...\alpha_r} = \sum_{\{\beta_1,...,\beta_r\}} \lambda_{\beta_1...\beta_r} \det\left[J^{-1}\right]^{\beta_1...\beta_r}_{\alpha_1...\alpha_r}$

where $\det\left[J^{-1}\right]^{\beta_1...\beta_r}_{\alpha_1...\alpha_r}$ is the determinant of the matrix $\left[J^{-1}\right]$ with elements row β_k column α_l

$G_r(\lambda, \mu) = \sum_{\{\alpha_1..\alpha_r\}\{\beta_1..\beta_r\}} \overline{\widehat{\lambda}}_{\alpha_1..\alpha_r} \widehat{\mu}_{\beta_1...\beta_r} \det\left[\widehat{g}^{-1}\right]^{\{\alpha_1..\alpha_r\},\{\beta_1..\beta_r\}}$

$= \sum_{\{\alpha_1..\alpha_r\}\{\beta_1..\beta_r\}} \sum_{\{\gamma_1,...,\gamma_r\}} \overline{\lambda}_{\gamma_1...\gamma_r} \det\left[J^{-1}\right]^{\gamma_1...\gamma_r}_{\alpha_1...\alpha_r}$

$\times \sum_{\{\eta_1,...,\eta_r\}} \widehat{\mu}_{\eta_1...\eta_r} \det\left[J^{-1}\right]^{\eta_1...\eta_r}_{\beta_1...\beta_r} \det\left[\widehat{g}^{-1}\right]^{\{\alpha_1..\alpha_r\},\{\beta_1..\beta_r\}}$

$= \sum_{\{\gamma_1,...,\gamma_r\}\{\eta_1,...,\eta_r\}} \overline{\lambda}_{\gamma_1...\gamma_r} \widehat{\mu}_{\eta_1...\eta_r}$

$\times \sum_{\{\alpha_1..\alpha_r\}\{\beta_1..\beta_r\}} \det\left[J^{-1}\right]^{\gamma_1...\gamma_r}_{\alpha_1...\alpha_r} \det\left[J^{-1}\right]^{\eta_1...\eta_r}_{\beta_1...\beta_r} \det\left[\widehat{g}^{-1}\right]^{\{\alpha_1..\alpha_r\},\{\beta_1..\beta_r\}}$

$\widehat{g}_{\alpha\beta} = \left[J^{-1}\right]^{\gamma}_{\alpha} \left[J^{-1}\right]^{\eta}_{\beta} g_{\gamma\eta}$

$\det\left[\widehat{g}^{-1}\right]^{\{\alpha_1..\alpha_r\},\{\beta_1..\beta_r\}} = \det\left[g^{-1}\right]^{\{\gamma_1..\gamma_r\},\{\eta_1..\eta_r\}} \det[J]^{\alpha_1...\alpha_r}_{\gamma_1...\gamma_r} \det[J]^{\beta_1...\beta_r}_{\eta_1...\eta_r}$

■

In an orthonormal basis : $G_r(\lambda, \mu) = \sum_{\{\alpha_1..\alpha_r\}\{\beta_1..\beta_r\}} \overline{\lambda}_{\alpha_1..\alpha_r} \mu_{\beta_1...\beta_r} \eta^{\alpha_1\beta_1}...\eta^{\alpha_r\beta_r}$

For $r = 1$ one gets the usual bilinear symmetric form over $\mathfrak{X}(\otimes^0_1 TM)$:
$G_1(\lambda, \mu) = \sum_{\alpha\beta} \overline{\lambda}_\alpha \mu_\beta g^{\alpha\beta}$

For r=m : $G_m(\lambda, \mu) = \overline{\lambda}\mu (\det g)^{-1}$

Theorem 1545 *For a 1 form π fixed in $\mathfrak{X}(\Lambda_1 TM^*)$, the map :*

$\lambda(\pi) : \mathfrak{X}(\Lambda_r TM^*) \to \mathfrak{X}(\Lambda_{r+1} TM^*) :: \lambda(\pi)\mu = \pi \wedge \mu$

has an adjoint with respect to the scalar product of forms : $G_{r+1}(\lambda(\pi)\mu, \mu') = G_r(\mu, \lambda^*(\pi)\mu')$ which is

$\lambda^*(\pi) : \mathfrak{X}(\Lambda_r TM^*) \to \mathfrak{X}(\Lambda_{r-1} TM^*) :: \lambda^*(\pi)\mu = i_{grad\pi}\mu$

It suffices to compute the two quantities.

Hodge duality

Theorem 1546 *On a m dimensional manifold (M,g) endowed with a scalar product, with the volume form ϖ_0 the map :*

$* : \mathfrak{X}(\Lambda_r TM^*) \to \mathfrak{X}(\Lambda_{m-r} TM^*)$ *defined by the condition*

$\forall \mu \in \mathfrak{X}(\Lambda_r TM^*) : *\lambda_r \wedge \mu = G_r(\lambda, \mu) \varpi_0$

is an anti-isomorphism

A direct computation gives the value of the **Hodge dual** $*\lambda$ in a holonomic basis :

$* \left(\sum_{\{\alpha_1...\alpha_r\}} \lambda_{\{\alpha_1...\alpha_r\}} dx^{\alpha_1} \wedge .. \wedge dx^{\alpha_r} \right)$

$= \sum_{\{\alpha_1..\alpha_{n-r}\}\{\beta_1..\beta_r\}} \epsilon(\beta_1..\beta_r, \alpha_1, ...\alpha_{m-r}) \overline{\lambda}^{\beta_1...\beta_r} \sqrt{|\det g|} dx^{\alpha_1} \wedge dx^{\alpha_2}... \wedge dx^{\alpha_{m-r}}$

With $\epsilon = sign \det[g]$ (which is always real)

For r=0:

$*\lambda = \overline{\lambda} \varpi_0$

For r=1 :

$* \left(\sum_\alpha \lambda_\alpha dx^\alpha \right) = \sum_{\beta=1}^m (-1)^{\beta+1} g^{\alpha\beta} \overline{\lambda}_\beta \sqrt{|\det g|} dx^1 \wedge ..\widehat{dx^\beta} \wedge ... \wedge dx^m$

For r=m-1:

$* \left(\sum_{\alpha=1}^m \lambda_{1..\widehat{\alpha}...m} dx^1 \wedge ..\widehat{dx^\alpha} \wedge ... \wedge dx^m \right) = \sum_{\alpha=1}^m (-1)^{\alpha-1} \overline{\lambda}^{1..\widehat{\alpha}...n} \sqrt{|\det g|} dx^\alpha$

For r=n:

$* \left(\lambda dx^1 \wedge .... \wedge dx^m \right) = \epsilon \frac{1}{\sqrt{|\det g|}} \overline{\lambda}$

Theorem 1547 *The inverse of the map * is :*

$*^{-1} : \mathfrak{X}(\Lambda_r TM^*) \to \mathfrak{X}(\Lambda_{m-r} TM^*) :: *^{-1}\lambda_r = \epsilon(-1)^{r(n-r)} * \lambda_r \Leftrightarrow * * \lambda_r = \epsilon(-1)^{r(n-r)} \lambda_r$

$G_q(\lambda, *\mu) = G_{n-q}(*\lambda, \mu)$

$G_{n-q}(*\lambda, *\mu) = G_q(\lambda, \mu)$

Codifferential

Definition 1548 *On a m dimensional manifold (M,g) endowed with a scalar product, with the volume form ϖ_0, the **codifferential** is the operator :*

$\delta : \mathfrak{X}(\Lambda_{r+1} TM^*) \to \mathfrak{X}(\Lambda_r TM^*) :: \delta\lambda = \epsilon(-1)^{r(m-r)+r} *d* \lambda = (-1)^r *d*^{-1}\lambda$

where $\epsilon = (-1)^p$ with p the number of - in the signature of g

2. It has the following properties :

$\delta^2 = 0$

For $f \in C(M; \mathbb{R}) : \delta f = 0$

For $\lambda_r \in \mathfrak{X}(\Lambda_{r+1} TM^*) : *\delta\lambda = (-1)^{m-r-1} d * \lambda$

$(\delta\lambda)_{\{\gamma_1..\gamma_{r-1}\}}$

$= \epsilon(-1)^{r(m-r)} \sqrt{|\det g|} \sum_{\{\eta_1..\eta_{m-r+1}\}} \epsilon(\eta_1..\eta_{m-r+1}, \gamma_1, ...\gamma_{r-1}) \sum g^{\eta_1\beta_1}...g^{\eta_{m-r+1}\beta_{m-r+1}} \times$

$$\sum_{k=1}^{r+1}(-1)^{k-1}\sum_{\{\alpha_1..\alpha_r\}}\epsilon\left(\alpha_1..\alpha_r,\beta_1,.\widehat{\beta}_k..\beta_{m-r+1}\right)\partial_{\beta_k}\left(\lambda^{\alpha_1...\alpha_r}\sqrt{|\det g|}\right)$$

For r=1 : $\delta\left(\sum_i\lambda_\alpha dx^\alpha\right)=(-1)^m\frac{1}{\sqrt{|\det g|}}\sum_{\alpha,\beta=1}^m\partial_\alpha\left(g^{\alpha\beta}\lambda_\beta\sqrt{|\det g|}\right)$

Proof. $\delta\left(\sum_i\lambda_\alpha dx^\alpha\right)=\epsilon(-1)^m*d*\left(\sum_i\lambda_\alpha dx^\alpha\right)$

$=\epsilon(-1)^m*d\left(\sum_{\alpha=1}^m(-1)^{\alpha+1}g^{\alpha\beta}\overline{\lambda_\beta}\sqrt{|\det g|}dx^1\wedge..\widehat{dx^\alpha}\wedge...\wedge dx^m\right)$

$=\epsilon(-1)^m*\sum_{\alpha,\beta,\gamma=1}^m(-1)^{\alpha+1}\partial_\beta\left(g^{\alpha\gamma}\overline{\lambda_\gamma}\sqrt{|\det g|}\right)dx^\beta\wedge dx^1\wedge..\widehat{dx^\alpha}\wedge...\wedge dx^m$

$=\epsilon(-1)^m*\sum_{\alpha,\beta=1}^m\partial_\alpha\left(g^{\alpha\beta}\overline{\lambda_\beta}\sqrt{|\det g|}\right)dx^1\wedge....\wedge dx^m$

$=(-1)^m\frac{1}{\sqrt{|\det g|}}\sum_{\alpha,\beta=1}^m\partial_\alpha\left(g^{\alpha\beta}\lambda_\beta\sqrt{|\det g|}\right)$ ∎

The codifferential is the adjoint of the exterior derivative with respect to the interior product G_r (see Functional analysis).

Laplacian

Definition 1549 *On a m dimensional manifold (M,g) endowed with a scalar product the **Laplace-de Rahm** operator is :*

$\Delta:\mathfrak{X}\left(\Lambda_rTM^*\right)\to\mathfrak{X}\left(\Lambda_rTM^*\right)::\Delta=-\left(\delta d+d\delta\right)=-\left(d+\delta\right)^2$

Remark : one finds also the definition $\Delta=\left(\delta d+d\delta\right)$.
Properties : see Functional analysis

19.1.3   Isometries

The isometries play a specific role in that they define the symmetries of the manifold.

Definition 1550 *A class 1 map $f:M\to N$ between the pseudo-riemannian manifolds (M,g),(N,h) is **isometric** at $p\in M$ if :*

$\forall u_p,v_p\in T_pM:h\left(f\left(p\right)\right)\left(f'(p)u_p,f'(p)v_p\right)=g\left(p\right)\left(u_p,v_p\right)$

Then f'(p) is injective. If f is isometric on M this is an immersion, and if it is bijective, this an embedding.

Definition 1551 *An **isometry** is a class 1 bijective map on the pseudo-riemannian manifolds (M,g) which is isometric for all p in M*

(Kobayashi I p.162)
An isometric immersion maps geodesics to geodesics
An isometry maps orthonormal bases to orthonormal bases.

Killing vector fields

Definition 1552 *A **Killing vector field** is a vector field on a pseudo-riemannian manifold which is the generator of a one parameter group of isometries.*

For any t in its domain of definition, $\Phi_V(t, p)$ is an isometry on M.

A Killing vector field is said to be complete if its flow is complete (defined over all $\mathbb{R}$).

Theorem 1553 *(Kobayashi I p.237) For a vector field V on a pseudo-riemannian manifold (M,g) the followings are equivalent :*

i) V is a Killing vector field

ii) £$_V g = 0$

iii) $\forall Y, Z \in \mathfrak{X}(TM) : g((\pounds_V - \nabla_V) Y, Z) = -g((\pounds_V - \nabla_V) Z, Y)$ where ∇ is the Levy-Civita connection

iv) $\forall \alpha, \beta : \sum_\gamma (g_{\gamma\beta} \partial_\alpha V^\gamma + g_{\alpha\gamma} \partial_\beta V^\gamma + V^\gamma \partial_\gamma g_{\alpha\beta}) = 0$

Theorem 1554 *(Wald p.442) If V is a Killing vector field and $c'(t)$ the tangent vector to a geodesic then $g(c(t))(c'(t), V) = Cte$*

Group of isometries

Theorem 1555 *(Kobayashi I p.238) The set of vector fields $\mathfrak{X}(M)$ over a m dimensional real pseudo-riemannian manifold M has a structure of Lie algebra (infinite dimensional) with the commutator as bracket. The set of Killing vector fields is a subalgebra of dimension at most equal to m(m+1)/2. If it is equal to m(m+1)/2 then M is a space of constant curvature.*

The set I(M) of isomometries over M, endowed with the compact-open topology (see Topology), is a Lie group whose Lie algebra is naturally isomorphic to the Lie algebra of all complete Killing vector fields (and of same dimension). The isotropy subgroup at any point is compact. If M is compact then the group I(M) is compact.

19.2   Lévi-Civita connection

On a pseudo-riemannian manifold there is a "natural" connection, called the Lévy-Civita connection.

19.2.1   Definitions

Metric connection

Definition 1556 *A covariant derivative ∇ on a pseudo-riemannian manifold (M,g) is said to be* ***metric*** *if $\nabla g = 0$*

Then we have similarly for the metric on the cotangent bundle : $\nabla g^* = 0$

So : $\forall \alpha, \beta, \gamma : \nabla_\gamma g_{\alpha\beta} = \nabla_\gamma g^{\alpha\beta} = 0$

By simple computation we have the theorem :

Theorem 1557 *For a covariant derivative ∇ on a pseudo-riemannian manifold (M,g) the following are equivalent :*

i) the covariant derivative is metric

ii) $\forall \alpha, \beta, \gamma : \partial_\gamma g_{\alpha\beta} = \sum_\eta \left(g_{\alpha\eta} \Gamma^\eta_{\gamma\beta} + g_{\beta\eta} \Gamma^\eta_{\gamma\alpha} \right)$

iii) the covariant derivative preserves the scalar product of transported vectors

iv) the riemann tensor is such that :
$\forall X, Y, Z \in \mathfrak{X}(TM) : R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0$

Lévi-Civita connection

Theorem 1558 *On a pseudo-riemannian manifold (M,g) there is a unique affine connection, called the **Lévi-Civita connection**, which is both torsion free and metric. It is fully defined by the metric, through the relations :*

$\Gamma^\alpha_{\beta\gamma} = \frac{1}{2} \sum_\eta (g^{\alpha\eta} (\partial_\beta g_{\gamma\eta} + \partial_\gamma g_{\beta\eta} - \partial_\eta g_{\beta\gamma}) = -\frac{1}{2} \sum_\eta (g_{\gamma\eta} \partial_\beta g^{\alpha\eta} + g_{\beta\eta} \partial_\gamma g^{\alpha\eta} + g^{\alpha\eta} \partial_\eta g_{\beta\gamma})$

There are many demonstrations, which is a straightforward computation, of this theorem for riemannian connection. For a connection of any signature see Wald.

Warning ! these are the most common definitions for Γ, but they can vary (mainly in older works) according to the definitions of the Christofell symbols. Those above are fully consistent with all other definitions used in this book.

19.2.2 Curvature

With the Lévi-Civita connection many formulas take a simple form :

Theorem 1559 $\Gamma_{\alpha\beta\gamma} = \sum_\eta g_{\alpha\eta} \Gamma^\eta_{\beta\gamma} = \frac{1}{2}(\partial_\beta g_{\gamma\alpha} + \partial_\gamma g_{\beta\alpha} - \partial_\alpha g_{\beta\gamma})$

$\partial_\alpha g_{\beta\gamma} = \sum_\eta g_{\gamma\eta} \Gamma^\eta_{\alpha\beta} + g_{\beta\eta} \Gamma^\eta_{\alpha\gamma}$

$\partial_\alpha g^{\beta\gamma} = -\sum_\eta g^{\beta\eta} \Gamma^\gamma_{\alpha\eta} + g^{\gamma\eta} \Gamma^\beta_{\alpha\eta}$

$\sum_\gamma \Gamma^\gamma_{\gamma\alpha} = \frac{1}{2} \frac{\partial_\alpha |\det g|}{|\det g|} = \frac{\partial_\alpha \left(\sqrt{|\det g|} \right)}{\sqrt{|\det g|}}$

Proof. $\frac{d}{d\xi^\alpha} \det g = \left(\frac{d}{dg_{\lambda\mu}} \det g \right) \frac{d}{d\xi^\alpha} g_{\lambda\mu} = g^{\mu\lambda} (\det g) \frac{d}{d\xi^\alpha} g_{\lambda\mu} = (\det g) g^{\mu\lambda} \partial_\alpha g_{\lambda\mu}$

$g^{\lambda\mu} (\partial_\alpha g_{\lambda\mu}) = g^{\lambda\mu} \left(g_{\mu l} \Gamma^l_{\alpha\lambda} + g_{\lambda l} \Gamma^l_{\alpha\mu} \right) = \left(g^{\lambda\mu} g_{\mu l} \Gamma^l_{\alpha\lambda} + g^{\lambda\mu} g_{\lambda l} \Gamma^l_{\alpha\mu} \right) = \left(\Gamma^\lambda_{\alpha\lambda} + \Gamma^\mu_{\alpha\mu} \right) = 2\Gamma^\gamma_{\gamma\alpha}$

$\partial_\alpha \det g = 2 (\det g) \Gamma^\gamma_{\gamma\alpha}$

$\frac{d}{d\xi^\alpha} \sqrt{|\det g|} = \frac{d}{d\xi^\alpha} \sqrt{|\det g|} = \frac{1}{2\sqrt{|\det g|}} \frac{d}{d\xi^\alpha} |\det g| = \frac{1}{2\sqrt{|\det g|}} \left(2 (-1)^p (\det g) \Gamma^\gamma_{\gamma\alpha} \right) = \sqrt{|\det g|} \Gamma^\gamma_{\gamma\alpha}$ ∎

Theorem 1560 $div(V) = \sum_\alpha \nabla_\alpha V^\alpha$

Proof. $div(V) = \sum_\alpha \partial_\alpha V^\alpha + V^\alpha \frac{1}{\sqrt{|\det g(p)|}} \partial_\alpha \sqrt{|\det g(p)|} = \sum_\alpha \partial_\alpha V^\alpha + V^\alpha \sum_\beta \Gamma^\beta_{\beta\alpha} =$
$\sum_\alpha \nabla_\alpha V^\alpha$ ∎

Theorem 1561 *For the volume form :* $\nabla \varpi_0 = 0$

Proof. $(\nabla \varpi_{12...n})_\alpha = \frac{\partial \varpi_{12...n}}{\partial x^\alpha} - \sum_{\beta=1}^n \sum_{l=1}^n \Gamma^\beta_{\alpha l} \varpi_{12..l-1,\beta,l+1...n}$
$= \partial_\alpha \sqrt{|\det g|} - \sum_{l=1}^n \Gamma^l_{\alpha l} \varpi_{12...n} = \partial_\alpha \sqrt{|\det g|} - \varpi_{12...n} \frac{1}{2} \frac{1}{\det g} \partial_\alpha \det g$
$= \partial_\alpha \sqrt{|\det g|} - \frac{1}{2} \sqrt{|\det g|} \frac{1}{\det g} \partial_\alpha (\det g)$
$|\det g| = (-1)^p \det g$
$(\nabla \varpi)_\alpha = \left(\frac{1}{2} \frac{(-1)^p \partial_\alpha \det g}{\sqrt{|\det g|}} - \frac{1}{2} \sqrt{|\det g|} \frac{1}{\det g} \partial_\alpha \det g \right) = \frac{1}{2} (\partial_\alpha \det g) \left(\frac{(-1)^p}{\sqrt{|\det g|}} - \frac{(-1)^p}{\sqrt{|\det g|}} \right) =$
0 ∎

Theorem 1562 $\partial_\beta \varpi_0 = \frac{1}{2} \frac{\partial_\beta |\det g|}{\varpi_0}$

Theorem 1563 $\sum_\gamma \Gamma^\gamma_{\gamma\alpha} = \frac{\partial_\alpha \varpi_0}{\varpi_0}$

The Riemann curvature is in the general case (see above) :
$R(X,Y,Z) = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z \in MV$
$R = \sum_{\{\gamma\eta\}} \sum_{\alpha\beta} R^\alpha_{\gamma\eta\beta} dx^\gamma \wedge dx^\eta \otimes dx^\beta \otimes \partial x_\alpha$ with $R^\alpha_{\gamma\eta\beta} = \partial_\gamma \Gamma^\alpha_{\eta\beta} - \partial_\eta \Gamma^\alpha_{\gamma\beta} +$
$\sum_\varepsilon \left(\Gamma^\alpha_{\gamma\varepsilon} \Gamma^\varepsilon_{\eta\beta} - \Gamma^\alpha_{\eta\varepsilon} \Gamma^\varepsilon_{\gamma\beta} \right)$
For the Levi-Civitta connection the Riemann tensor, is :
$R^l_{ijk} = \frac{1}{2} \sum_\eta \{ g^{l\eta} \left(\partial_i \partial_k g_{j\eta} - \partial_i \partial_\eta g_{jk} - \partial_j \partial_k g_{i\eta} + \partial_j \partial_\eta g_{ik} \right)$
$+ \left(\partial_i g^{l\eta} + \frac{1}{2} g^{l\varepsilon} g^{s\eta} \left(\partial_i g_{s\varepsilon} + \partial_s g_{i\varepsilon} - \partial_\varepsilon g_{is} \right) \right) \left(\partial_j g_{k\eta} + \partial_k g_{j\eta} - \partial_\eta g_{jk} \right)$
$- \left(\partial_j g^{l\eta} - \frac{1}{2} g^{l\varepsilon} g^{s\eta} \left(\partial_j g_{s\varepsilon} + \partial_s g_{j\varepsilon} - \partial_\varepsilon g_{js} \right) \right) \left(\partial_i g_{k\eta} + \partial_k g_{i\eta} - \partial_\eta g_{ik} \right) \}$
It has the properties :

Theorem 1564 *(Wald p.39)*
$R^l_{ijk} + R^l_{jki} + R^l_{kij} = 0$
Bianchi's identity : $\nabla_a R^e_{bcd} + \nabla_c R^e_{abd} + \nabla_b R^e_{acd} = 0$
$R_{ijkl} = g_{im} R^m_{jkl}$
$R_{ijkl} = -g_{im} R^m_{kjl} = -R_{ikjl}$
$g_{ml} R^l_{ijk} + g_{ml} R^l_{jki} + g_{ml} R^l_{kij} = 0 \Leftrightarrow R_{mijk} + R_{mjki} + R_{mkij} = 0$
$\left(\nabla (\nabla g_{ij})_\beta \right)_\alpha - \left(\nabla (\nabla g_{ij})_\alpha \right)_\beta = R^\varepsilon_{\alpha\beta i} g_{\varepsilon j} + R^\varepsilon_{\alpha\beta j} g_{i\varepsilon} = 0 = R_{j\alpha\beta i} + R_{i\alpha\beta j} \Rightarrow$
$R_{\alpha\beta\gamma\delta} = -R_{\delta\beta\gamma\alpha}$
$\nabla (R^e_{bcd} - R^e_{cbd})_a + \nabla (R^e_{abd} - R^e_{bad})_c + \nabla (R^e_{acd} - R^e_{cad})_b = 0$

The Weyl's tensor is C such that :
$R_{\alpha\beta\gamma\delta} = C_{\alpha\beta\gamma\delta} + \frac{2}{n-2} \left(g_{\alpha[\gamma} R_{\delta]\beta} - g_{\beta[\gamma} R_{\delta]\alpha} \right) - \frac{2}{(n-1)(n-2)} R g_{\alpha[\gamma} g_{\delta]\beta}$

The Ricci tensor is :

$\text{Ric} = \sum_{\alpha\gamma} Ric_{\alpha\gamma} dx^\alpha \otimes dx^\gamma$

$Ric_{\alpha\gamma} = \sum_\beta R^\beta_{\alpha\beta\gamma} = \sum_\eta \left(\partial_\alpha \Gamma^\eta_{\eta\gamma} - \partial_\eta \Gamma^\eta_{\alpha\gamma} + \sum_\varepsilon \left(\Gamma^\eta_{\alpha\varepsilon} \Gamma^\varepsilon_{\eta\gamma} - \Gamma^\eta_{\eta\varepsilon} \Gamma^\varepsilon_{\alpha\gamma} \right) \right)$

So it is symmetric : $Ric_{\alpha\gamma} = Ric_{\gamma\alpha}$

Definition 1565 *The **scalar curvature** is* :$R = \sum_{\alpha\beta} g^{\alpha\beta} Ric_{\alpha\beta} \in C(M;\mathbb{R})$

19.2.3 Sectional curvature

(Kobayashi I p.200)

Definition 1566 *On a pseudo-riemannian manifold (M,g) the **sectional curvature** $K(p)$ at $p \in M$ is the scalar : $K(p) = g(p)(R(u_1, u_2, u_2), u_1)$ where u_1, u_2 are two orthonormal vectors in $T_p M$.*

K(p) depends only on the plane P spanned by u_1, u_2.

Definition 1567 *If K(p) is constant for all planes P in p, and for all p in M, then M is called a **space of constant curvature**.*

So there are positive (resp.negative) curvature according to sign of K(p).

Then :

$R(X, Y, Z) = K(g(Z,Y)X - g(Z,X)Y)$

$R^\alpha_{\beta\gamma\eta} = K\left(\delta^\alpha_\gamma g_{\beta\eta} - \delta^\alpha_\eta g_{\beta\gamma} \right)$

If M is connected, with dimension >2, and K(p) depends only on p (and not P), then M is a space of constant curvature.

Definition 1568 *A Riemannian manifold whose Levi-Civita connection is flat (the torsion and the riemann tensor vanish on M) is said to be locally euclidean.*

Definition 1569 *An Einstein manifold is a pseudo-riemannian manifold whose Ricci tensor is such that : $R_{\alpha\beta}(p) = \lambda(p) g_{\alpha\beta}$*

Then $R = Cte ; \lambda = Cte$

19.2.4 Geodesics

Geodesics can be defined by different ways :

- by a connection, as the curves whose tangent is paralllel transported
- by a metric on a topological space
- by a scalar product on a pseudo-riemannian manifold

The three concepts are close, but not identical. In particular geodesics in pseudo-riemannian manifolds have special properties used in General Relativity.

Length of a curve

A curve C is a 1 dimensional submanifold of a manifold M defined by a class 1 path : $c : [a, b] \rightarrow M$ such that c'(t)$\neq$ 0.

The volume form on C induced by a scalar product on (M,g) is

$\lambda(t) = \sqrt{|g(c(t))(c'(t), c'(t))|} dt = c_* \varpi_0$.

So the "volume" of C is the length of the curve :

$\ell(a, b) = \int_{[a,b]} \varpi_0 = \int_C c_* \varpi_0 = \int_a^b \sqrt{|g(c(t))(c'(t), c'(t))|} dt$, which is always finite because the function is the image of a compact by a continuous function, thus bounded. And it does not depend on the parameter.

The sign of $g(c(t))(c'(t), c'(t))$ does not depend of the parameter, it is >0 if M is Riemannian, but its sign can change on c if not. If, for $t \in [a, b]$:

$g(c(t))(c'(t), c'(t)) < 0$ we say that the path is time like

$g(c(t))(c'(t), c'(t)) > 0$ we say that the path is space like

$g(c(t))(c'(t), c'(t)) = 0$ we say that the path is light like

Theorem 1570 *In a pseudo-riemannian manifold (M,g), the curve of external length, among all the class 1 path from p to q in M of the same type, is a geodesic for the Lévy-Civita connection.*

Proof. So we restrict ourselves to the set P of paths $c \in C_1([a, b]; M)$ such that $g(c(t))(c'(t), c'(t))$ has a constant sign ($\epsilon = +1$ or -1), and c(a)=p,c(b)=q. At first a,b are fixed.

To find an extremal curve is a problem of variational calculus.

The space $C_1([a, b]; M)$ is a Banach, because [a,b] is compact. So the subset of P : P^+ such that $g(c(t))(c'(t), c'(t)) > 0$ and P^- such that $g(c(t))(c'(t), c'(t)) < 0$ are open in this space.

The map c in $P^{+/-}$ for which the functionnal : $\ell(a, b) = \int_a^b \sqrt{\epsilon g(c(t))(c'(t), c'(t))} dt$ is extremum is such that the derivative vanishes.

The Euler-Lagrange equations give with c'(t)=$\sum_\alpha u^\alpha(t) \partial x_\alpha$:

For α : $\frac{1}{2} \frac{\epsilon(\partial_\alpha g_{\beta\gamma}) u^\gamma u^\beta}{\sqrt{\epsilon g_{\lambda\mu} u^\lambda u^\mu}} - \frac{d}{dt} \left(\frac{\epsilon g_{\alpha\beta} u^\beta}{\sqrt{\epsilon g_{\lambda\mu} u^\lambda u^\mu}} \right) = 0$

Moreover the function : $L = \sqrt{\epsilon g(c(t))(c'(t), c'(t))}$ is homogeneous of degree 1 so we have the integral $\sqrt{\epsilon g(c(t))(c'(t), c'(t))} = \theta = Ct$

The equations become :

$\frac{1}{2}(\partial_\alpha g_{\beta\gamma}) u^\gamma u^\beta = \frac{d}{dt}(g_{\alpha\beta} u^\beta) = \left(\frac{d}{dt} g_{\alpha\beta}\right) u^\beta + g_{\alpha\beta} \frac{d}{dt} u^\beta$

using : $\partial_\alpha g_{\beta\gamma} = g_{\eta\gamma} \Gamma^\eta_{\alpha\beta} + g_{\beta\eta} \Gamma^\eta_{\alpha\gamma}$ and $\frac{du^\beta}{dt} = u^\gamma \partial_\gamma u^\beta, \frac{dg_{\alpha\beta}}{dt} = (\partial_\gamma g_{\alpha\beta}) u^\gamma = \left(g_{\eta\beta} \Gamma^\eta_{\gamma\alpha} + g_{\alpha\eta} \Gamma^\eta_{\gamma\beta}\right) u^\gamma$

$\frac{1}{2}\left(g_{\eta\gamma} \Gamma^\eta_{\alpha\beta} + g_{\beta\eta} \Gamma^\eta_{\alpha\gamma}\right) u^\gamma u^\beta = \left(g_{\eta\beta} \Gamma^\eta_{\gamma\alpha} + g_{\alpha\eta} \Gamma^\eta_{\gamma\beta}\right) u^\gamma u^\beta + g_{\alpha\beta} u^\gamma \partial_\gamma u^\beta$

$g_{\eta\gamma}(\nabla_\alpha u^\eta - \partial_\alpha u^\eta) u^\gamma + g_{\beta\eta}(\nabla_\alpha u^\eta - \partial_\alpha u^\eta) u^\beta = 2\left(g_{\eta\beta}(\nabla_\alpha u^\eta - \partial_\alpha u^\eta) + g_{\alpha\eta}(\nabla_\beta u^\eta - \partial_\beta u^\eta)\right) u^\beta + 2 g_{\alpha\beta} u^\gamma \partial_\gamma u^\beta$

$-2 g_{\alpha\gamma} u^\beta \nabla_\beta u^\gamma = 0$

$\nabla_u u = 0$

Thus a curve with an extremal length must be a geodesic, with an affine parameter.

If this an extremal for [a,b] it will be an extremal for any other affine parameter. ∎

Theorem 1571 *The quantity $g\left(c\left(t\right)\right)\left(c'(t),c'(t)\right)$ is constant along a geodesic with an affine parameter.*

Proof. Let us denote : $\theta(t) = g_{\alpha\beta}u^\alpha u^\beta$

$\dfrac{d\theta}{dt} = \left(\partial_\gamma\theta\right)u^\gamma = \left(\nabla_\gamma\theta\right)u^\gamma = u^\alpha u^\beta u^\gamma\left(\nabla_\gamma g_{\alpha\beta}\right) + g_{\alpha\beta}(\nabla_\gamma u^\alpha)u^\beta u^\gamma + g_{\alpha\beta}(\nabla_\gamma u^\beta)u^\alpha u^\gamma = 2g_{\alpha\beta}(\nabla_\gamma u^\alpha)u^\beta u^\gamma = 0$ ∎

So a geodesic is of constant type, which is defined by its tangent at any point. If there is a geodesic joining two points it is unique, and its type is fixed by its tangent vector at p. Moreover at any point p has a convex neighborhood n(p).

To sum up :

Theorem 1572 *On a pseudo-riemannian manifold M, any point p has a convex neighborhood n(p) in which the points q can be sorted according to the fact that they can be reached by a geodesic which is either time like, space like or light like. This geodesic is unique and is a curve of extremal length among the curves for which $g\left(c\left(t\right)\right)\left(c'(t),c'(t)\right)$ has a constant sign.*

Remarks :

i) there can be curves of extremal length such that $g\left(c\left(t\right)\right)\left(c'(t),c'(t)\right)$ has not a constant sign.

ii) one cannot say if the length is a maximum or a minimum

iii) as a loop cannot be a geodesic the relation p,q are joined by a geodesic is not an equivalence relation, and therefore does not define a partition of M (we cannot have $p \sim p$).

General relativity context

In the context of general relativity (meaning g is of Lorentz type and M is four dimensional) the set of time like vectors at any point is disconnected, so it is possible to distinguish future oriented and past oriented time like vectors. The manifold is said to be **time orientable** if it is possible to make this distinction in a continuous manner all over M.

The future of a given point p is the set I(p) of all points q in p which can be reached from p by a curve whose tangent is time like, future oriented.

There is a theorem saying that, in a convex neighborhood n(p) of p, I(p) consists of all the points which can be reached by future oriented geodesic staying in n(p).

In the previous result we could not exclude that a point q reached by a space like geodesic could nevertheless be reached by a time like curve (which cannot be a geodesic). See Wald p.191 for more on the subject.

Riemannian manifold

Theorem 1573 *(Kobayashi I p. 168) For a riemannian manifold (M,g) the map : $d : M \times M \to \mathbb{R}_+ :: d(p,q) = \min_c \ell(p,q)$ for all the piecewise class 1 paths from p to q is a metric on M, which defines a topology equivalent to the topology of M. If the length of a curve between p and q is equal to d, this is a geodesic. Any maping f in M which preserves d is an isometry.*

Theorem 1574 *(Kobayashi I p. 172) For a connected Riemannian manifold M, the following are equivalent :*
  i) the geodesics are complete (defined for t$\in \mathbb{R}$)
  ii) M is a complete topological space with regard to the metric d
  iii) every bounded (with d) subset of M is relatively compact
  iv) any two points can be joined by a geodesic (of minimal length)
  v) any geodesic is infinitely extendable

As a consequence:
- a compact riemannian manifold is complete
- the affine parameter of a geodesic on a riemannian manifold is the arc length ℓ.

19.3 Submanifolds

On a pseudo-riemannian manifold (N,g) g induces a bilinear symmetric form in the tangent space of any submanifold M, but this form can be degenerate if it is not definite positive. Similarly the Lévy-Civita connection does not always induces an affine connection on a submanifold, even if N is riemmanian.

19.3.1 Induced scalar product

Theorem 1575 *If (N,g) is a real, finite dimensional pseudo-riemannian manifold, any submanifold M, embedded into N by f, is endowed with a bilinear symmetric form which is non degenerate at p$\in f(M)$ iff $\det [f'(p)]^t [g(p)][f'(p)] \neq 0$. It is non degenerate on f(M) if N is riemannian.*

Proof. Let us denote f(M)=$\widehat{M}$ as a subset of N.

Because $\widehat{M}$ is a submanifold any vector $v_q \in T_q N, q \in \widehat{M}$ has a unique decomposition : $v_q = v'_q + w_q$ with $v'_q \in T_q \widehat{W}$

Because f is an embedding, thus a diffeomorphism, for any vector $v'_q \in T_q \widehat{W}$ there is $u_p \in T_p M, q = f(p) : v'_q = f'(p) u_p$

So $v_q = f'(p) u_p + w_q$

g reads for the vectors of $T_q \widehat{W}$: $g(f(p))(f'(p) u_p, f'(p) u'_p) = f_* g(p)(u_p, u'_p)$

$h = f_* g$ is a bilinear symmetric form on M. And $T_q \widehat{W}$ is endowed with the bilinear symmetric form which has the same matrix : $[h(p)] = [f'(p)]^t [g(p)][f'(p)]$ in an adapted chart.

If g is positive definite : $\forall u'_p : g\left(f\left(p\right)\right)\left(f'\left(p\right)u_p, f'\left(p\right)u'_p\right) = 0 \Rightarrow f'\left(p\right)u_p = 0 \Leftrightarrow u_p \in \ker f'(p) \Rightarrow u_p = 0$ ∎

If N is n dimensional, M m<n dimensional, there is a chart in which $[f'(p)]_{n\times m}$ and $h_{\lambda\mu} = \sum_{\alpha\beta=1}^{n} g_{\alpha\beta}\left(f\left(p\right)\right)[f'\left(p\right)]_{\lambda}^{\alpha}[f'\left(p\right)]_{\mu}^{\beta}$

If g is riemannian there is an orthogonal complement to each vector space tangent at M.

19.3.2   Covariant derivative

1. With the notations of the previous subsection, h is a symmetric bilinear form on M, so, whenever it is not degenerate (meaning $[h]$ invertible) it defines a Levi-Civita connection $\widehat{\nabla}$ on M:

$\widehat{\Gamma}_{\mu\nu}^{\lambda} = \dfrac{1}{2}\sum_{\rho}(h^{\lambda\rho}\left(\partial_{\mu}h_{\nu\rho} + \partial_{\nu}h_{\mu\rho} - \partial_{\rho}h_{\mu\nu}\right)$

A lenghty but straightforward computation gives the formula :

$\widehat{\Gamma}_{\mu\nu}^{\lambda} = \sum_{\alpha\beta\gamma} G_{\alpha}^{\lambda}\left(\partial_{\nu}F_{\mu}^{\alpha} + \Gamma_{\beta\gamma}^{\alpha}F_{\mu}^{\beta}F_{\nu}^{\gamma}\right)$

with $[F] = [f'\left(p\right)]$ and $[G\left(p\right)]_{m\times n} = [h]_{m\times m}^{-1}[F]_{m\times n}^{t}[g]_{n\times n} \Rightarrow [G][F] = [h]^{-1}[F]^{t}[g][F] = I_{m}$

This connection $\widehat{\nabla}$ is symmetric and metric, with respect to h.

2. A vector field U∈ $\mathfrak{X}\left(TM\right)$ gives by push-forward a vector field on $T\widehat{M}$ and every vector field on $T\widehat{M}$ is of this kind.

The covariant derivative of such a vector field on N gives :

$\nabla_{f^{*}V}\left(f^{*}U\right) = X + Y \in \mathfrak{X}\left(TN\right)$ with $X = f^{*}U'$ for some U'∈ $\mathfrak{X}\left(TM\right)$

On the other hand the push forward of the covariant derivative on M gives : $f^{*}\left(\widehat{\nabla}_{V}U\right) \in \mathfrak{X}\left(T\widehat{M}\right)$

It can be schown (Lovelock p.269) that $\nabla_{f^{*}V}\left(f^{*}U\right) = f^{*}\left(\widehat{\nabla}_{V}U\right) + S\left(f^{*}U, f^{*}V\right)$

S is a bilinear map called the **second fundamental form.** If g is riemannian $S : \mathfrak{X}\left(T\widehat{M}\right) \times \mathfrak{X}\left(T\widehat{M}\right) \to \mathfrak{X}\left(T\widehat{M}^{\perp}\right)$ sends $u_q, v_q \in T_q\widehat{M}$ to the orthogonal complement $T_q\widehat{M}^{\perp}$ of $T_q\widehat{M}$ (Kobayashi II p.11).

So usually the induced connection is not a connection on $T\widehat{M}$, as $\nabla_{f^{*}V}\left(f^{*}U\right)$ has a component out of $T\widehat{M}$.

19.3.3   Vectors normal to a hypersurface

Theorem 1576 *If (N,g) is a real, finite dimensional pseudo-riemannian manifold, M a hypersurface embedded into N by f, then the symmetric bilinear form h induced in M is not degenerate at q∈f(M) iff there is a normal ν at f(M) such that $g(q)(\nu, \nu) \neq 0$*

Proof. 1. If f(M) is a hypersurface then $[F]$ is a nx(n-1) matrix of rank n-1, the system of n-1 linear equations : $[\mu]_{1\times n}[F]_{n\times n-1} = 0$ has a non null solution, unique up to a multiplication by a scalar. If we take : $[\nu] = [g]^{-1}[\mu]^{t}$ we have the components of a vector orthogonal to $T_q\widehat{M} : [F]^{t}[g][\nu] = 0 \Rightarrow \forall u_p \in T_pM :$

$g\left(f\left(p\right)\right)\left(f'(p)u_p, \nu_q\right) = 0$. Thus we have a non null normal, unique up to a scalar. Consider the matrix $\widehat{F} = [F, F_\lambda]_{n\times n}$ were the last column is any column of F. It is of rank n-1 and by development along the last column we gets :

$\det \widehat{F} = 0 = \sum_\alpha (-1)^{\alpha+n} \, \widehat{F}_\lambda^\alpha \det \left[\widehat{F}\right]_{(1...n\backslash n)}^{(1...n^\backslash\alpha)} = \sum_\alpha (-1)^{\alpha+n} \, F_\lambda^\alpha \det [F]^{(1...n^\backslash\alpha)}$

And the component expression of vectors normal to f(M) is :

$\nu = \sum_{\alpha\beta} (-1)^\alpha \, g^{\alpha\beta} \det [F]^{(1...n^\backslash\beta)} \partial y_\alpha$

2. If h is not degenerate at p, then there is an orthonormal basis $(e_i)_{i=1}^{n-1}$ at p in M, and $[h] = I_{n-1} = [F]^t [g] [F]$

If $\nu \in T_q\widehat{M}$ we would have $[\nu] = [F][u]$ for some vector $u_p \in T_pM$ and $[F]^t [g] [\nu] = [F]^t [g] [F] [u] = [u] = 0$

So $\nu \notin T_q\widehat{M}$. If $g(q)(\nu, \nu) = 0$ then, because $((f'(p)e_i, i = 1...n - 1), \nu)$ are linearly independant, they constitute a basis of T_qN. And we would have $\forall u \in T_qN : \exists v_q \in T_q\widehat{M}, y \in \mathbb{R} : u = v_p + y\nu$ and $g(q)(\nu, v_p + y\nu) = 0$ so g would be degenerate.

3. Conversely let $g(q)(\nu, \nu) \neq 0$ .If $\nu \in T_q\widehat{M} : \exists u \in T_pM : \nu = f'(p)u$. As ν is orthogonal to all vectors of $T_q\widehat{M}$ we would have : $g(q)(f'(p)u, f'(p)u) = 0$ .So $\nu \notin T_q\widehat{M}$ and the nxn matrix : $\widehat{F} = [F, \nu]_{n\times n}$ is the matrix of coordinates of n independant vectors.

$\left[\widehat{F}\right]^t [g] \left[\widehat{F}\right] = \begin{bmatrix} [F]^t [g] [F] & [F]^t [g] [\nu] \\ [\nu]^t [g] [F] & [\nu]^t [g] [\nu] \end{bmatrix} = \begin{bmatrix} [F]^t [g] [F] & 0 \\ 0 & [\nu]^t [g] [\nu] \end{bmatrix}$

$\det \left[\widehat{F}\right]^t [g] \left[\widehat{F}\right] = \det \left([F]^t [g] [F]\right) \det \left([\nu]^t [g] [\nu]\right) = g(p)(v, \nu) (\det [h]) = \left(\det \left[\widehat{F}\right]\right)^2 \det [g] \neq 0$

$\Rightarrow \det [h] \neq 0$ ■

Theorem 1577 *If M is a connected manifold with boundary ∂M in a real, finite dimensional pseudo-riemannian manifold (N,g), given by a function $f \in C_1(M; \mathbb{R})$, the symmetric bilinear form h induced in M by g is not degenerate at q$\in \partial M$ iff $g(gradf, gradf) \neq 0$ at q, and then the unitary, outward oriented normal vector to ∂M is : $\nu = \frac{gradf}{|g(gradf,gradf)|}$*

Proof. On the tangent space at p to $\partial M : \forall u_p \in T_p\partial M : f'(p)u_p = 0$

$\forall u \in T_pN : g(gradf, u) = f'(p)u$ so $f'(p)u_p = 0 \Leftrightarrow g(gradf, u_p) = 0$

So the normal N is proportional to gradf and the metric is non degenerate iff $g(gradf, gradf) \neq 0$

If M and ∂M are connected then for any transversal outward oriented vector v : $f'(p)v > 0$ so a normal outward oriented N is such that : $f'(p)N = g(gradf, N) > 0$ with N=kgradf : $g(gradf, gradf) k > 0$ ■

A common notation in this case is with the normal $\nu = \frac{gradf}{|g(gradf,gradf)|}$:

$\forall \varphi \in C_1(M; \mathbb{R}) : \frac{\partial\varphi}{\partial\nu} = g(grad\varphi, \nu) = \sum_{\alpha\beta} g_{\alpha\beta}g^{\alpha\gamma} (\partial_\gamma\varphi) \nu^\beta = \sum_\alpha (\partial_\alpha\varphi) \nu^\alpha = \varphi'(p)\nu$

If N is compact and we consider the family of manifolds with boundary : $M_t = \{p \in N : f(p) \le t\}$ then the flow of the vector ν is a diffeomorphism between the boundaries : $\partial M_t = \Phi_\nu (t - a, \partial M_a)$ (see Manifold with boundary).

19.3.4 Volume form

Theorem 1578 *If (N,g) is a real, finite dimensional pseudo-riemannian manifold with its volume form ϖ_0, M a hypersurface embedded into N by f, and the symmetric bilinear form h induced in f(M) is not degenerate at $q\in f(M)$, then the volume form ϖ_1 induced by h on f(M) is such that $\varpi_1 = i_\nu \varpi_0$, where ν is the outgoing unitary normal to f(M). Conversely $\varpi_0 = \nu^* \Lambda \varpi_1$ where ν^* is the 1-form $\nu^*_\alpha = \sum_\beta g_{\alpha\beta} \nu^\beta$*

Proof. if the metric h is not degenerate at q, there is an orthonormal basis $(\varepsilon_i)_{i=1}^{n-1}$ in $T_q f(M)$ and a normal unitary vector ν, so we can choose the orientation of ν such that $(\nu, \varepsilon_1, ..., \varepsilon_{n-1})$ is direct in TN, and :

$\varpi_0 (\nu, \varepsilon_1, ..., \varepsilon_{n-1}) = 1 = i_\nu \varpi_0$

Let us denote $\nu^* \in T_q N^*$ the 1-form such that $\nu^*_\alpha = \sum_\beta g_{\alpha\beta} \nu^\beta$

$\nu^* \wedge \varpi_1 \in \Lambda_n TN^*$ and all n-forms are proportional so : $\nu^* \wedge \varpi_1 = k\varpi_0$ as ϖ_0 is never null.

$(\nu^* \wedge \varpi_1)(n, \varepsilon_1, ..., \varepsilon_{n-1}) = k\varpi_0 (n, \varepsilon_1, ..., \varepsilon_{n-1}) = k$

$= \frac{1}{(n-1)!} \sum_{\sigma \in \mathfrak{S}_{n-1}} \epsilon(\sigma) \nu^*(n) \varpi_1 (\varepsilon_{\sigma(1)}, ...\varepsilon_{\sigma(q)}) = 1$

So $\nu^* \wedge \varpi_1 = \varpi_0$ ∎

Notice that this result needs only that the induced metric be not degenerate on M (we do not need the Levy Civita connection)

The volume form on N is : $\varpi_0 = \sqrt{|\det g|} dy^1 \wedge ... dy^n$

The volume form on $\widehat{M} = f(M)$ is $\varpi_1 = \sqrt{|\det h|} du^1 \wedge ... \wedge du^{n-1}$

19.3.5 Stockes theorem

Theorem 1579 *If M is a connected manifold with boundary ∂M in a real, pseudo-riemannian manifold (N,g), given by a function $f \in C_1(M; \mathbb{R})$, if $g(gradf, gradf) \ne 0$ on ∂M then for any vector field on N : $\int_M (divV) \varpi_0 = \int_{\partial M} g(V, \nu) \varpi_1$ where ϖ_1 the volume form induced by g on ∂M and ν the unitary, outward oriented normal vector to ∂M*

Proof. The boundary is an hypersurface embedded in N and given by f(p)=0.

In the conditions where the Stockes theorem holds, for a vector field V on M and $\varpi_0 \in \mathfrak{X}(\Lambda_n TN^*)$ the volume form induced by g in TN : $\int_M (divV) \varpi_0 = \int_{\partial M} i_V \varpi_0$

$i_V \varpi_0 = i_V (\nu^* \wedge \varpi_1) = (i_V \nu^*) \wedge \varpi_1 + (-1)^{\deg \nu^*} \nu^* \wedge (i_V \varpi_1) = g(V, \nu) \varpi_1 - \nu^* \wedge (i_V \varpi_1)$

On $\partial M : \nu^* \wedge (i_V \varpi_1) = 0$

$\int_M (divV) \varpi_0 = \int_{\partial M} g(V, \nu) \varpi_1$ ∎

The unitary, outward oriented normal vector to ∂M is : $\nu = \frac{gradf}{|g(gradf, gradf)|}$

$g\left(V,\nu\right)=\sum_{\alpha\beta}g_{\alpha\beta}V^{\alpha}\nu^{\beta}=\frac{1}{\|gradf\|}\sum_{\alpha\beta}g_{\alpha\beta}V^{\alpha}\sum_{\beta}g^{\beta\gamma}\partial_{\gamma}f=\frac{1}{\|gradf\|}\left(\sum_{\alpha}V^{\alpha}\partial_{\alpha}f\right)=\frac{1}{\|gradf\|}f'(p)V$

$\int_{M}\left(divV\right)\varpi_0=\int_{\partial M}\frac{1}{\|gradf\|}f'(p)V\varpi_1$

If V is a transversal outgoing vector field : $f'(p)V>0$ and $\int_{M}\left(divV\right)\varpi_0>0$

Notice that this result needs only that the induced metric be not degenerate on the boundary. If N is a riemannian manifold then the condition is always met.

Let $\varphi\in C_1\left(M;\mathbb{R}\right),V\in\mathfrak{X}\left(M\right)$ then for any volume form (see Lie derivative) :

$div\left(\varphi V\right)=\varphi'(V)+\varphi div\left(V\right)$

but φ'(V) reads : $\varphi'(V)=g\left(grad\varphi,V\right)$

With a manifold with boundary it gives the usual formula :

$\int_{M}div\left(\varphi V\right)\varpi_0=\int_{\partial M}g\left(\varphi V,n\right)\varpi_1=\int_{\partial M}\varphi g\left(V,n\right)\varpi_1=\int_{M}g\left(grad\varphi,V\right)\varpi_0+\int_{M}\varphi div\left(V\right)\varpi_0$

If φ or V has a support in the interior of M then $\int_{\partial M}\varphi g\left(V,n\right)\varpi_1=0$ and $\int_{M}g\left(grad\varphi,V\right)\varpi_0=-\int_{M}\varphi div\left(V\right)\varpi_0$

20 SYMPLECTIC MANIFOLDS

Symplectic manifolds,used in mechanics, are a powerful tool to study lagrangian models. Some of the concepts (such as Poisson fields) are also a key ingredient in quantum theory of fields. We consider only finite dimensional manifolds. The extension of symplectic structures to infinite dimensional manifolds does not seem promising. One of the key point of symplectic manifold (as of pseudo-riemanian manifolds) is the isomorphism with the dual, which requires finite dimensional manifolds. In fact Hilbert structures (quite normal for infinite dimensional manifolds) is the best extension.

We will follow mainly Hofer.

20.0.6 Symplectic manifold

Definition 1580 *A **symplectic manifold** M is a finite dimensional real manifold endowed with a 2-form ϖ closed non degenerate, called the **symplectic form**:*

$\varpi \in \mathfrak{X}(\Lambda_2 TM^*)$ is such that :
it is non degenerate : $\forall u_p \in TM : \forall v_p \in T_p M : \varpi(p)(u_p, v_p) = 0 \Rightarrow u_p = 0$
it is closed :$d\varpi = 0$

As each tangent vector space is a symplectic vector space a *necessary condition is that the dimension of M be even*, say 2m.

M must be at least of class 2.

Any open subset M of a symplectic manifold (N,ϖ) is a symplectic manifold (M,$\varpi|_M$).

Definition 1581 *The **Liouville form** on a 2m dimensional symplectic manifold (M, ϖ) is $\Omega = (\wedge \varpi)^m$. This is a volume form on M.*

So if M is the union of countably many compact sets it is orientable

Theorem 1582 *The product $M = M_1 \times M_2$ of symplectic manifolds $(M_1, \varpi_1), (M_2, \varpi_2)$ is a symplectic manifold:*

$V_1 \in VM_1, V_2 \in VM_2, V = (V_1, V_2) \in VM_1 \times VM_2$
$\varpi(V, W) = \varpi_1(V_1, W_1) + \varpi_2(V_2, W_2)$

Symplectic maps

Definition 1583 *A map $f \in C_1(M_1; M_2)$ between two symplectic manifolds $(M_1, \varpi_1), (M_2, \varpi_2)$ is **symplectic** if it preserves the symplectic forms : $f^* \varpi_2 = \varpi_1$*

f'(p) is a symplectic linear map, thus it is injective and we must have :
$\dim M_1 \leq \dim M_2$

f preserves the volume form : $f^*\Omega_2 = \Omega_1$ so we must have :
$\int_{M_1} f^*\Omega_2 = \int_{M_2} \Omega_2 = \int_{M_1} \Omega_1$

Total volumes measures play a special role in symplectic manifolds, with a special kind of invariant called capacity (see Hofer).

Symplectic maps are the morphisms of the category of symplectic manifolds.

Theorem 1584 *There cannot be a symplectic maps between compact smooth symplectic manifolds of different dimension.*

Proof. If the manifolds are compact, smooth and oriented :
$\exists k(f) \in \mathbb{Z} : \int_{M_1} f^*\Omega_2 = k(f) \int_{M_2} \Omega_2 = \int_{M_1} \Omega_1$

where k(f) is the degree of the map. If f is not surjective then k(f)=0. Thus if $\dim M_1 < \dim M_2$ then $\int_{M_1} \Omega_1 = 0$ which is not possible for a volume form.

if $\dim M_1 = \dim M_2$ then f is a local diffeomorphism ∎

Conversely: from the Moser's theorem, if M is a compact, real, oriented, finite dimensional manifold with volume forms ϖ, π such that : $\int_M \varpi = \int_M \pi$ then there is a diffeomorphism $f : M \to M$ such that $\pi = f^*\varpi$. So f is a symplectomorphism.

Canonical symplectic structure

The symplectic structure on $\mathbb{R}^{2m}$ is given by $\varpi_0 = \sum_{k=1}^m e^k \wedge f^k$ with matrix $J_m = \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix}_{2m \times 2m}$ and any basis $(e_k, f_k)_{k=1}^m$ and its dual $(e^k, f^k)_{k=1}^m$ (see Algebra)

Theorem 1585 *Darboux's theorem (Hofer p.10) For any symplectic manifold (M, ϖ) there is an atlas $(\mathbb{R}^{2m}, (O_i, \varphi_i)_{i \in I})$ such that, if $\mathbb{R}^{2m}$ is endowed with its canonical symplectic structure with the symplectic form ϖ_0, the transitions maps $\varphi_j^{-1} \circ \varphi_i$ are symplectomorphisms on $\mathbb{R}^{2m}$*

They keep invariant $\varpi_0 : (\varphi_j^{-1} \circ \varphi_i)^* \varpi_0 = \varpi_0$

The maps φ_i are symplectic diffeomorphisms.

Then there is a family $(\varpi_i)_{i \in I} \in \mathfrak{X}(\Lambda_2 TM^*)^I$ such that $\forall p \in O_i : \varpi(p) = \varpi_i(p)$ and $\varpi_i = \varphi_i^* \varpi_0$

We will denote $(x_\alpha, y_\alpha)_{\alpha=1}^m$ the coordinates in $\mathbb{R}^{2m}$ associated to the maps φ_i :and the canonical symplectic basis. Thus there is a holonomic basis of M which is also a canonical symplectic basis : $dx^\alpha = \varphi_i^*(e^\alpha), dy^\alpha = \varphi_i^*(f^\alpha)$ and :
$\varpi = \sum_{\alpha=1}^m dx^\alpha \wedge dy^\alpha$

The Liouville forms reads : $\Omega = (\wedge \varpi)^m = dx^1 \wedge ... \wedge dx^m \wedge dy^1 \wedge .. \wedge dy^m$.

So locally all symplectic manifolds of the same dimension look alike.

Not all manifolds can be endowed with a symplectic structure (the spheres S_n for n>1 have no symplectic structure).

Generating functions

(Hofer p.273)

In $\left(\mathbb{R}^{2m}, \varpi_0\right)$ symplectic maps can be represented in terms of a single function, called generating function.

Let $f : \mathbb{R}^{2m} \to \mathbb{R}^{2m} :: f(\xi, \eta) = (x, y)$ be a symplectic map with $x = \left(x^i\right)_{i=1}^m$,...

$x = X(\xi, \eta)$

$y = Y(\xi, \eta)$

If $\det\left[\frac{\partial X}{\partial \xi}\right] \neq 0$ we can change the variables and express f as :

$\xi = A(x, \eta)$

$y = B(x, \eta)$

Then f is symplectic if there is a function $W : \mathbb{R}^{2m} \to \mathbb{R}^{2m} :: W(x, \eta)$ such that :

$\xi = A(x, \eta) = \frac{\partial W}{\partial \eta}$

$y = B(x, \eta) = \frac{\partial W}{\partial x}$

20.0.7 Hamiltonian vector fields

Isomorphism between the tangent and the cotangent bundle

Theorem 1586 *On a symplectic manifold (M, ϖ), there is a canonical isomorphism at any point p between the tangent space T_pM and the cotangent space T_pM^*, and between the tangent bundle TM and the cotangent bundle TM^*.*

$j : TM \to TM^* :: u_p \in T_pM \to \mu_p = \jmath(u_p) \in T_pM^* :: \forall v_p \in T_pM :$
$\varpi(p)(u_p, v_p) = \mu_p(v_p)$

Hamilonian vector fields

As a particular case, if f is a function then its differential is a 1-form.

Definition 1587 *The **Hamiltonian vector field** V_f associated to a function $f \in C_1(M; \mathbb{R})$ on a symplectic manifold (M, ϖ) is the unique vector field such that : $i_{V_f}\varpi = -df \Leftrightarrow \forall W \in \mathfrak{X}(TM) : \varpi(V_f, W) = -df(W)$*

Theorem 1588 *The flow of a Hamiltonian vector field preserves the symplectic form and is a one parameter group of symplectic diffeomorphisms. Conversely Hamiltonian vector fields are the infinitesimal generators of one parameter group of symplectic diffeomorphisms.*

Proof. We have $\pounds_{V_f}\varpi = i_{V_f}d\varpi + d\circ i_{V_f}\varpi = 0$ so the flow of a Hamiltonian vector field preserves the symplectic form.

$\forall t : \Phi_{V_f}(t, .)^* \varpi = \varpi$: the flow is a one parameter group of symplectic diffeomorphisms. ∎

So symplectic structures show a very nice occurence : the infinitesimal generators of one parameter groups of diffeomorphisms which preserve the structure (the form ϖ) are directly related to functions f on M.

Theorem 1589 *The divergence of Hamiltonian vector fields is null.*

Proof. Hamiltonian vector field preserves the Liouville form : $\pounds_{Vf}\Omega = 0$. So we have $\pounds_{Vf}\Omega = (divV_f)\,\Omega \Rightarrow divV_f = 0$ ∎

Theorem 1590 *Symplectic maps S maps hamiltonian vector fields to hamiltonian vector fields*

$$V_{f\circ S} = S^*V_f : \Phi_{V_f}(t,.) \circ S = S \circ \Phi_{V_{f\circ S}}(t,.)$$

Poisson brackets

Definition 1591 *The **Poisson bracket** of two functions f,h $\in C_1(M;\mathbb{R})$ on a symplectic manifold (M,ϖ) is the function :*
$$(f,h) = \varpi(grad(f),grad(h)) = \varpi(JV_f,JV_h) = \varpi(V_f,V_h) \in C(M;\mathbb{R})$$

Theorem 1592 *With the Poisson bracket the vector space $C_\infty(M;\mathbb{R})$ is a Lie algebra (infinite dimnsional)*

i) The Poisson bracket is a an antisymmetric, bilinear map.:
$\forall f_1, f_2 \in C_1(M;\mathbb{R}), k, k' \in \mathbb{R}$:
$(f_1, f_2) = -(f_2, f_1)$
$(kf_1 + k'f_2, h) = k(f_1, h) + k'(f_2, h)$
$(f_1 f_2, h) = (f_1, h)f_2 + (f_2, h)f_1$
ii) $(f_1, (f_2, f_3)) + (f_2, (f_3, f_1) + (f_3, (f_1, f_2)) = 0$
Furthermore the Poisson bracket has the properties:
i) for any function : $\phi \in C_1(\mathbb{R};\mathbb{R}) : (f, \phi \circ h) = \phi'(h)(f,h)$
ii) If f is constant then : $\forall h \in\in C_1(M;\mathbb{R}) : (f,h) = 0$
iii) $(f,h) = \pounds_{grad(h)}(f) = -\pounds_{grad(f)}(h)$

20.0.8   Complex structure

Theorem 1593 *(Hofer p.14) On a symplectic manifold (M,ϖ) there is an almost complex structure J and a Riemannian metric g such that : $\forall u, p \in T_pM$: $\varpi(p)(u, Jv) = g(p)(u,v)$*

$\varpi(p)(Ju, Jv) = \varpi(p)(u,v)$ so J(p) is a symplectic map in the tangent space
$J^2 = -Id, J^* = J^{-1}$ where J* is the adjoint of J by g : $g(Ju, v) = g(u, J^*v)$
Finite dimensional real manifolds generally admit Riemannian metrics (see Manifolds), but g is usually not this metric. The solutions J,g are not unique.
So a symplectic manifold has an almost complex Kähler manifold structure.
In the holonomic symplectic chart :
$[J^*] = [g]^{-1}[J][g]$
$\varpi(p)(u, Jv) = [u]^t[J]^t[J_m][v] = [u]^t[g][v] \Leftrightarrow [g] = [J]^t[J_m]$
J is not necessarily J_m
$\det g = \det J \det J_m = 1$ because $[J]^2 = [J_m]^2 = -I_{2m}$

So the volume form for the Riemannian metric is identical to the Liouville form Ω.

If V_f is a Hamiltonian vector field : $\varpi\left(V_f, W\right) = -df\left(W\right) = \varpi\left(W, -J^2 V_f\right) = g\left(W, -JV_f\right)$ so : $JV_f = grad\left(f\right)$.

20.0.9 Symplectic structure on the cotangent bundle

Definition

Theorem 1594 *The cotangent bundle TM* of a m dimensional real manifold can be endowed with the structure of a symplectic manifold*

The symplectic form is $\varpi = \sum_\alpha dy^\alpha \wedge dx^\alpha$ where $\left(x^\alpha, y^\beta\right)_{\alpha,\beta=1}^m$ are the coordinates in TM*

Proof. Let $\left(\mathbb{R}^m, \left(O_i, \varphi_i\right)_{i\in I}\right)$ be an atlas of M, with coordinates $(x^\alpha)_{\alpha=1}^m$

The cotangent bundle TM* has the structure of a 2m dimensional real manifold with atlas $\left(\mathbb{R}^m \times \mathbb{R}^{m*}, \left(O_i \times \cup_{p\in O_i} T_p M^*, \left(\varphi_i, \left(\varphi_i'\right)^t\right)\right)\right)$

A point $\mu_p \in TM^*$ can be coordinated by the pair $\left(x^\alpha, y^\beta\right)_{\alpha,\beta=1}^m$ where x stands for p and y for the components of μ_p in the holonomic basis. A vector $Y \in T_{\mu_p} TM^*$ has components $(u^\alpha, \sigma_\alpha)_{\alpha\in A}$ expressed in the holonomic basis $(\partial x_\alpha, \partial y_\alpha)$

Let be the projections :

$\pi_1 : TM^* \to M :: \pi_1\left(\mu_p\right) = p$

$\pi_2 : T\left(TM^*\right) \to TM^* :: \pi_2\left(Y\right) = \mu_p$

So : $\pi_1'\left(\mu_p\right) : T_{\mu_p} T_p M^* \to T_p M :: \pi_1'\left(\mu_p\right) Y = u \in T_p M$

Define the 1-form over $TM^* : \lambda\left(\mu_p\right) \in L\left(T_{\mu_p} T_p M^*; \mathbb{R}\right)$

$\lambda\left(\mu_p\right)\left(Y\right) = \pi_2\left(Y\right)\left(\pi'\left(\mu_p\right)\left(Y\right)\right) = \mu_p\left(u\right) \in \mathbb{R}$

It is a well defined form. Its components in the holonomic basis associated to the coordinates $(x^\alpha, y_\alpha)_{\alpha\in A}$ are :

$\lambda\left(\mu_p\right) = \sum_\alpha y_\alpha dy^\alpha$

The components in dy^α are zero.

The exterior differential of λ is $\varpi = d\lambda = \sum_\alpha dy^\alpha \wedge dx^\alpha$ so ϖ is closed, and it is not degenerate. $\blacksquare$

If $X, Y \in T_{\mu_p} TM^* : X = \sum_\alpha \left(u^\alpha \partial x_\alpha + \sigma^\alpha \partial y^\alpha\right); Y = \sum_\alpha \left(v^\alpha \partial x_\alpha + \theta^\alpha \partial y^\alpha\right)$

$\varpi\left(\mu_p\right)\left(X, Y\right) = \sum_\alpha \left(\sigma^\alpha u^\alpha - \theta^\alpha v^\alpha\right)$

Application to analytical mechanics

In analytic mechanics the state of the system is described as a point q in some m dimensional manifold M coordinated by m positional variables $\left(q^i\right)_{i=1}^m$ which are coordinates in M (to account for the constraints of the system modelled as liaisons between variables) called the configuration space. Its evolution is some path $\mathbb{R} \to M : q(t)$ The quantities $\left(q^i, \frac{dq^i}{dt} = \dot{q}^i\right)$ belong to the tangent bundle TM.

The dynamic of the system is given by the principle of least action with a Lagrangian : $L \in C_2\left(M \times \mathbb{R}^m; \mathbb{R}\right) : L\left(q_i, u_i\right)$

q(t) is such that : $\int_{t_0}^{t_1} L(q\left(t\right), \dot{q}\left(t\right))dt$ is extremal.

The Euler-Lagrange equations give for the solution :

$\frac{\partial L}{\partial q^i} = \frac{d}{dt}\frac{\partial L}{\partial u^i} = \frac{\partial}{\partial u^i}\frac{d}{dt}L$

If the Hessian $\left[\frac{\partial^2 L}{\partial u_i \partial u_j}\right]$ has a determinant which is non zero it is possible to implement the change of variables :

$\left(q^i, u^i\right) \to \left(q^i, p_i\right) : p_i = \frac{\partial L}{\partial u^i}$

and the equations become : $q^{i\prime} = \frac{\partial H}{\partial p_i}; p_i' = -\frac{\partial H}{\partial q^i}$ with the Hamiltonian :

$H(q,p) = \sum_{i=1}^n p_i u^i - L(q,u)$

The new variables $p_i \in TM^*$ are called the moments of the system and the cotangent bundle TM* is the phase space. The evolution of the system is a path :$C : \mathbb{R} \to TM^* :: \left(q^i\left(t\right), p_i\left(t\right)\right)$ and $C'(t) = \left(\dot{q}^i, \dot{p}_i\right)$

The Hamiltonian vector field V_H associated with H has the components :

$V_H = (\frac{\partial H}{\partial p_i}\partial q_i, -\frac{\partial H}{\partial q^i}\partial p_i)_{i=1..m}$

So the solutions C(t) are just given by the flow of V_H.

20.0.10   Surfaces of constant energy

As infinitesimal generators of the one parameter group of diffeomorphisms (Hamiltonian vector fields) are related to functions, the submanifolds where this function is constant play a specific role. They are manifolds with boundary. So manifolds with boundary are in some way the integral submanifolds of symplectic structures.

In physics usually a function, the energy H, plays a special role, and one looks for the evolutions of a system such that this energy is constant.

Principles

If a system is modelled by a symplectic manifold (M,ϖ) (such as above) with a function $H \in C_1\left(M; \mathbb{R}\right)$:

1. The value of H is constant along the integral curves of its Hamiltonian vector field V_H of H

The Hamiltonian vector field and its flow Φ_{V_H} are such that :

$\varpi\left(V_H, \frac{\partial}{\partial t}\Phi_{V_H}\left(t,.\right)|_{t=\theta}\right) = -dH\frac{\partial}{\partial t}\Phi_{V_H}\left(t,.\right)|_{t=\theta} = 0 = \frac{d}{dt}H\left(\Phi_{V_H}\left(t,p\right)\right)|_{t=\theta}$

So : $\forall t, \forall p : H\left(\Phi_{V_H}\left(t,p\right)\right) = H\left(p\right) \Leftrightarrow \Phi_{V_H}\left(t,.\right)_* H = H$.

2. The divergence of V_H is null.

3. H defines a folliation of M with leaves the surfaces of constant energy $\partial S_c = \{p \in M : H\left(p\right) = c\}$

If H'(p)=0 then $V_H = 0$ because $\forall W : \varpi\left(V_H, W\right) = -dH\left(W\right) = 0$

If H'(p)$\neq$0 the sets : $S_c = \{p \in M : H\left(p\right) \leq c\}$ are a family of manifolds with boundary the hypersurfaces ∂S. The vector V_H belongs to the tangent space to ∂S_c The hypersurfaces ∂S_c are preserved by the flow of V_H.

4. If there is a Riemannian metric g (as above) on M then the unitary normal outward oriented ν to ∂S_c is : $\nu = \frac{JV_H}{\varpi\left(V_H, JV_H\right)}$

$\nu = \frac{gradH}{|g(gradH,gradH)|}$ with $gradH = JV_H \Rightarrow g(gradH, gradH) = g(JV_H, JV_H) = g(V_H, V_H) = \varpi(V_H, JV_H) > 0$

The volume form on ∂S_t is $\Omega_1 = i_\nu \Omega \Leftrightarrow \Omega = \Omega_1 \wedge \nu$

If M is compact then H(M)=[a,b] and the flow of the vector field ν is a diffeomorphism for the boundaries $\partial S_c = \Phi_\nu(\partial S_a, c)$. (see Manifolds with boundary).

Periodic solutions

If t is a time parameter and the energy is constant, then the system is described by some curve c(t) in M, staying on the boundary ∂S_c. There is a great deal of studies about the kind of curve that can be found, depending of the surfaces S. They must meet the equations :

$\forall u \in T_{c(t)}M : \varpi(V_H(c(t)), u) = -dH(c(t))u$

A T **periodic solution** is such that if : $c'(t) = V_H(t)$ on M, then c(T)=c(0). The system comes back to the initial state after T.

There are many results about the existence of periodic solutions and about ergodicity. The only general theorem is :

Theorem 1595 *Poincaré's recurrence theorem (Hofer p.20): If M is a Hausdorff, second countable, symplectic (M, ϖ) manifold and $H \in C_1(M; \mathbb{R})$ such that H'(p) is not null on M, then for almost every point p of $\partial S_c = \{p \in M : H(p) = c\}$ there is an increasing sequence $(t_n)_{n \in \mathbb{N}}$ such that $\lim_{n \to \infty} \Phi_{V_H}(t_n, p) = p$*

the null measure is with respect to the measure Ω_1. The proof in Hofer can easily be extended to the conditions above.

One says that p is a recurring point : if we follow an integral curve of the hamiltonian vector, then we will come back infinitely often close to any point.

Part V

PART 5 : LIE GROUPS

The general properties and definitions about groups are seen in the Algebra part. Groups with countably many elements have been classified. When the number of elements is uncountable the logical way is to endow the set with a topological structure : when the operations (product and inversion) are continuous we get the continuous groups. Further if we endow the set with a manifold structure compatible with the group structure we get Lie groups. The combination of the group and manifold structures gives some stricking properties. First the manifold is smooth, and even analytic. Second, the tangent space at the unity element of the group in many ways summarize the group itself, through a Lie algebra. Third, most (but not all) Lie groups are isomorphic to a linear group, meaning a group of matrices, that we get through a linear representation. So the study of Lie groups is closely related to the study of Lie algebras and linear representations of groups.

In this part we start with Lie algebras. As such a Lie algebra is a vector space endowed with an additional internal operation (the bracket). The most common example of Lie algebra is the set of vector fields over a manifold equipped with the commutator. In the finite dimensional case we have more general results, and indeed all finite dimensional Lie algebras are isomorphic to some algebra of matrices, and have been classified. Their study involves almost uniquely algebraic methods. Thus the study of finite dimensional Lie groups stems mostly from their Lie algebra.

The theory of linear representation of Lie groups is of paramount importance in physics. There is a lot of litterature on the subject, but unfortunately it is rather confusing. The point is that this is a fairly technical subject, with many traditional notations and conventions which are not very helpful. I will strive to put some ligth on the topic, with the main purpose to give to the reader the most useful and practical grasp on these questions.

21 LIE ALGEBRAS

I will follow mainly Knapp.

21.1 Definitions

21.1.1 Lie algebra

Definition 1596 *A **Lie algebra G over a field K** is a vector space A over K endowed with a bilinear map (**bracket**) $: [] : A \times A \to A$*

$\forall X, Y, Z \in A, \forall \lambda, \mu \in K : [\lambda X + \mu Y, Z] = \lambda [X, Z] + \mu [Y, Z]$ *such that :*

$[X, Y] = - [Y, X]$

$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ *(Jacobi identities)*

Notice that a Lie algebra is not an algebra because the bracket is not associative. But any algebra becomes a Lie algebra with the bracket : $[X, Y] = X \cdot Y - Y \cdot X$.

The dimension of the Lie algebra is the dimension of the vector space. In the following A can be infinite dimensional if not otherwise specified.

A Lia algebra is said to be **abelian** if it is commutative, then the bracket is null : $\forall X, Y : [X, Y] = 0$

Notation 1597 **ad** *is the linear map* $ad : A \to L(A; A) :: ad(X)(Y) = [X, Y]$ *induced by the bracket*

Classical examples

The set of linear endomorphisms over a vector space L(E;E) with the bracket ; $[f, g] = f \circ g - g \circ f$.

The set K(r) of square rxr matrices endowed with the bracket : $[X, Y] = [X][Y] - [Y][X]$.

The set of vector fields VM over a manifold endowed with the commutator : $[V, W]$

Any vector space with the trivial bracket : $[X, Y] = 0$. Indeed any commutative Lie algebra has this bracket.

Structure coefficients

If $(e_i)_{i \in I}$ is a basis of the vector space A, then : $\forall i, j : [e_i, e_j] = \sum_{k \in I} C_{ij}^k e_k$ where the family $\left(C_{ij}^k\right)_{k \in I}$ of scalars in K, called the **structure coefficients**, has at most finitely many non zero elements. In a vector space the components of a vector can take any value. This is not the case with the structure coefficients. We must have $C_{ij}^k = -C_{ji}^k$ and due to the Jacobi identities :

$\forall i, j, k, m : \sum_{l \in I} \left(C_{jk}^l C_{il}^m + C_{ki}^l C_{jl}^m + C_{ij}^l C_{kl}^m\right) = 0$

So the structure coefficients have special properties. Conversely a family of structure coefficients meeting these relations define, in any basis of a vector space, a bracket and so a unique Lie algebra structure (with the understanding that the structure coefficients change accordingly in a change of basis).

The system of equations for the structure coefficients has not always a solution : not all Lie algebras structures are possible in a vector space over a field K with a given dimension. This is the starting point for the classification of Lie algebras.

21.1.2 Morphisms of Lie algebras

Definition

Definition 1598 *A **Lie algebra morphism** (also called a homomorphism) is a linear map f between Lie algebras $(A, []_A)$,$(B, []_B)$ which preserves the bracket :*

$f \in L(A; B) : \forall X, Y \in A : f([X, Y]_A) = [f(X), f(Y)]_B$
They are the morphisms of the category of Lie algebras over the same field.
We will denote $\hom(A, B)$ the set of morphisms from A to B
A Lie algebra isomorphism is an **isomorphism** of vector spaces which is also a Lie algebra morphism.
Two Lie algebras A,B are **isomorphic** if there is an isomorphism of Lie algebra : $A \to B$

Spaces of Lie algebras morphisms

Theorem 1599 *The set L(A;A) of linear maps over a Lie algebra A is a Lie algebra with the composition law and the map $ad : A \to L(A; A)$ is a Lie algebra morphism :*

Proof. i) $\forall X, Y \in A : f([X, Y]) = [f(X), f(Y)], g([X, Y]) = [g(X), g(Y)]$
$h([X, Y]) = f \circ g([X, Y]) - g \circ f([X, Y]) = f([g(X), g(Y)]) - g([f(X), f(Y)]) = [f \circ g(X), f \circ g(Y)] - [g \circ f(X), g \circ f(Y)]$
So : $\forall f, g \in \hom(A; A) : h = [f, g] = f \circ g - g \circ f \in \hom(A; A)$
ii) Take $U \in A : ad(X) \circ ad(Y)(U) - ad(Y) \circ ad(X)(U) = [X, [Y, U]] - [Y, [X, U]]$
$= [X, [Y, U]] + [Y, [U, X]] = -[U, [X, Y]] = [[X, Y], U] = ad([X, Y])(U)$
So : $ad \in \hom(A, L(A; A)) : [ad(X), ad(Y)]_{L(A;A)} = ad(X) \circ ad(Y) - ad(Y) \circ ad(X) = ad([X, Y]_A)$ ∎

Definition 1600 *An **automorphism** over a Lie algebra $(A, [])$ is a linear automorphism of vector space (thus it must be inversible) which preserves the bracket*

$f \in GL(A) : \forall X, Y \in A : f([X, Y]) = [f(X), f(Y)]$ .Then f^{-1} is a Lie algebra automorphism.
The set of automorphisms over A is a group with the composition law.

Definition 1601 *A **derivation** over a Lie algebra $(A, [])$ is an endomorphism D such that :*
$D \in L(A; A) : \forall X, Y \in A : D([X, Y]) = [D(X), Y] + [X, D(Y)]$

For any X the map ad(X) is a derivation.

Theorem 1602 *(Knapp p.38) The set of derivations over a Lie algebra $(A, [])$, denoted Der(A), has the structure of a Lie algebra with the composition law :*
$D, D' \in Der(A) : D \circ D' - D' \circ D \in Der(A)$
The map : $ad : A \to Der(A)$ is a Lie algebra morphism

21.1.3 Killing form

Definition 1603 *The Killing form over a finite dimensional Lie algebra $(A, [])$ is the bilinear map B :*
$B \in L^2(A, A; K) : A \times A \to K :: B(X, Y) = Trace(ad(X) \circ ad(Y))$

(see tensors).

Theorem 1604 *The Killing form is a bilinear symmetric form*

Proof. In a basis $(e_i)_{i \in I}$ of A and its dual $\left(e^i\right)_{i \in I}$ the map ad can be read as the tensor :
$ad(X) = \sum_{i,j,k \in I} C_{kj}^i x^k e_i \otimes e^j$
$ad(Y) = \sum_{i,j,k \in I} C_{kj}^i y^k e_i \otimes e^j$
$ad(X) \circ ad(Y)(Z) = \sum_{i,j,k \in I} ad(X) \left(C_{kj}^i y^k z^j e_i\right) = \sum_{i,j,k \in I} C_{kj}^i y^k z^j ad(X)(e_i) = \sum_{i,j,k \in I} C_{kj}^i y^k z^j \sum_{l,m \in I} C_{mi}^l x^m e_l$
$ad(X) \circ ad(Y) = \sum_{i,j,k,l,m \in I} C_{kj}^i C_{mi}^l y^k x^m e^j \otimes e_l$
$Trace(ad(X) \circ ad(Y)) = \sum_{i,j,k,m \in I} C_{kj}^i C_{mi}^j y^k x^m$
So $Trace(ad(X) \circ ad(Y)) = Trace(ad(Y) \circ ad(X))$
And $B = \sum_{i,j,k,m \in I} C_{kj}^i C_{mi}^j e^k \otimes e^m$ ∎

Theorem 1605 *The Killing form is such that : $\forall X, Y, Z \in A : B([X,Y], Z) = B(X, [Y,Z])$*

It comes from the Jacobi identities

Theorem 1606 *(Knapp p.100) Any automorphism of a Lie algebra $(A, [])$ preserves the Killing form*

$$f([X,Y]) = [f(X), f(Y)] \Rightarrow B(f(X), f(Y)) = B(X, Y)$$

21.1.4 Subsets of a Lie algebra

Subalgebra

Definition 1607 *A **Lie subalgebra** (say also subalgebra) of a Lie algebra $(A, [])$ is a vector subspace B of A which is closed under the bracket operation : $\forall X, Y \in B : [X, Y] \in B$.*

415

Notation 1608 $[B, C]$ *denotes the vector space* $Span \{[X, Y], X \in B, Y \in C\}$ *generated by all the brackets of elements of B,C in A*

Theorem 1609 *If B is a subalgebra and f an automorphism then f(B) is a subalgebra.*

Definition 1610 *The **normalizer** N(B) of a subalgeba B of the Lie algebra* $(A, [])$ *is the set of vectors :* $N(B) = \{X \in A : \forall Y \in B : [X, Y] \in B\}$

Ideal

Definition 1611 *An **ideal** of of the Lie algebra* $(A, [])$ *is a vector subspace B of A such that :* $[A, B] \sqsubseteq B$

So an ideal is a subalgebra (the converse is not true)

If B,C are ideals then the sets $B + C, [B, C], B \cap C$ are ideal

If A,B are Lie algebras and $f \in \hom(A, B)$ then ker f is an ideal of A.

If B is an ideal the quotient set A/B : $X \sim Y \Leftrightarrow X - Y \in B$ is a Lie algebra with the bracket : $[[X], [Y]] = [X, Y]$ because $\forall U, V \in B, \exists W \in B :$ $[X + U, Y + V] = [X, Y] + W$ .Then the map : $A \rightarrow A/B$ is a Lie algebra morphism

Center

Definition 1612 *The **centralizer** Z(B) of a subset B of the Lie algebra* $(A, [])$ *is the set :* $Z(B) = \{X \in A : \forall Y \in B : [X, Y] = 0\}$

*The **center** Z(A) of the Lie algebra* $(A, [])$ *is the centralizer of A itself*

So : $Z(A) = \{X \in A : \forall Y \in A : [X, Y] = 0\}$ is the set (which is an ideal) of vectors which commute with any vector of A.

Z(A) is an ideal.

21.1.5 Complex and real Lie algebra

(see Complex vector spaces in the Algebra part)

Complexified

There are two ways to define a complex vector space structure on a real vector space and so for a real Lie algbra.

1. Complexification:

Theorem 1613 *Any real Lie algebra* $(A, [])$ *can be endowed with the structure of a complex Lie algebra, called its **complexified** $(A_{\mathbb{C}}, []_{\mathbb{C}})$ which has same basis and structure coefficients as A.*

Proof. i) It is always possible to define the complexified vector space $A_{\mathbb{C}} = A \oplus iA$ over A

ii) define the bracket by :

$[X + iY, X' + iY']_{\mathbb{C}} = [X, X'] - [Y, Y'] + i([X, Y'] + [X', Y])$ ∎

Definition 1614 *A **real form of a complex Lie algebra** $(A, [])$ is a real Lie algebra $(A_0, [])$ such that its complexified is equal to $A : A \equiv A_0 \oplus iA_0$*

2. Complex structure:

Theorem 1615 *A complex structure J on a real Lie algebra $(A, [])$ defines a structure of complex Lie algebra on the set A iff $J \circ ad = ad \circ J$*

Proof. J is a linear map $J \in L(E; E)$ such that $J^2 = -Id_E$, then for any $X \in A : iX$ is defined as J(X).

The complex vector space structure A_J is defined by $X = x + iy \Leftrightarrow X = x + J(y)$ then the bracket

$[X, X']_{A_J} = [x, x'] + [J(y), J(y')] + [x, J(y')] + [x', J(y)] = [x, x'] - [Jy, y'] + i([x, y'] + [x', y])$

if $[x, J(y')] = -J([x, y']) \Leftrightarrow \forall x \in A_1 : J(ad(x)) \circ ad(J(x)) \Leftrightarrow J \circ ad = ad \circ J$ ∎

If A is finite dimensional a necessary condition is that its dimension is even.

Real structure

Theorem 1616 *Any real structure on a complex Lie algebra $(A, [])$ defines a structure of real Lie algebra with same bracket on the real kernel.*

Proof. There are always a real structure, an antilinear map σ such that $\sigma^2 = Id_A$ and any vector can be written as : $X = \operatorname{Re} X + i \operatorname{Im} X$ where $\operatorname{Re} X, \operatorname{Im} X \in A_{\mathbb{R}}$. The real kernel $A_{\mathbb{R}}$ of A is a real vector space, subset of A, defined by $\sigma(X) = X$. It is simple to check that the bracket is closed in $A_{\mathbb{R}}$. ∎

Notice that there are two real vector spaces and Lie algebras $A_{\mathbb{R}}, iA_{\mathbb{R}}$ which are isomorphic (in A) by multiplication with i. The **real form** of the Lie algebra is : $A_\sigma = A_{\mathbb{R}} \times iA_{\mathbb{R}}$ which can be seen either as the direct product of two real algebras, or a real algebras of two times the complex dimension of A.

If σ is a real structure of A then $A_{\mathbb{R}}$ is a real form of A.

21.2   Sum and product of Lie algebras

21.2.1   Free Lie algebra

Definition 1617 *A **free Lie algebra** over any set X is a pair $(L, \jmath)$ of a Lie algebra L and a map : $\jmath : X \to L$ with the universal property : whatever the Lie algebra A and the map : $f : X \to A$ there is a unique Lie algebra morphism $F : L \to A$ such that : $f = F \circ j$*

$$\begin{array}{ccccc}
 & & \text{\scriptsize J} & & \\
\text{X} & \rightarrow & \rightarrow & L & \\
 & \searrow & & \downarrow & \\
 & \text{f} & \searrow & \downarrow & \text{F} \\
 & & & \text{A} &
\end{array}$$

Theorem 1618 *(Knapp p.188) For any non empty set X there is a free algebra over X and the image ȷ(X) generates L. Two free algebras over X are isomorphic.*

21.2.2 Sum of Lie algebras

Definition 1619 *The sum of two Lie algebras $(A, []_A), (B, []_B)$ is the vector space $A \oplus B$ with the bracket :*

$\quad X, Y \in A, X', Y' \in B : [X + X', Y + Y']_{A \oplus B} = [X, Y]_A + [X', Y']_B$

then A'=(A,0), B'=(0,B) are ideals of $A \oplus B$

Definition 1620 *A Lie algebra A is said to be **reductive** if for any ideal B there is an ideal C such that $A = B \oplus C$*

A real Lie algebra of matrices over the fields $\mathbb{R}, \mathbb{C}, H$ which is closed under the operation conjugate / transpose is reductive.

21.2.3 Semi-direct product

Theorem 1621 *(Knapp p.38) If $(A, []_A), (B, []_B)$ are two Lie algebras over the same field, F a Lie algebra morphism $F : A \rightarrow Der(B)$ where B is the set of derivations over B, then there is a unique Lie algebra structure over $A \oplus B$ called **semi-direct product** of A,B, denoted $C = A \oplus_F B$ such that :*

$\quad \forall X, Y \in A : [X, Y]_C = [X, Y]_A$
$\quad \forall X, Y \in B : [X, Y]_C = [X, Y]_B$
$\quad \forall X \in A, Y \in B : [X, Y]_C = F(X)(Y)$

Then A is a subalgebra of C, and B is an ideal of C.
The direct sum is a special case of semi-direct product with F=0.

21.2.4 Universal envelopping algebra

A Lie algebra is not an algebra. It entails that the computations in a Lie algebras, when they involve many brackets, become quickly unmanageable. This is the case with the linear representations (F, ρ) of Lie algebras where it is natural to deal with products of the kind $\rho(X_1) \rho(X_2) ... \rho(X_p)$ which are product of matrices, image of tensorial products $X_1 \otimes ... \otimes X_p$. Moreover it is useful to be able to use some of the many theorems about "true" algebras. But to build an algebra over a Lie algebra A requires to use many copies of A.

Definition

Definition 1622 *The universal envelopping algebra of order r $U_r(A)$ of a Lie algebra $(A, [])$ over the field K is the quotient space: $(T(A))^r$ of the tensors of order r over A, by the two sided ideal :$J = \{X \otimes Y - Y \otimes X - [X, Y], X, Y \in T^1(A)\}$*

*The **universal envelopping algebra** $U(A)$ is the direct sum : $U(A) = \oplus_{r=0}^{\infty} U_r(A)$*

Theorem 1623 *(Knapp p.214) With the tensor product U(A) is a unital algebra over the field K*

The scalars K belong to U(A) and the unity element is 1.

So all elements of the kind : $X \otimes Y - Y \otimes X - [X, Y] \sim 0$

The subset $U_r(A)$ of the elements of U(A) which can be written as products of exactly r elements of A is a vector subspace of U(A).

U(A) *is not a Lie algebra*. Notice that A can be infinite dimensional.

The map : $\imath : A \to U(A)$ is one-one with the founding identity :

$\imath[X, Y] = \imath(X)\imath(Y) - \imath(Y)\imath(X)$

Theorem 1624 *(Knapp p.215) The universal envelopping algebra of a Lie algebra $(A, [])$ over the field K has the universal property that, whenever L is a unital algebra on the field K and $\rho : A \to L$ a map such that $\rho(X)\rho(Y) - \rho(Y)\rho(X) = \rho[X, Y]$, there is a unique algebra morphism $\widetilde{\rho}$ such that : $\widetilde{\rho} : U(A) \to L : \rho = \widetilde{\rho} \circ \imath$*

Properties

Theorem 1625 *Poincaré-Birkhoff-Witt (Knapp p.217):If A is a Lie algebra with basis $(e_i)_{i \in I}$ where the set I has some total ordering, then the set of monomials : $(\imath(e_{i_1}))^{n_1} (\imath(e_{i_2}))^{n_2} .... (\imath(e_{i_p}))^{n_p}, i_1 < i_2... < i_p \in I, n_1, ...n_p \in \mathbb{N}$ is a basis of its universal envelopping algebra U(A).*

Theorem 1626 *(Knapp p.216) **Transpose** is the unique automorphism ^{t} : $U(A) \to U(A)$ on the universal enveloping algebra U(A) of a Lie algebra : such that : $\imath(X)^t = -\imath(X)$*

Theorem 1627 *(Knapp p.492) If $(A, [])$ is a finite dimensional Lie algebra , then the following are equivalent for any element U of its universal envelopping algebra U(A):*

i) U is in the center of U(A)

ii) $\forall X \in A : XU = UX$

iii) $\forall X \in A : \exp(ad(X))(U) = U$

Theorem 1628 *(Knapp p.221) If B is a Lie subalgebra of A, then the associative subalgebra of U(A) generated by 1 and B is canonically isomorphic to U(B).*

So if $A = A_1 \oplus A_2$ then $U(A) \simeq U(A_1) \otimes_K U(A_2)$

And we have also : $U(A) \simeq S_{\dim A_1}(A_1) S_{\dim A_2}(A_2)$ with the symmetrization operator on U(A) :

$S_r(U) = \sum_{(i_1...i_r)} U^{i_1...i_r} \sum_{\sigma \in \mathfrak{s}_r} e_{\sigma(1)}..e_{\sigma(r)}$

Theorem 1629 *(Knapp p.230) If A is the Lie algebra of a Lie group G then U(A) can be identified with the left invariant differential operators on G, A can be identified with the first order operators .*

Theorem 1630 *If the Lie algebra A is also a Banach algebra (possibly infinite dimensional), then U(A) is a Banach algebra and a C*-algebra with the involution :* $* : U(A) \to U(A) : U^* = U^t$

Proof. the tensorial product is a Banach vector space and the map ı is continuous. ∎

Casimir elements

Casimir elements are commonly used to label representations of groups.

Definition 1631 *(Knapp p.293) If (F, ρ) is a finite dimensional representation of the finite dimensional Lie algebra $(A, [])$, and the Killing form B on A is non degenerate the Casimir element of order r is :*

$\Omega_r = \sum_{(i_1,...i_r)=1}^{n} Tr(\rho(e_{i_1} \cdot e_{i_2}... \cdot e_{i_{2r}})) \imath(E_{i_1})...\imath(E_{i_{2r}}) \in U(A)$

where :

$(e_i)_{i=1}^{n}$ is a basis of A

E_i is the vector of A such that : $B(E_i, e_j) = \delta_{ij}$

Warning ! the basis $(E_i)_{i=1}^{n}$ is a basis of A and not a basis of its dual A*, so Ω_1 is just an element of U(A) and not a bilinear form.

The matrix of the components of $(E_i)_{i=1}^{n}$ is just $[E] = [B]^{-1}$ where $[B]$ is the matrix of B in the basis. So the vectors $(E_i)_{i=1}^{n}$ are another basis of A

For r=1 and $\rho = Id$ we have :

$\Omega_1 = \sum_{i,j=1}^{n} B(e_i, e_j) \imath(E_i) \imath(E_j) \in U(A)$

The Casimir element has the following properties :

i) it does not depend of the choice of a basis

ii) it belongs to the center Z(U(A)) of U(A), so it commutes with any element of A

21.3   Classification of Lie algebras

Foundamental theorems

Theorem 1632 *Third Lie-Cartan theorem (Knapp p.663) : Every finite dimensional real Lie algebra is isomorphic to the Lie algebra of an analytic real Lie group.*

Theorem 1633 *Ado's Theorem (Knapp p.663) : Let A be a finite dimensional real Lie algebra, n its unique largest nilpotent ideal. Then there is a one-one finite dimensional representation (E,f) of A on a complex vector space E such that f(X) is nilpotent for every X in n. If A is complex then f can be taken complex linear.*

Together these theorems show that any finite dimensional Lie algebra on a field K is the Lie algebra of some group on the field K and can be represented as a Lie algebra of matrices. We will see that this not true for topological groups which are a much more diversified breed.

Thus the classification of finite dimensional Lie algebras is an endeavour which makes sense : the scope of the mathematical structures to explore is well delimited, and we have many tools to help our quest. But the path is not so easy and rather technical. However the outcome is much simpler, and that is the most important for all practical purposes.

The existence of brackets leads to some relations between elements of a Lie algebra. Indeed we have : $[e_i, e_j] = \sum C_{ij}^k e_k$ meaning that a vector of A can be defined, either as a linear combination of vectors of a basis (as in any vector space), or through the bracket operation. Thus we can consider the possibility to define a "set of generators", meaning a set of vectors such that, by linear combination or brackets, they give back A. This set of generators, if it exists, will be of a cardinality at most equal to the space vector dimension of A. This is the foundation of the classification of Lie algebras. We are lead to give a special attention to the subsets which are generated by the brackets of vectors (what we have previously denoted $[B, C]$).

The classification of Lie algebras is also the starting point to the linear representation of both Lie algebras and Lie groups and in fact the classification is based upon a representation of the algebra on itself throuh the operator ad.

The first step is to decompose Lie algebras in more elementary bricks, meaning simple Lie algebras.

21.3.1 Solvable and nilpotent algebras

Definition 1634 *For any Lie algebra $(A, [])$ we define the sequences :*
$A^0 = A \sqsupseteq A^1 = \left[A^0, A^0\right] \supseteq ... A^{k+1} = \left[A^k, A^k\right] ..$
$A_0 = A \supseteq A_1 = [A, A_0] ... \supseteq A_{k+1} = [A, A_k] , ...$

Theorem 1635 *(Knapp p.42) For any Lie algebra $(A, [])$:*
$A^k \sqsubseteq A_k$
Each A^k, A_k is an ideal of A.

Solvable algebra

Definition 1636 *A Lie algebra is said **solvable** if $\exists k : A^k = 0$. Then A^{k-1} is abelian.*

Theorem 1637 *Any 1 or 2 dimensional Lie algebra is solvable*

Theorem 1638 *If B is a solvable ideal and A/B is solvable, then A is solvable*

Theorem 1639 *The image of a solvable algebra by a Lie algebra morphism is solvable*

Theorem 1640 *(Knapp p.40) A n dimensional Lie algebra is solvable iff there is a decreasing sequence of subalgebras B_k :*
$B_0 = B \supseteq B_1... \supseteq B_{k+1}.. \supseteq B_n = 0$
such that B_{k+1} is an ideal of B_k and $dim(B_k/B_{k+1}) = 1$

Theorem 1641 *Cartan (Knapp p.50) : A finite dimensional Lie algebra is solvable iff its Killing form B is such that : $\forall X \in A, Y \in [A,A] : B(X,Y) = 0$*

Theorem 1642 *Lie (Knapp p.40) : If A is a solvable Lie algebra on a field K and (E,f) a finite dimensional representation of A, then there is a non null vector u in E which is a simulaneous eigen vector for $f(X), X \in A$ if all the eigen values are in the field K.*

So if A is solvable, it can be represented in a finite dimensional vector space as a set of triangular matrices.

Theorem 1643 *(Knapp p.32) If A is finite dimensional Lie algebra, there is a unique solvable ideal, called the **radical** of A which contains all the solvable ideals.*

Nilpotent algebra

Definition 1644 *A Lie algebra A is said **nilpotent** if $\exists k : A_k = 0$.*

Theorem 1645 *A nilpotent algebra is solvable, has a non null center Z(A) and $A_{k-1} \subseteq Z(A)$.*

Theorem 1646 *The image of a nilpotent algebra by a Lie algebra morphism is nilpotent*

Theorem 1647 *(Knapp p.46) A Lie algebra A is nilpotent iff $\forall X \in A : ad(X)$ is a nilpotent linear map (meaning that $\exists k : (adX)^k = 0$).*

Theorem 1648 *(Knapp p.49) Any finite dimensional Lie algebra has a unique largest nilpotent ideal n, which is contained in the radical rad(A) of A and [A,rad(A)]$\sqsubseteq$ n. Any derivation D is such that D(rad(A))$\sqsubseteq$n.*

Theorem 1649 *(Knapp p.48, 49) A finite dimensional solvable Lie algebra A :*
i) has a unique largest nilpotent ideal n, namely the set of elements X for which ad(X) is nilpotent. For any derivation D : D(A)$\sqsubseteq$ n
ii) [A,A] is nilpotent

Theorem 1650 *Engel (Knapp p.46) If E is a finite dimensional vector space, A a sub Lie algebra of L(E;E) of nilpotent endomorphisms, then A is nilpotent and there is a non null vector u of E such that f(u)=0 for any f in A. There is a basis of E such that the matrices of f are triangular with 0 on the diagonal.*

21.3.2 Simple and semi-simple Lie algebras

Definition 1651 *A Lie algebra is :*
  simple *if it is non abelian and has no non zero ideal.*
  semi-simple *if it has no non zero solvable ideal.*

A simple algebra is semi-simple, the converse is not true.

Theorem 1652 *(Knapp p.33) If A is simple then $[A, A] = A$*

Theorem 1653 *(Knapp p.33) Every semi-simple Lie algebra has for center Z(A)=0.*

Theorem 1654 *(Knapp p.32) A finite dimensional Lie algebra A is semi-simple iff rad(A)=0.*

There are no complex semi-simple Lie algebra of dimension 4,5 or 7.

Theorem 1655 *(Knapp p.54) If A_0 is the real form of a complex Lie algebra A, then A_0 is a semi simple real Lie algebra iff A is a semi simple complex Lie algebra.*

Theorem 1656 *(Knapp p.50,54) For a finite dimensional Lie algebra A :*
  i) A/rad(A) is semi-simple.
  ii) A is semi-simple iff the Killing form B is non degenerate
  iii) A is solvable iff the Killing form B is such that : $\forall X \in A, Y \in [A, A]$: $B(X, Y) = 0$
  iv) A is semi simple iff $A = A_1 \oplus A_2 ... \oplus A_k$ where the A_i are ideal and simple Lie subalgebras. Then the decomposition is unique and the only ideals of A are sum of some A_i.

Thus the way to classify finite dimensional Lie algebras is the following :
i) for any algebra H=A/rad(A) is semi-simple.
ii) rad(A) is a solvable Lie algebra, and can be represented as a set of triangular matrices
iii) any semi simple algebra is the sum of simple Lie algebras
iv) from H we get back A by semi-direct product
If we have a classification of simple Lie algebras we have a classification of all finite dimensional Lie algebras.

21.3.3 Abstract roots system

The classification is based upon the concept of abstract roots system, which is...abstract and technical, but is worth a look because it is used extensively in many topics related to the representation of groups. We follow Knapp (II p.124).

Abstract roots system
 1. Definition:

Definition 1657 *An **abstract roots system** is a finite set Δ of non null vectors of a finite dimensional real vector space $(V, \langle \rangle)$ endowed with an inner product (definite positive), such that :*

i) Δ spans V

ii) $\forall \alpha, \beta \in \Delta : 2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} \in \mathbb{N}$

iii) the set W of linear maps on V defined by $s_\alpha(\beta) = \beta - 2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}\alpha$ for $\alpha \in \Delta$ carries Δ on itself

   W is a subgroup of the orthonormal group of V, comprised of reflexions, called the **Weyl's group**.
   The vectors of Δ are linearly dependant, and the identities above are very special, and indeed they can be deduced from only 9 possible configurations. The next step is to simplify Δ.

 2. Reduced system:
   It is easy to see that any integer multiple of the vectors still meet the identities. So we can restrict a bit the definition :

Definition 1658 *An abstract roots sytem is said to be **reduced** if $\alpha \in \Delta \Rightarrow 2\alpha \notin \Delta$.*

   Most of the results will be given for reduced roots system.

Definition 1659 *An abstract roots system is said to be **reducible** if it is the direct sum of two sets, which are themselves abstract roots systems, and are orthogonal. It is **irreducible** if not.*

   An irreducible system is necessarily reduced, but the converse is not true.
   The previous conditions give rise to a great number of identities. The main result is that $\alpha \in \Delta \Rightarrow -\alpha \in \Delta$ even if the system is reduced.

 3. Ordering:

Definition 1660 *An **ordering** on a finite dimensional real vector space is an order relation where :*

a set of positive vectors is identified

if u is in V, either u or -u is positive

if u and v are positive then u+v is positive.

There are many ways to achieve that, the simplest is the lexicographic ordering. Take a basis $(e_i)_{i=1}^l$ of V and say that u>0 if there is a k such that $\langle u, e_i \rangle = 0$ for i=1...k-1 and $\langle u, e_k \rangle > 0$.

4. Simple system:

Definition 1661 *A root $\alpha \in \Delta$ of an ordered abstract root system is **simple** if $\alpha > 0$ and there is no $\beta, \beta' > 0$ such that $\alpha = \beta + \beta'$.*

Theorem 1662 *For any abstract roots system Δ there is a set $\Pi = (\alpha_1, ..., \alpha_l)$ with l=dim(V) of linearly independant simple roots, called a **simple system**, which fully defines the system:*
i) for any root $\beta \in \Delta : \exists (n_i)_{i=1}^l, n_i \in \mathbb{N} :$ either $\beta = \sum_{i=1}^l n_i \alpha_i$ or $\beta = -\sum_{i=1}^l n_i \alpha_i$
ii) W is generated by the $s_{\alpha_i}, i = 1...l$
iii) $\forall \alpha \in \Delta, \exists w \in W, \alpha_i \in \Pi : \alpha = w\alpha_i$
iv) if Π, Π' are simple systems then $\exists w \in W, w$ unique $: \Pi' = w\Pi$

The set of positive roots is denoted $\Delta^+ = \{\alpha \in \Delta : \alpha > 0\}$.
Remarks :
i) $\Pi \sqsubset \Delta^+ \sqsubset \Delta$ but Δ^+ is usually larger than Π
ii) any $\pm \left(\sum_{i=1}^l n_i \alpha_i \right)$ does not necessarily belong to Δ

Definition 1663 *A vector $\lambda \in V$ is said to be **dominant** if : $\forall \alpha \in \Delta^+ : \langle \lambda, \alpha \rangle \geq 0$.*

For any vector λ of V there is always a simple system Π for which it is dominant, and there is always $w \in W$ such that $w\lambda$ is dominant.

Abstract Cartan matrix
The next tool is a special kind of matrix, adjusted to abstract roots systems.
1. For an abstract root system represented by a simple system $\Pi = (\alpha_1, ..., \alpha_l)$ the matrix :
$[A]_{ij} = 2\frac{\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_i, \alpha_i \rangle}$
has the properties :
i) $[A]_{ij} \in \mathbb{Z}$
ii) $[A]_{ii} = 2; \forall i \neq j : [A]_{ij} \leq 0$
iii) $[A]_{ij} = 0 \Leftrightarrow [A]_{ji} = 0$
iv) there is a diagonal matrix D with positive elements such that DAD^{-1} is symmetric definite positive.
v) does not depend ot the choice of positive ordering, up to a permutation of the indices (meaning up to conjugation by permutation matrices).
2. A matrix meeting the conditions i) through iv) is called a **Cartan matrix**. To any Cartan matrix one can associate a unique simple system, thus a reduced abstract root system, unique up to isomorphism.

3. By a permutation of rows and columns it is always possible to bring a Cartan matrix in the block diagonal form : a triangular matrix which is the assembling of triangular matrices above the main diagonal. The matrix has a unique block iff the associated abstract root system is irreducible and then is also said to be irreducible.

The diagonal matrix D above is unique up a multiplicative scalar on each block, thus D is unique up to a multiplicative scalar if A is irreducible.

Dynkin's diagram

Dynkin's diagram are a way to represent abstract root systems. They are also used in the representation of Lie algebras.

1. The **Dymkin's diagram** of a simple system of a reduced abstract roots Π is built as follows :

i) to each simple root α_i we associate a vertex of a graph

ii) to each vertex we associate a weigth $w_i = k \langle \alpha_i, \alpha_i \rangle$ where k is some fixed scalar

iii) two vertices i, j are connected by $[A]_{ij} \times [A]_{ji} = 4 \frac{\langle \alpha_i, \alpha_j \rangle^2}{\langle \alpha_i, \alpha_i \rangle \langle \alpha_j, \alpha_j \rangle}$ edges

The graph is connected iff the system is irreducible.

2. Conversely given a Dynkin's diagram the matrix A is defined up to a multiplicative scalar for each connected component, thus it defines a unique reduced abstract root system up to isomorphism.

3. There are only 9 types of connected Dynkin's diagrams, which define all the irreducible abstract roots systems. They are often represented in books about Lie groups (see Knapp p.182).

4. To understand the usual representation of Dynkin's diagram :

i) the abstract roots system is a set of vectors of a finite dimensional real vector space V. So, up to isomorphism, we can take V as a n dimensional subspace of $\mathbb{R}^m$ for some m (it is almost always $\mathbb{R}^m$ itself but there are exceptions).

ii) we can define the inner product $\langle \rangle$ through an orthonormal basis, so we take the canonical basis $(e_i)_{i=1}^m$ of $\mathbb{R}^m$ with the usual euclidian inner product.

iii) thus a simple roots system is defined as a special linear combination of the $(e_i)_{i=1}^m$

5. The 9 irreducible roots systems are the following :

a) $A_n : n \geq 1 : V = \sum_{k=1}^{n+1} x_k e_k, \sum_{k=1}^{n+1} x_k = 0$
$\Delta = e_i - e_j, i \neq j$
$\Pi = \{e_1 - e_2, e_2 - e_3, .. e_n - e_{n+1}\}$
b) $B_n : n \geq 2 : V = \mathbb{R}^n$
$\Delta = \{\pm e_i \pm e_j, i < j\} \cup \{\pm e_k\}$
$\Pi = \{e_1 - e_2, e_2 - e_3, .. e_{n-1} - e_n, e_n\}$
c) $C_n : n \geq 3 : V = \mathbb{R}^n$
$\Delta = \{\pm e_i \pm e_j, i < j\} \cup \{\pm 2e_k\}$
$\Pi = \{e_1 - e_2, e_2 - e_3, .. e_{n-1} - e_n, 2e_n\}$
d) $D_n : n \geq 4 : V = \mathbb{R}^n$

$\Delta = \{\pm e_i \pm e_j, i < j\}$

$\Pi = \{e_1 - e_2, e_2 - e_3, ..e_{n-1} - e_n, e_{n-1} + e_n\}$

e) 5 exceptional types (with the dimension of V) : $E_6\,(6)\,, E_7\,(7)\,, E_8\,(8)\,, F_4\,(4)\,, G_2\,(2)$

21.3.4 Classification of semi-simple complex Lie algebras

The procedure is to exhibit for any complex semi-simple Lie algebra an abstract roots system, which gives also a set of generators of the algebra. And conversely to prove that a Lie algebra can be associated to any abstract roots system.

In the following A is a complex semi-simple finite dimensional Lie algebra with dimension n.

Cartan subalgebra

Definition 1664 *A **Cartan subalgebra** of a complex Lie algebra $(A, [])$ is an abelian subalgebra h such that there is a set of linearly independant eigen vectors X_k of A such that $A = Span(X_k)$ and $\forall H \in h : ad_H X_k = \lambda\,(H)\,X_k$ with an eigen value $\lambda\,(H)$ which can depend on H*

We assume that h is a maximal Cartan algebra : it does not contain any other subset with these properties.

Theorem 1665 *(Knapp p.134) Any complex semi-simple finite dimensional Lie algebra A has a Cartan subalgebra. All Cartan subalgebras of A have the same dimension, called the rank of A*

Cartan subalgebras are not necessarily unique. If h,h' are Cartan subalgebras of A then there is some automorphism $a \in Int(A) : h' = ah$

Definition 1666 *A Cartan subalgebra h of a real Lie algebra $(A, [])$ is a subalgebra of h such that the complexified of h is a Cartan subalgebra of the complexified of A.*

Theorem 1667 *(Knapp p.384) Any real semi simple finite dimensional Lie algebra has a Cartan algebra . All Cartan subalgebras have the same dimension.*

Root-space decomposition

1. Let h be a Cartan subalgebra, then we have the following properties :

i) h itself is an eigen space of ad_H with eigen value 0 : indeed $\forall H, H' \in h :$ $[H, H'] = ad_H H' = 0$

ii) the eigenspaces for the non zero eigen values are unidimensional

iii) the non zero eigen values are linear functions $\alpha\,(H)$ of the vectors H

2. Thus there are n linearly independant vectors denoted X_k such that :

for k=1...l $(X_k)_{k=1}^l$ is a basis of h and $\forall H \in h : ad_H X_k = [H, X_k] = 0$

for k=l+1,...n : $\forall H \in h : ad_H X_k = [H, X_k] = \alpha_k\,(H)\,X_k$ where : $\alpha_k : h \to \mathbb{C}$ is linear, meaning that $\alpha_k \in h^*$

These functions do not depend on the choice of X_k in the eigenspace because they are unidimensional. Thus it is customary to label the eigenspaces by the function itself : indeed they are no more than vectors of the dual, and we have exactly n-l of them. And one writes :

$\Delta(h) = \{\alpha_k \in h^*\}$

$A_\alpha = \{X \in A : \forall H \in h : ad_H X = \alpha(H) X\}$

$A = h \oplus_{\alpha \in \Delta} A_\alpha$

The functionals $\alpha \in \Delta$ are called the **roots**, the vectors of each A_α are the **root vectors**, and the equation above is the root-space decomposition of the algebra.

3. Let B be the Killing form on A. Because A is semi-simple B is non degenerate thus it can be used to implement the duality between A and A*, and h and h*, both as vector spaces on $\mathbb{C}$.

Then : $\forall H, H' \in h : B(H, H') = \sum_{\alpha \in \Delta} \alpha(H) \alpha(H')$

Define :

i) V the linear real span of Δ in h*: $V = \{\sum_{\alpha \in \Delta} x_\alpha \alpha; x_\alpha \in \mathbb{R}\}$

ii) the n-l B-dual vectors H_α of α in h : $H_\alpha \in h : \forall H \in h : B(H, H_\alpha) = \alpha(H)$

iii) the bilinear symmetric form in V :

$\langle H_\alpha, H_\beta \rangle = B(H_\alpha, H_\beta) = \sum \sum_{\gamma \in \Delta} \gamma(H_\alpha) \gamma(H_\beta)$

$\langle u, v \rangle = \sum_{\alpha, \beta \in \Delta} x_\alpha y_\beta B(H_\alpha, H_\beta)$

iv) h_0 the real linear span of the H_α : $h_0 = \{\sum_{\alpha \in \Delta} x_\alpha H_\alpha; x_\alpha \in \mathbb{R}\}$

Then :

i) V is a real form of h*: $h^* = V \oplus iV$

ii) h_0 is a real form of h: $h = h_0 \oplus ih_0$ and V is exactly the set of covectors such that $\forall H \in h_0 : u(H) \in \mathbb{R}$ and V is real isomorphic to h_0^*

iii) $\langle \rangle$ is a definite positive form, that is an inner product, on V

iv) the set Δ is an abstract roots system on V, with $\langle \rangle$

v) up to isomorphism this abstract roots system does not depend on a choice of a Cartan algebra

vi) the abstract root system is irreducible iff the Lie algebra is simple

4. Thus, using the results of the previous subsection there is a simple system of roots $\Pi = (\alpha_1, ... \alpha_l)$ with l roots, because $V = span_\mathbb{R}(\Delta), \dim_\mathbb{R} V = \dim_\mathbb{C} h^* = \dim_\mathbb{C} h = l$

Define :

$h_i = \frac{2}{\langle \alpha_i, \alpha_i \rangle} H_{\alpha_i}$

$E_i \neq 0 \in A_{\alpha_i} : \forall H \in h : ad_H E_i = \alpha_i(H) E_i$

$F_i \neq 0 \in A_{-\alpha_i} : \forall H \in h : ad_H F_i = -\alpha_i(H) F_i$

Then the set $\{h_i, E_i, F_i\}_{i=1}^l$ generates A as a Lie algebra (by linear combination and bracket operations). As the abstract roots systems have been classified, so are the semi-simple complex Lie algebras.

So *a semi-simple complex Lie algebra has a set of at most 3 × rank generators*. But notice that it can have fewer generators : indeed if dim(A)=n<3l.

This set of generators follows some specific identities, called Serre's relations, expressed with a Cartan matrix C, which can be useful (see Knapp p.187):

$$[h_i, h_j] = 0$$
$$[E_i, F_j] = \delta_{ij} h_i$$
$$[h_i, E_j] = C_{ij} E_j$$
$$[h_i, F_j] = -C_{ij} F_j$$
$$(adE_i)^{-C_{ij}+1} E_j = 0 \text{ when } i \neq j$$
$$(adF_i)^{-A_{ij}+1} F_j = 0 \text{ when } i \neq j$$

Example sl($\mathbb{C}$,3)

This the Lie algebra of 3x3 complex matrices such that Tr(X)=0. It is a 8 dimensional, rank 2 algebra.

Its Killing form is : $X, Y \in sl(\mathbb{C}, 3) : B(X, Y) = 6Tr(XY)$

The Cartan algebra comprises of diagonal matrices and we take as basis of h :

$$H_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} ; H_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

and take as other elements of a basis of sl(c,3) :

$$E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} ; E_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} ; E_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$F_1 = E_1^t; F_2 = E_2^t; F_3 = E_3^t$

E_1, E_2, E_3 are common eigen vectors of ad(h) :

$ad(H_1) E_1 = 2E_1 = \alpha_1(H_1) E_1; ad(H_2) E_1 = -E_1 = \alpha_1(H_2) E_1$

$ad(H_1) E_2 = -E_2 = \alpha_2(H_1) E_2; ad(H_2) E_2 = 2E_2 = \alpha_2(H_2) E_2$

$ad(H_1) E_3 = E_3 = \alpha_3(H_1) E_3; ad(H_2) E_3 = E_3 = \alpha_3(H_2) E_3$

It is easy to see that : $ad_H F_i = -\alpha_i(H) F_i$

Each of the root spaces are generated by one of the vectors $E_1, E_2, E_3, F_1, F_2, F_3$

So : $sl(\mathbb{C}, 3) = h \oplus \left(\oplus_{i=1}^3 E_i \right) \oplus \left(\oplus_{i=1}^3 F_i \right)$

We have the roots : $\Delta = \{\alpha_1, \alpha_2, \alpha_3, -\alpha_1, -\alpha_2, -\alpha_3\}$ which are not independant : we have also $\alpha_3 = \alpha_1 + \alpha_2$

A simple system is given by : $\Pi = \{\alpha_1, \alpha_2\}$ and the positive racines are : $\Delta^+ = \{\alpha_1, \alpha_2, \alpha_3\}$ and the negative : $\Delta^- = \{-\alpha_1, -\alpha_2, -\alpha_3\}$

The generators of sl(C,3) are : $\{H_i, E_i, F_i\}_{i=1,2}$

Take as basis of the dual vector space sl(C,3)* : $(\lambda_i)_{i=1}^8 : i = 1, 2, 3 : \lambda_i(X) = [X]^i$ for $X = H_1, H_2$

$\lambda_i \left(\sum_{j=1}^2 x_j H_j + \sum_{j=1}^3 y_j E_j + z_j F_j \right) = \sum_{j=1}^2 x_i \lambda_i(H_j) + \sum_{j=1}^3 y_j \lambda_i(E_j) + z_j \lambda_i(F_j)$

then: $\alpha_1 = \lambda_1 - \lambda_2, \alpha_2 = \lambda_2 - \lambda_3, \alpha_3 = \lambda_1 - \lambda_3$

The λ_i play the same role as the e_i in the Dynkin diagram.

Classification of semi-simple complex Lie algebras

1. The classification of Lie algebras follows the classification of abstract roots systems (Knapp p.683). The Lie algebras are expressed as matrices algebras in their standard linear representation (see below):

$A_n, n \geq 1 : sl(n+1, C)$
$B_n, n \geq 2 : so(2n+1, C)$
$C_n, n \geq 3 : sp(n, C)$
$D_n, n \geq 4 : so(2n, C)$

The exceptional systems give rise to 5 exceptional Lie algebras (their dimension is in the index) :E_6, E_7, E_8, F_4, G_2

2. Conversely if we start with an abstract roots system it can be proven :

i) Given an abstract Cartan matrix C there is a complex semi-simple Lie algebra whose roots system has C as Cartan matrix

ii) and that this Lie algebra is unique, up to isomorphism. More precisely :

let A,A' be complex semi-simple algebras with Cartan subalgebras h, h', and roots systems Δ, Δ'. Assume that there is a vector space isomorphism :$\varphi : h \to h'$ such that its dual $\varphi^* : h'^* \to h^* :: \varphi^* (\Delta') = \Delta$. For $\alpha \in \Delta$ define $\alpha' = \varphi^{*-1} (\alpha)$. Take a simple system $\Pi \subset \Delta$, root vectors $E_\alpha, E_{\alpha'}$ then there is one unique Lie algebra isomorphism $\Phi : A \to A'$ such that $\Phi|_h = \varphi$ and $\Phi (E_\alpha) = E_{\alpha'}$.

3. Practically : *usually there is no need for all the material above.* We know that any finite dimensional Lie algebra belongs to one the 9 types above, and we can proceed directly with them.

21.3.5   Compact algebras

This the only topic for which we use analysis concepts in Lie algebra study.

Definition of Int(A)

Let A be a Lie algebra such that A is also a Banach vector space (it will be the case if A is a finite dimensional vector space). Then :

i) For any continuous map f$\in \mathcal{L}(A; A)$ the map : $\exp f = \sum_{n=0}^{\infty} \frac{1}{n!} f^n \in \mathcal{L}(A; A)$ (f^n is the n iterate of f) is well defined (see Banach spaces) and has an inverse.

ii) If f is a continuous morphism then $f^n ([X, Y]) = [f^n (X), f^n (Y)]$ and $\exp(f)$ is a continuous automorphism of A

iii) if for any X the map ad(X) is continuous then : $\exp ad(X)$ is a continuous automorphism of A

iv) the subset of continuous (then smooth) automorphisms of A is a Lie group whose connected component of the identity is denoted **Int(A)**.

Compact Lie algebra

Definition 1668 *A Banach Lie algebra A is said to be* **compact** *if Int(A) is compact with the topology of $\mathcal{L}(A; A)$.*

Int(A) is a manifold, if it is compact, it must be locally compact, so it cannot be infinite dimensional. Therefore there is no compact infinite dimensional Lie algebra. Moreover :

Theorem 1669 *(Duistermaat p.151) A compact complex Lie algebra is abelian.*

So the story of compact Lie algebras is limited to real finite dimensional Lie algebras.

Theorem 1670 *The Lie algebra of a compact Lie group is compact.*

Theorem 1671 *(Duistermaat p.149) For a real finite dimensional Lie algebra A the following are equivalent :*
i) A is compact
ii) its Killing form is negative semi-definite and its kernel is the center of A
iii) A is the Lie algebra of a compact group

Theorem 1672 *(Duistermaat p.151) For a real finite dimensional Lie algebra A the following are equivalent :*
i) A is compact and semi-simple
ii) A is compact and has zero center
iii) its Killing form is negative definite
iv) Every Lie group with Lie algebra Lie isomorphic to A is compact

The simplest criterium to identify compact algebras lays upon the Killing form :
i) the Lie algebra of a real compact Lie group is always compact and its Killing form is *semi-definite* negative.
ii) conversely if the Killing form of a real Lie algebra A is *negative definite* then A is compact (and semi-simple).

21.3.6   Structure of semi-simple real Lie algebra

The structure and classification of real Lie algebras are a bit more complicated than complex ones. They are based upon the fact that any semi-simple complex Lie algebra can be decomposed into two compact real Lie algebras.

Compact real forms

Theorem 1673 *(Knapp p.434) The isomorphisms classes of compact, semi simple real finite dimensional Lie algebras A_0 and the isomorphisms classes of complex semi simple finite dimensional Lie algebras A are in one-one correspondance : A is the complexification of A_0 and A_0 is a compact real form of A. Under this correspondance simple Lie algebras correspond to simple Lie algebras.*

So *any finite dimensional complex semi-simple Lie algebra A has a compact real form u_0.* A can be written : $A = u_0 \oplus iu_0$ where u_0 is a real compact Lie algebra. Any two compact real forms are conjugate via Int(A)

Using the previous classification we can list *all the real forms* : $A \rightarrow u_0$

$A_n, n \geq 1 : sl(n+1, \mathbb{C}) \rightarrow su(n+1, \mathbb{C})$

$B_n, n \geq 2 : so(2n+1, \mathbb{C}) \rightarrow so(2n+1, \mathbb{R})$

$C_n, n \geq 3 : sp(n, \mathbb{C}) \rightarrow sp(n, H) \simeq sp(n, \mathbb{C}) \cap u(2n)$

$D_n, n \geq 4 : so(2n, \mathbb{C}) \rightarrow so(2n, \mathbb{R})$

and similarly for the exceptional Lie algebras (Knapp p.413).

Given an abstract Cartan matrix C there is a unique, up to isomorphism, compact real semi simple algebra such that its complexified has C as Cartan matrix.

Cartan involution

Definition 1674 *A **Cartan involution** on a real semi-simple Lie algebra A is an automorphism on A, such that $\theta^2 = Id$ and $B_\theta : B_\theta(X, Y) = -B(X, \theta Y)$ is positive definite.*

Let A be a semi-simple complex Lie algebra, u_0 its compact real form. Then $\forall Z \in A, \exists x, y \in u_0 : Z = x + iy$.

Define :$\theta : A \rightarrow A :: \theta(Z) = x - iy$ this is a Cartan involution on $A_\mathbb{R} = (u_0, u_0)$ and all the Cartan involutions are of this kind.

Theorem 1675 *(Knapp p.445) Any real semi-simple finite dimensional Lie algebra A_0 has a Cartan involution. And any two Cartan involutions are conjugate via Int(A_0).*

Cartan decomposition:

Definition 1676 *A Cartan decomposition of a real finite dimensional Lie algebra $(A, [])$ is a pair of vector subspaces l_0, p_0 of A such that :*

i) $A = l_0 \oplus p_0$

ii) l_0 is a subalgebra of A

iii) $[l_0, l_0] \sqsubseteq l_0, [l_0, p_0] \sqsubseteq p_0, , [p_0, p_0] \sqsubseteq l_0,$

iv) the Killing form B of A is negative definite on l_0 and positive definite on $p._0$

v) l_0, p_0 are orthogonal under B and B_θ

Theorem 1677 *Any real semi-simple finite dimensional Lie algebra A has a Cartan decomposition*

Proof. Any real semi-simple finite dimensional Lie algebra A has a Cartan involution θ, wich has two eigenvalues : ± 1. Taking the eigenspaces decomposition of A with respect to $\theta :\theta(l_0) = l_0, \theta(p_0) = -p_0$ we have a Cartan decomposition. ∎

Moreover l_0, p_0 are orthogonal under B_θ.

Conversely a Cartan decomposition gives a Cartan involution with the definition $\theta = +Id$ on l_0 and $\theta = -Id$ on p_0

If $A = l_0 \oplus p_0$ is a Cartan decomposition of A, then the real Lie algebra $A = l_0 \oplus ip_0$ is a compact real form of the complexified $(A_0)_{\mathbb{C}}$.

Theorem 1678 *(Knapp p.368) A finite dimensional real semi simple Lie algebra is isomorphic to a Lie algebra of real matrices that is closed under transpose. The isomorphism can be specified so that a Cartan involution is carried to negative transpose.*

Classification of simple real Lie algebras

The procedure is to go from complex semi simple Lie algebras to real forms by Cartan involutions. It uses Vogan diagrams, which are Dynkin diagrams with additional information about the way to get the real forms. The results are the following (Knapp p.421) :

Up to isomorphism, any *simple* real finite dimensional Lie algebra belongs to one the following types :

i) the real structures of the complex semi simple Lie algebras (considered as real Lie algebras) :

$A_n, n \geq 1 : su(n+1, \mathbb{C}) \oplus isu(n+1, \mathbb{C})$
$B_n, n \geq 2 : so(2n+1, \mathbb{R}) \oplus iso(2n+1, \mathbb{R})$
$C_n, n \geq 3 : (sp(n, \mathbb{C}) \cap u(2n)) \oplus i\,(sp(n, \mathbb{C}) \cap u(2n))$
$D_n, n \geq 4 : so(2n, \mathbb{R}) \oplus iso(2n, \mathbb{R})$
and similarly for E_6, E_7, E_8, F_4, G_2

ii) the compact real forms of the complex semi simple Lie algebras
$A_n, n \geq 1 : su(n+1, \mathbb{C})$
$B_n, n \geq 2 : so(2n+1, \mathbb{R})$
$C_n, n \geq 3 : sp(n, \mathbb{C}) \cap u(2n)$
$D_n, n \geq 4 : so(2n, \mathbb{R})$
and similarly for E_6, E_7, E_8, F_4, G_2

iii) the classical matrix algebras :
$su(p, q, \mathbb{C}) : p \geq q > 0, p + q > 1$
$so(p, q, \mathbb{R}) : p > q > 0, p + q$ *odd and* > 4 *or* $p \geq q > 0, p + q$ *even* $p + q > 7$
$sp(p, q, H) : p \geq q, p + q > 2$
$sp(n, R) : n > 2$
$so^*(2n, \mathbb{C}) : n > 3$
$sl(n, \mathbb{R}) : n > 2$
$sl(n, H) : n > 1$

iv) 12 non complex, non compact exceptional Lie algebras (p 416)

22 LIE GROUPS

22.1 General definitions and results

22.1.1 Definition of Lie groups

Whereas Lie algebras involve quite essentially only algebraic tools, at the group level analysis is of paramount importance. And there is two ways to deal with this : with simple topological structure and we have the topological groups, or with manifold structure and we have the Lie groups. As it is common in the litterature to mix both cases, it is necessary to understand the difference between both structures.

Topological group

Definition 1679 *A **topological group** is a Hausdorff topological space, endowed with an algebraic structure such that the operations product and inverse are continuous.*

With such structure we can handle all the classic concepts of general topology : convergence, integration, continuity of maps over a group,...What we will miss is what is related to derivatives. Of course Lie algebras can be related to Lie groups only.

Definition 1680 *A discrete group is a group endowed with the discrete topology.*

Any set endowed with an algebraic group structure can be made a topological group with the discrete topology.A discrete group which is second-countable has necessarily countably many elements. A discrete group is compact iff it is finite. A finite topological group is necessarily discrete.

Theorem 1681 *(Wilansky p.240) Any product of topological groups is a topological group*

Theorem 1682 *(Wilansky p.243) A topological group is a regular topological space*

Theorem 1683 *(Wilansky p.250) A locally compact topological group is paracompact and normal*

Lie group

Definition 1684 *A **Lie group** is a class r manifold G, modeled on a Banach space E over a field K, endowed with a group structure such that the product and the inverse are class r maps.*

Moreover we will assume that G is a normal, Hausdorff, second countable topological space, which is equivalent to say that G is a metrizable, separable manifold (see Manifolds).

The manifold structure (and thus the differentiability of the operations) are defined with respect to the field K. As E is a Banach we need the field K to be complete (practically K=$\mathbb{R}$ or $\mathbb{C}$). While a topological group is not linked to any field, a Lie group is defined over a field K, through its manifold structure that is necessary whenever we use derivative on G.

The **dimension of the Lie group** is the dimension of the manifold. Notice that we do not assume that the manifold is finite dimensional : we will precise this point when it is necessary. Thus if G is infinite dimensional, following the Henderson theorem, it can be embedded as an open subset of an infinite dimensional, separable, Hilbert space.

For the generality of some theorems we take the convention that finite groups with the discrete topology are Lie groups of dimension zero.

Lie groups are locally connected, but usually not connected. The **connected component of the identity**, denoted usually G_0 is of a particular importance. This is a group and a manifold, so it has its own Lie group structure (we will see that it is a Lie subgroup of G).

Example : $GL(\mathbb{R}, 1) = (\mathbb{R}, \times)$ has two connected components, $GL_0(\mathbb{R}, 1) = \{x, x > 0\}$

We will denote :

the operation $G \times G \to G :: xy = z$

the inverse : $G \to G :: x \to x^{-1}$

the unity : 1

There are some general theorems :

A Lie group is locally compact iff it is finite dimensional.

If G has a complex manifold structure then it is a smooth manifold, and the operations being C-differentiable are holomorphic.

Theorem 1685 *Montgomery and Zippin (Kolar p.43) If G is a separable, locally compact topological group, with a neighbourhood of 1 which does not contain a proper subgroup then G is a Lie group .*

Theorem 1686 *Gleason, Montgomery and Zippin (Knapp p.99 for the real case) : For a real finite dimensional Lie group G there is exactly one analytic manifold structure on G which is consistent with the Lie group structure*

As the main advantage of the manifold structure (vs the topological structure) is the use of derivatives in the following we will always assume that a Lie group has the structure of a smooth (real or complex) manifold, with smooth group operations.

Warning ! Mathematicians who are specialists of Lie groups use freely the name "analytic group". For Knapp "an analytic group is a connected Lie group" (p.69). This is rather confusing. The manifold structure to be analytic is a property which is rarely used, whereas it is really useful that it is smooth. So a

smooth manifold will do, and if the analycity is required it is possible to revert to the theorems above. On the other hand few Lie groups are connected, and many theorems stand only for connected Lie groups. To use these theorems one takes the component of the identity (which is the most useful part of the group) and extend the results by some products. Similarly we will always specify if the manifold structure is real or complex, and finite or infinite dimensional if there is some restriction. If there is no such specification that means that the definition or the result stands for any Lie group, real or complex, finite or infinite dimensional. So here :

Lie group = any kind of smooth manifold with smooth group operations

Connected Lie group = any kind of Lie group whose manifold is connected

Real Lie group = Lie group whose manifold is modelled on a Banach real vector space

Complex Lie group = Lie group whose manifold is modelled on a Banach complex vector space and whose operations are C-differentiable

Finite dimensional Lie group = Lie group whose manifold is finite dimensional

Examples of Lie groups

1. The group GL(K,n) of square nxn inversible matrices over a field K : it is a vector subspace of K^{n^2} which is open as the preimage of $\det X \neq 0$ so it is a manifold, and the operations are smooth.

2. A Banach vector space is an abelian Lie group with addition

3. Let E be a Banach vector space, then the set $\mathcal{L}(E; E)$ of continuous linear map is a Banach vector space, thus a manifold. The set $G\mathcal{L}(E; E)$ of continuous automorphisms over E is an open subset of $\mathcal{L}(E; E)$ thus a manifold. It is a group with the composition law and the operations are differentiable (see derivatives). So $G\mathcal{L}(E; E)$ is a Lie group (but not a Banach algebra).

22.1.2   Translations

Basic operations

1. The **translations** over a group are just the rigth (R) and left (L) products (same definition as for any group - see Algebra). They are smooth diffeomorphisms. There is a commonly used notation for them :

Notation 1687 R_a *is the right multiplication by a :* $R_a : G \to G : R_a x = xa$
L_a *is the left multiplication by a :* $L_a : G \to G :: L_a x = ax$

and $R_a x = xa = L_x a$
These operations commute : $L_a \circ R_b = R_b \circ L_a$
Because the product is associative we have the identities :
$abc = R_c ab = R_c (R_a b) = L_a bc = L_a (L_b c)$
$L_{ab} = L_a \circ L_b; R_a \circ R_b = R_{ab};$
$L_{a^{-1}} = (L_a)^{-1}; R_{a^{-1}} = (R_a)^{-1};$
$L_{a^{-1}}(a) = 1; R_{a^{-1}}(a) = 1$

$L_1 = R_1 = Id$

2. The **conjugation** with respect to a is the map : $Conj_a : G \to G ::$
$Conj_a x = axa^{-1}$

Notation 1688 $Conj_a x = L_a \circ R_{a^{-1}}(x) = R_{a^{-1}} \circ L_a(x)$

If the group is commutative then $Conj_a x = x$
Conjugation is an inversible map.

Derivatives

1. If G is a Lie group all these operations are smooth diffeomorphisms over G so we have the linear bijective maps :

Notation 1689 $L'_a x$ *is the derivative of* $L_a (g)$ *with respect to g, at g=x;* $L'_a x \in GL(T_x G; T_{ax} G)$
$R'_a x$ *is the derivative of* $R_a (g)$ *with respect to g, at g=x;* $R'_a x \in GL(T_x G; T_{xa} G)$

2. The product can be seen as a two variables map : $M : G \times G \to G :$
$M(x, y) = xy$ with partial derivatives :
$u \in T_x G, v \in T_y G : M'(x, y)(u, v) = R'_y(x)u + L'_x(y)v \in T_{xy} G$
$\frac{\partial}{\partial x}(xy) = \frac{\partial}{\partial z}(R_y(z))|_{z=x} = R'_y(x)$
$\frac{\partial}{\partial y}(xy) = \frac{\partial}{\partial z}(L_x(z))|_{z=y} = L'_x(y)$
Let g,h be differentiable maps $G \to G$ and :
$f : G \to G :: f(x) = g(x)h(x) = M(g(x), h(x))$
$f'(x) = \frac{d}{dx}(g(x) h(x)) = R'_{h(x)}(g(x)) \circ g'(x) + L'_{g(x)}(h(x)) \circ h'(x)$
3. Similarly for the inverse map :$\Im : G \to G :: \Im(x) = x^{-1}$
$\frac{d}{dx}\Im(x)|_{x=a} = \Im'(a) = -R'_{a^{-1}}(1) \circ L'_{a^{-1}}(a) = -L'_{a^{-1}}(1) \circ R'_{a^{-1}}(a)$
and for the map : $f : G \to G :: f(x) = g(x)^{-1} = \Im \circ g(x)$
$\frac{d}{dx}(g(x)^{-1}) = f'(x) = -R'_{g(x)^{-1}}(1) \circ L'_{g(x)^{-1}}(g(x)) \circ g'(x) = -L'_{g(x)^{-1}}(e) \circ R'_{g(x)^{-1}}(g(x)) \circ g'(x)$
4. From the relations above we get the useful identities :
$(L'_g 1)^{-1} = L'_{g^{-1}}(g); (R'_g 1)^{-1} = R'_{g^{-1}}(g)$
$(L'_g h)^{-1} = L'_{g^{-1}}(gh); (R'_g h)^{-1} = R'_{g^{-1}}(hg)$
$L'_g(h) = L'_{gh}(1)L'_{h^{-1}}(h); R'_g(h) = R'_{hg}(1)R'_{h^{-1}}(h)$
$L'_{gh}(1) = L'_g(h)L'_h(1); R'_{hg}(1) = R'_g(h)R'_h(1)$
$(L'_g(h))^{-1} = L'_h(1) L'_{(gh)^{-1}}(gh)$

Group of invertible endomorphisms of a Banach

If E is a Banach vector space, then the set $GL(E; E)$ of continuous automorphisms over E with the composition law is a Lie group. The derivative of the composition law and the inverse are (see derivatives) :
$M : GL(E; E) \times GL(E; E) \to GL(E; E) :: M(f, g) = f \circ g$
$M'(f, g)(\delta f, \delta g) = \delta f \circ g + f \circ \delta g$
$\Im : GL(E; E) \to GL(E; E) :: \Im(f) = f^{-1}$
$(\Im(f))'(\delta f) = -f^{-1} \circ \delta f \circ f^{-1}$

Thus we have :
$$M'(f,g)(\delta f, \delta g) = R'_g(f)\delta f + L'_f(g)\delta g$$
$$R'_g(f)\delta f = \delta f \circ g = R_g(\delta f),$$
$$L'_f(g)\delta g = f \circ \delta g = L_f(\delta g)$$

Linear groups

The set of matrices GL(n,K) over a field K is a Lie group. The operations :
$$M : GL(n,K) \times GL(n,K) \to GL(n,K) :: M(X,Y) = XY = L_X Y = R_Y X$$
$$\Im : GL(n,K) \to GL(n,K) :: \Im(X) = X^{-1}$$
have similarly the derivatives :
$$R'_X(Y) = R_X; L'_X(Y) = L_X$$
$$(\Im(X))'(u) = -X^{-1}uX^{-1}$$

Tangent bundle of a Lie group

The tangent bundle $TG = \cup_{x \in G} T_x G$ of a Lie group is a manifold GxE, on the same field K with dimTG=2xdimG. We can define a multiplication on TG as follows :
$$M : TG \times TG \to TG :: M(U_x, V_y) = R'_y(x)U_x + L'_x(y)V_y \in T_{xy}G$$
$$\Im : TG \to TG :: \Im(V_x) = -R'_{x^{-1}}(1) \circ L'_{x^{-1}}(x)V_x = -L'_{x^{-1}}(x) \circ R'_{x^{-1}}(x)V_x \in T_{x^{-1}}G$$

Identity : $U_x = 0_1 \in T_1 G$

Notice that the operations are between vectors on the tangent bundle TG, and not vector fields (the set of vector fields is denoted $\mathfrak{X}(TG)$).

The operations are well defined and smooth. So TG has a Lie group structure. It is isomorphic to the semi direct group product : $TG \simeq (T_1 G, +) \propto_{Ad} G$ with the map $Ad : G \times T_1 G \to T_1 G$ (Kolar p.98).

22.1.3   Lie algebra of a Lie group

Subalgebras of invariant vector fields

Theorem 1690 *The subspace of the vector fields on a Lie group G over a field K, which are invariant by the left translation have a structure of Lie subalgebra over K with the commutator of vector fields as bracket. Similarly for the vector fields invariant by the right translation*

Proof. i) Left translation, right translation are diffeomorphisms, so the push forward of vector fields is well defined.

A left invariant vector field is such that: $X \in \mathfrak{X}(TG) : \forall g \in G : L_{g*}X = X \Leftrightarrow L'_g(x) X(x) = X(gx)$

so with x=1 : $\forall g \in G : L'_g(1) X(1) = X(g)$

The set of left smooth invariant vector fields is

LVG=$\left\{ X \in \mathfrak{X}_\infty(TG) : X(g) = L'_g(1) u, u \in T_1 G \right\}$

Similarly for the smooth right invariant vector fields :

RVG=$\left\{ X \in \mathfrak{X}_\infty(TG) : X(g) = R'_g(1) u, u \in T_1 G \right\}$

ii) The set $\mathfrak{X}_{\infty}(TG)$ of smooth vector fields over G has an infinite dimensional Lie algebra structure over the field K, with the commutator as bracket (see Manifolds). The push forward of a vector field over G preserves the commutator (see Manifolds).

$X, Y \in LVG : [X, Y] = Z \in \mathfrak{X}_{\infty}(TG)$

$L_{g*}(Z) = L_{g*}([X, Y]) = [L_{g*}X, L_{g*}Y] = [X, Y] = Z \Rightarrow [X, Y] \in LVG$

So the sets LVG of left invariant and RVG right invariant vector fields are both Lie subalgebras of $\mathfrak{X}_{\infty}(TG)$. ■

Lie algebra structure of T_1G

Theorem 1691 *The derivative $L'_g(1)$ of the left translation at x=1 is an isomorphism to the set of left invariant vector fields. The tangent space T_1G becomes a Lie algebra over K with the bracket*

$[u, v]_{T_1G} = L'_{g^{-1}}(g)\left[L'_g(1)u, L'_g(1)v\right]_{LVG}$

So for any two left invariant vector fields X,Y :

$[X, Y](g) = L'_g(1)\left([X(1), Y(1)]_{T_1G}\right)$

Proof. The map : $\lambda : T_1G \to LVG :: \lambda(u)(g) = L'_g(1)u$ is an injective linear map

It has an inverse: $\forall X \in LVG, \exists u \in T_1G : X(g) = L'_g(1)u$

$\lambda^{-1} : LVG \to T_1G :: u = L'_{g^{-1}}(g)X(g)$ which is linear.

So we map : $[]_{T_1G} : T_1G \times T_1G \to T_1G :: [u, v]_{T_1G} = \lambda^{-1}\left([\lambda(u), \lambda(v)]_{LVG}\right)$ is well defined and it is easy to check that it defines a Lie bracket. ■

Remarks :

i) if M is finite dimensional, there are several complete topologies available on $\mathfrak{X}_{\infty}(TG)$ so the continuity of the map λ is well assured. There is no such thing if G is infinite dimensional, however λ and the bracket are well defined algebraically.

ii) some authors (Duistermaat) define the Lie bracket through rigth invariant vector fields.

With this bracket the tangent space T_1G becomes a Lie algebra, the **Lie algebra of the Lie group G** on the field K, with the same dimension as G as manifold, which is Lie isomorphic to the Lie algebra LVG.

Notation 1692 *T_1G is the Lie algebra of the Lie group G*

Right invariant vector fields

The right invariant vector fields define also a Lie algebra structure on T_1G, which is Lie isomorphic to the previous one and the bracket have opposite signs (Kolar p.34).

Theorem 1693 *If X,Y are two smooth right invariant vector fields on the Lie group G, then $[X, Y](g) = -R'_g(1)[X(1), X(1)]_{T_1G}$*

Proof. The derivative of the inverse map (see above) is :

$\frac{d}{dx}\Im(x)|_{x=g} = \Im'(g) = -R'_{g^{-1}}(1) \circ L'_{g^{-1}}(g) \Rightarrow \Im'(g^{-1}) = -R'_g(1) \circ L_g(g^{-1})$

$R'_g(1) = -\Im'(g^{-1}) \circ L'_{g^{-1}}(1)$

So if X is a right invariant vector field : $X(g) = R'_g(1)X(1)$ then $X(g) = -\Im'(g^{-1})X_L$ with $X_L = L'_{g^{-1}}(1)X(1)$ a left invariant vector field.

For two right invariant vector fields :

$[X,Y] = \left[-\Im'(g^{-1})X_L, -\Im'(g^{-1})Y_L\right] = \Im'(g^{-1})[X_L,Y_L]$

$= \Im'(g^{-1})\left[L'_{g^{-1}}(1)X(1), L'_{g^{-1}}(1)X(1)\right]$

$= \Im'(g^{-1})L'_{g^{-1}}(1)[X(1),X(1)]_{T_1G} = -R'_g(1)[X(1),X(1)]_{T_1G}$ ∎

Theorem 1694 *(Kolar p.34) If X,Y are two smooth vector fields on the Lie group G, respectively right invariant and left invariant, then $[X,Y] = 0$*

Lie algebra of the group of automorphisms of a Banach space

Theorem 1695 *For a Banach vector space E, the Lie algebra of $G\mathcal{L}(E;E)$ is $\mathcal{L}(E;E)$. It is a Banach Lie algebra with bracket $[u,v] = u \circ v - v \circ u$*

Proof. If E is a Banach vector space then the set $G\mathcal{L}(E;E)$ of continuous automorphisms over E with the composition law is an open of the Banach vector space $\mathcal{L}(E;E)$. Thus the tangent space at any point is $\mathcal{L}(E;E)$.

A left invariant vector field is : $f \in G\mathcal{L}(E;E), u \in \mathcal{L}(E;E) : X_L(f) = L'_f(1)u = f \circ u = L_f(1)u$

The commutator of two vector fields V,W:$G\mathcal{L}(E;E) \to \mathcal{L}(E;E)$ is (see Differential geometry) :

$[V,W](f) = \left(\frac{d}{df}W\right)(V(f)) - \left(\frac{d}{df}V\right)(W(f))$

with the derivative : $V'(f) : \mathcal{L}(E;E) \to \mathcal{L}(E;E)$ and here : $X_L(f) = f \circ u = R_u(f) \Rightarrow (X_L)'(f) = \frac{d}{df}(R_u(f)) = R_u$

Thus : $\left[L'_f(1)u, L'_f(1)v\right] = R_v(f \circ u) - R_u(f \circ v) = f \circ (u \circ v - v \circ u) = L_f(1)[u,v]_{\mathcal{L}(E;E)} = f \circ [u,v]_{\mathcal{L}(E;E)}$

So the Lie bracket on the Lie algebra $\mathcal{L}$(E;E) is : $u,v \in \mathcal{L}(E;E) : [u,v] = u \circ v - v \circ u$. This is a continuous bilinear map because the composition of maps is itself a continuous operation. ∎

Linear groups

If G is a Lie group of matrices, meaning some subset of GL(K,n) with a Lie group structure, as a manifold it is embedded in the vector space of square matrices K(n) and its tangent space at any point is some vector subspace L(n) of K(n). Left invariant vector fields are of the kind : $X_L(g) = [g] \times [u]$ where $g \in G, u \in L(n)$. The commutator is, as above :

$[X_L, Y_L]_{VG} = [g] \times [u] \times [v] - [g] \times [v] \times [u] \Rightarrow [u,v]_{L(n)} = [u] \times [v] - [v] \times [u].$

The group of automorphisms of the Lie algebra

Following the conditions imposed to the manifold structure of G, the tangent space $T_x G$ at any point can be endowed with the structure of a Banach space (see Differential geometry), which is diffeomorphic to E itself. So this is the case for $T_1 G$ which becomes a Banach Lie algebra, linear diffeomorphic to E.

The set $\mathcal{L}(T_1 G; T_1 G)$ of continuous maps over $T_1 G$ has a Banach algebra structure with composition law and the subset $G\mathcal{L}(T_1 G; T_1 G)$ of continuous automorphisms of $T_1 G$ is a Lie group. Its component of the identity $\text{Int}(T_1 G)$ is also a Lie group. Its Lie algebra is $\mathcal{L}(T_1 G; T_1 G)$ endowed with the bracket : $f, g \in \mathcal{L}(T_1 G; T_1 G) :: [f, g] = f \circ g - g \circ f$

One consequence of these results is :

Theorem 1696 *A Lie group is a parallellizable manifold*

Proof. take any basis $(e_\alpha)_{\alpha \in A}$ of the Lie algebra and transport the basis in any point x by left invariant vector fields : $(L'_x(1)e_\alpha)_{\alpha \in A}$ is a basis of $T_x G$. ∎

22.1.4 Adjoint map

Definition

Definition 1697 *The **adjoint map** over a Lie group G is the derivative of the conjugation taken at x=1*

Notation 1698 *Ad is the adjoint map : $Ad : G \to \mathcal{L}(T_1 G; T_1 G) :: Ad_g = (Conj_g(x))'|_{x=1} = L'_g(g^{-1}) \circ R'_{g^{-1}}(1) = R'_{g^{-1}}(g) \circ L'_g(1)$*

Properties

Theorem 1699 *The **adjoint map** over a Lie group G is a bijective, continuous linear map, and $\forall x \in G : Ad_x \in G\mathcal{L}(T_1 G; T_1 G)$ is a continuous automorphism of Lie algebra and belongs to its connected component $\text{Int}(T_1 G)$.*

Theorem 1700 *(Knapp p.79) For any Lie group Ad is a smooth Lie homomorphism from G to $G\mathcal{L}(T_1 G; T_1 G)$*

It is easy to check that :
$Ad_{xy} = Ad_x \circ Ad_y$
$Ad_1 = Id$
$(Ad_x)^{-1} = Ad_{x^{-1}}$
$\forall u, v \in T_1 G, x \in G : Ad_x[u, v] = [Ad_x u, Ad_x v]$
If G is the set $G\mathcal{L}(E; E)$ of automorphims of a Banach vector space or a subset of a matrices group, then we have seen that the derivatives of the translations are the translations. Thus :
$G\mathcal{L}(E; E) : Ad_x u = x \circ u \circ x^{-1}$
Matrices : $Ad_x[u] = [x][u][x]^{-1}$

Derivative of the adjoint map

Conjugation being differentiable at any order, we can compute the derivative of Ad_x with respect to x :

$\frac{d}{dx}Ad_x|_{x=e} \in \mathcal{L}(T_1G; \mathcal{L}(T_1G; T_1G)) = \mathcal{L}^2(T_1G; T_1G)$

$\left(\frac{d}{dx}Ad_xu\right)|_{x=1}(v) = [u,v]_{T_1G} = ad(u)(v)$

If M is a manifold on the same field K as G, $f : M \to G$ a smooth map then for a fixed $u \in T_1G$ let us define :

$\phi_u : M \to T_1G :: \phi_u(p) = Ad_{f(p)}u$

The value of its derivative for $v_p \in T_pM$ is :

$\phi'_u(p)(v_p) = Ad_{f(p)}\left[L'_{f(p)^{-1}}(f(p))f'(p)v_p, u\right]_{T_1G}$

22.1.5 Exponential map

The exponential map $\exp : \mathcal{L}(E; E) \to G\mathcal{L}(E; E)$ is well defined on the set of continuous linear maps on a Banach space E . It is related to one paramer groups, meaning the differential equation $\frac{dU}{dt} = SU(t)$ between operators on E, where S is the infinitesimal generator and U(t)=exptS. On manifolds the flow of vector fields provides another concept of one parameter group of diffeomorphisms. Lie group structure gives a nice unified view of these concepts.

One parameter subgroup

Theorem 1701 *(Kolar p.36) On a Lie group G, left and right invariant vector fields are the infinitesimal generators of one parameter groups of diffeomorphism. The flow of these vector fields is complete.*

Proof. i) Let $\phi : \mathbb{R} \to G$ be a smooth map such that : $\forall s, t, s+t \in \mathbb{R}$: $\phi(s+t) = \phi(s)\phi(t)$ so $\phi(0) = 1$

Then : $F_R : \mathbb{R} \times G \to G :: F_R(t,x) = \phi(t)x = R_x\phi(t)$ is a one parameter group of diffeomorphism on the manifold G, as defined previously (see Differential geometry). Similarly with $F_L(t,x) = \phi(t)x = L_x\phi(t)$

So F_R, F_L have an infinitesimal generator, which is given by the vector field :

$X_L(x) = L'_x(1)\left(\frac{d\phi}{dt}|_{t=0}\right) = L'_x(1)u$

$X_R(x) = R'_x(1)\left(\frac{d\phi}{dt}|_{t=0}\right) = R'_x(1)u$

with $u = \frac{d\phi}{dt}|_{t=0} \in T_1G$

2. And conversely any left (right) invariant vector field gives rise to the flow $\Phi_{X_L}(t,x) (\Phi_{X_R}(t,x))$ which is defined on some domain $D(\Phi_{X_L}) = \cup_{x \in G}\{J_x \times \{x\}\} \subset \mathbb{R} \times G$ which is an open neighborhood of 0xG

Define : $\phi(t) = \Phi_{X_L}(t,1) \Rightarrow \phi(s+t) = \Phi_{X_L}(s+t,1) = \Phi_{X_L}(s, \Phi_{X_L}(t,1)) = \Phi_{X_L}(s, \phi(t))$ thus ϕ is defined over $\mathbb{R}$

Define $F_L(t,x) = L_x\phi(t)$ then : $\frac{\partial}{\partial t}F_L(t,x)|_{t=0} = L'_x(1)\frac{d\phi}{dt}|_{t=0} = X_L(x)$ so this is the flow of X_L and

$F_L(t+s,x) = F_L(t, F_L(s,x)) = L_{F_L(s,x)}\phi(t) = L_{x\phi(s)}\phi(t) = x\phi(s)\phi(t) \Rightarrow$
$\phi(s+t) = \phi(s)\phi(t)$

Thus the flow of left and rigth invariant vector fields are complete. $\blacksquare$

The exponential map

Definition 1702 *The* **exponential map** *on a Lie group is the map :*
$\exp : T_1G \rightarrow G :: \exp u = \Phi_{X_L}(1,1) = \Phi_{X_R}(1,1)$ *with* $X_L(x) = L'_x(1)u, X_R(x) = R'_x(1)u, u \in T_1G$

From the definition and the properties of the flow :

Theorem 1703 *On a Lie group G over the field K the exponential has the following properties:*
i) $\exp(0) = 1, (\exp u)'|_{u=0} = Id_{T_1G}$
ii) $\exp((s+t)u) = (\exp su)(\exp tu); \exp(-u) = (\exp u)^{-1}$
iii) $\frac{\partial}{\partial t}\exp tu|_{t=\theta} = L'_{\exp \theta u}(1)u = R'_{\exp \theta u}(1)u$
iv) $\forall x \in G, u \in T_1G : \exp(Ad_x u) = x(\exp u)x^{-1} = Conj_x(\exp u)$
v) For any left X_L and right X_R invariant vector fields : $\Phi_{X_L}(x,t) = x\exp tu; \Phi_{X_R}(x,t) = (\exp tu)x$

Theorem 1704 *(Kolar p.36) On a finite dimensional Lie group G the exponential has the following properties:*
i) it is a smooth map from the vector space T_1G to G,
ii) it is a diffeomorphism of a neighbourhood $n(0)$ of 0 in T_1G to a neighborhood of 1 in G. The image $\exp n(0)$ generates the connected component of the identity.

Remark : the theorem still holds for infinite dimensional Lie groups, if the Lie algebra is a Banach algebra, with a continuous bracket (Duistermaat p.35). This is not usually the case, except for the automorphisms of a Banach space.

Theorem 1705 *(Knapp p.91) On a finite dimensional Lie group G for any vectors $u,v \in T_1G$:*
$[u,v]_{T_1G} = 0 \Leftrightarrow \forall s,t \in \mathbb{R} : \exp su \circ \exp tv = \exp tv \circ \exp su$

Warning !
i) we do not have exp(u+v)=(expu)(expv) and the exponential do not commute. See below the formula.
ii) usually exp is not surjective : there can be elements of G which cannot be written as $g = \exp X$. But the subgroup generated (through the operation of the group) by the elements $\{\exp v, v \in n(0)\}$ is the component of the identity G_0. See coordinates of the second kind below.
iii) the derivative of expu with respect to u is *not* expu (see below logarithmic derivatives)
iv) this exponential map is not related to the exponential map deduced from geodesics on a manifold with connection.

Theorem 1706 *Campbell-Baker-Hausdorff formula (Kolar p.40) : In the Lie algebra T_1G of a finite dimensional group G there is a neighborhood of 0 such that $\forall u, v \in n(0) : \exp u \exp v = \exp w$ where*

$$w = u + v + \frac{1}{2}[u,v] + \sum_{n=2}^{\infty} \frac{(-1)^n}{n+1} \int_0^1 \left(\sum_{k,l \geq 0; k+l \geq 1} \frac{t^k}{k!l!} (adu)^k (adv)^l \right)^n (u)\, dt$$

Group of automorphisms of a Banach space

Theorem 1707 *The exponential map on the group of continuous automorphisms $G\mathcal{L}(E;E)$ of Banach vector space E is the map :*

$\exp : \mathbb{R} \times \mathcal{L}(E;E) \to G\mathcal{L}(E;E) :: \exp tu = \sum_{n=0}^{\infty} \frac{t^n}{n!} u^n$ *and* $\|\exp tu\| \leq \exp t \|u\|$

where the power is understood as the n iterate of u.

Proof. $G\mathcal{L}(E;E)$ is a Lie group, with Banach Lie algebra $\mathcal{L}(E;E)$ and bracket : $u, v \in \mathcal{L}(E;E) :: [u,v] = u \circ v - v \circ u$

For $u \in \mathcal{L}(E;E)$, $X_L = L'_f(1)u = f \circ u$ fixed, the map : $\phi : \mathbb{R} \times \mathcal{L}(E;E) \to \mathcal{L}(E;E) :: \phi(t) = f^{-1}\Phi_{X_L}(f,t) = \exp tu$ with the relations : $\phi(0) = Id, \phi(s+t) = \phi(s) \circ \phi(t)$ is a one parameter group over $\mathcal{L}(E;E)$ (see Banach spaces). It is uniformly continuous : $\lim_{t \to 0} \|\phi(t) - Id\| = \lim_{t \to 0} \|\exp tu - Id\| = 0$ because exp is smooth ($\mathcal{L}(E;E)$ is a Banach algebra). So there is an infinitesimal generator : $S \in \mathcal{L}(E;E) : \phi(t) = \exp tS$ with the exponential defined as the series : $\exp tS = \sum_{n=0}^{\infty} \frac{t^n}{n!} S^n$. Thus we can identified the two exponential maps. The exponential map has all the properties seen in Banach spaces :$\|\exp tu\| \leq \exp t \|u\|$ and if E is finite dimensional :$\det(\exp u) = \exp(Tr(u))$ ∎

Linear Group of matrices

If G is a Lie group of matrices, meaning some subset of GL(K,n) with a Lie algebra L(n) then the flow of a left invariant vector field is given by the equation : $\frac{d}{dt}\Phi_{X_L}(t,g)|_{t=\theta} = X_L(\Phi_{X_L}(\theta,g)) = \Phi_{X_L}(\theta,g) \times u$ whose solution is : $\Phi_{X_L}(t,g) = g \exp tu$ where the exponential is computed as $\exp tu = \sum_{p=0}^{\infty} \frac{t^p}{p!}[u]^p$ .Thus the exponential can be computed as the exponential of a matrix.

Logarithmic derivatives

Definition 1708 *(Kolar p.38) : For a map $f \in C_{\infty}(M;G)$ from a manifold M to a Lie group G, on the same field,*

*the **right logarithmic derivative** of f is the map : $\delta_R f : TM \to T_1G :: \delta_R f(u_p) = R'_{f(p)^{-1}}(f(p))f'(p)u_p$*

*the **left logarithmic derivative** of f is the map : $\delta_L f : TM \to T_1G :: \delta_L f(u_p) = L'_{f(p)^{-1}}(f(p))f'(p)u_p$*

$\delta_R, \delta_L \in \Lambda_1(M;T_1G)$: they are 1-form on M valued in the Lie algebra of G

If f=Id$_G$ then $\delta_L\left(f\right)\left(x\right) = L'_x\left(x^{-1}\right) \in \Lambda_1\left(G;T_1G\right)$ is the **Maurer-Cartan form** of G

If f,g$\in C_\infty\left(M;G\right)$:

$\delta_R\left(fg\right)\left(p\right) = \delta_R f\left(p\right) + Ad_{f(p)}\delta_R g\left(p\right)$

$\delta_L\left(fg\right)\left(p\right) = \delta_L g\left(p\right) + Ad_{g(p)^{-1}}\delta_L f\left(p\right)$

Theorem 1709 *(Kolar p.39,Duistermaat p.23) If T_1G is a Banach Lie algebra of a Lie group G, then*

i) The derivative of the exponential is given by :

$\left(\exp u\right)' = R'_{\exp u}(1)\circ\int_0^1 e^{sad(u)}ds = L'_{\exp u}(1)\circ\int_0^1 e^{-sad(u)}ds \in \mathcal{L}\left(T_1G;T_{\exp u}G\right)$

with :

$\int_0^1 e^{sad(u)}ds = \left(ad(u)\right)^{-1}\circ\left(e^{ad(u)}-I\right) = \sum_{n=0}^\infty \dfrac{\left(ad(u)\right)^n}{(n+1)!} \in \mathcal{L}\left(T_1G;T_1G\right)$

$\int_0^1 e^{-sad(u)}ds = \left(ad(u)\right)^{-1}\circ\left(I-e^{-ad(z)}\right) = \sum_{n=0}^\infty \dfrac{\left(-ad(u)\right)^n}{(n+1)!} \in \mathcal{L}\left(T_1G;T_1G\right)$

the series, where the power is understood as the n iterate of ad(u), being convergent if ad(u) is inversible

ii) thus we have :

$\delta_R\left(\exp\right)\left(u\right) = \int_0^1 e^{sad(u)}ds$

$\delta_L\left(\exp\right)\left(u\right) = \int_0^1 e^{-sad(u)}ds$

iii) The eigen values of $\int_0^1 e^{sad(u)}ds$ are $\frac{e^z-1}{z}$ where z is eigen value of ad(u)

iv) The map $\frac{d}{dv}\left(\exp v\right)|_{v=u} : T_1G \to T_1G$ is bijective except for the u which are eigen vectors of ad(u) with eigenvalue of the form $\pm i2k\pi$ with $k \in \mathbb{Z}/0$

Coordinates of the second kind

Definition 1710 *On a n dimensional Lie group on a field K, there is a neighborhood n(0) of 0 in K^n such that the map to the connected component of the identity $G_0 : \phi : n\left(0\right) \to G_0 :: \phi\left(t_1,..t_n\right) = \exp t_1e_1 \times \exp t_2e_2... \times \exp t_ne_n$ is a diffeomorphism. The map ϕ^{-1} is a **coordinate system of the second kind** on G.*

Warning ! The product is not commutative.

22.1.6 Morphisms

Definitions

1. As usual when we have two different structures over a set, morphisms are map which are consistent with both structures.

Definition 1711 *A **group morphism** is a map f between two groups G,H such that : $\forall x,y \in G : f\left(xy\right) = f\left(x\right)f\left(y\right), f\left(x^{-1}\right) = f\left(x\right)^{-1}$*

Definition 1712 *A morphism between topological groups is a group morphism which is also continuous*

Definition 1713 *A class s **Lie group morphism** between class r Lie groups over the same field K is a group morphism which is a class s differentiable map between the manifolds underlying the groups.*

If not precised otherwise the Lie groups and the Lie morphisms are assumed to be smooth.

A morphism is usually called also a homomorphism.

2. Thus the categories of :

i) topological groups, comprises topological groups and continuous morphisms

ii) Lie groups comprises Lie groups on the same field K as objects, and smooth Lie groups morphisms as morphisms.

The set of continuous (resp. Lie) groups morphisms between topological (resp.Lie) groups G,H is denoted $\hom(G; H)$.

3. If a continuous (resp.Lie) group morphism is bijective and its inverse is also a continuous (resp.Lie) group morphism then it is a continuous (resp.Lie) group **isomorphism**. An isomorphism over the same set is an automorphism.

If there is a continuous (resp.Lie) group isomorphism between two topological (resp.Lie) groups they are said to be isomorphic.

Lie group morphisms

1. The most important theorems are the following :

Theorem 1714 *(Kolar p.36) If f is a smooth Lie group morphism $f \in \hom(G, H)$ then its derivative at the unity f'(1) is a Lie algebra morphism $f'(1) \in \hom(T_1 G, T_1 H)$.*

The following diagram commutes :

$$
\begin{array}{ccccc}
T_1 G & \to & \text{f'(1)} & \to & T_1 H \\
\downarrow & & & & \downarrow \\
\exp_G & & & & \exp_H \\
\downarrow & & & & \downarrow \\
G & \to & f & \to & H
\end{array}
$$

$\forall u \in T_1 G : f(\exp_G u) = \exp_H f'(1)u$

and conversely:

Theorem 1715 *(Kolar p.42) If $f : T_1 G \to T_1 H$ is Lie algebra morphism between the Lie algebras of the finite dimensional Lie groups G,H, there is a Lie group morphism F locally defined in a neighborhood of 1_G such that $F'(1_G) = f$. If G is simply connected then there is a globally defined morphism of Lie group with this property.*

Theorem 1716 *(Knapp p.90) Any two simply connected Lie groups whose Lie algebras are Lie isomorphic are Lie isomorphic.*

Notice that in the converse there is a condition : G must be simply connected.

Warning ! two Lie groups with isomorphic Lie algebras are not Lie isomorphic in general, so even if they have the same universal cover they are not necessarily Lie isomorphic.

2. A continuous group morphism between Lie groups is smooth:

Theorem 1717 *(Kolar p.37, Duistermaat p.49, 58) A continuous group morphism f_1 between the Lie groups G,H on the same field K:*

i) f_1 is a smooth Lie group morphism

ii) if f_1 is bijective and H has only many countably connected components, then it is a smooth diffeomorphism and a Lie group isomorphism.

iii) if : at least G or H has finitely many connected components, $f_1 \in \hom(G; H), f_2 \in \hom(H, G)$ are continuous injective group morphisms. Then $f_1(G) = H, f_2(H) = G$ and f_1, f_2 are Lie group isomorphisms.

3. Exponential of ad:

Theorem 1718 *On a Lie group G the map Ad is the exponential of the map ad in the following meaning :*

$$\forall u \in T_1 G : Ad_{\exp_G u} = \exp_{G\mathcal{L}(T_1 G; T_1 G)} ad(u)$$

Proof. $\forall x \in G, Ad_x \in G\mathcal{L}(T_1 G; T_1 G)$: so Ad is a Lie a Lie group morphism : $Ad : G \to G\mathcal{L}(T_1 G; T_1 G)$ and we have :

$\forall u \in T_1 G : Ad_{\exp_G u} = \exp_{G\mathcal{L}(T_1 G; T_1 G)} (Ad_x)'_{x=1} u = \exp_{G\mathcal{L}(T_1 G; T_1 G)} ad(u)$ ∎

The exponential over the Lie group $G\mathcal{L}(T_1 G; T_1 G)$ is computed as for any group of automorphisms over a Banach vector space :

$\exp_{G\mathcal{L}(T_1 G; T_1 G)} ad(u) = \sum_{n=0}^{\infty} \frac{1}{n!} (ad(u))^n$ where the power is understood as the n iterate of ad(u).

And we have :

$\det(\exp ad(u)) = \exp(Tr(ad(u))) = \det Ad_{\exp u}$

22.1.7 Action of a group on a set

Definitions

These definitions are mainly an adaptation of those given in Algebra (groups).

Definition 1719 *Let G be a topological group, E a topological space.*

*A **left-action** of G on E is a continuous map : $\lambda : G \times E \to E$ such that :*
$\forall x \in E, \forall g, g' \in G : \lambda(g, \lambda(g', p)) = \lambda(g \cdot g', p); \lambda(1, p) = p$

*A **right-action** of G on E is a continuous map : $\rho : E \times G \to E$ such that :*
$\forall x \in E, \forall g, g' \in G : \rho(\rho(p, g'), g)) = \rho(p, g' \cdot g); \rho(p, 1) = p$

For g fixed, the maps $\lambda(g, .) : E \to E, \rho(., g) : E \to E$ are bijective.

In the following every definition holds for a right action.

If G is a Lie group and E is a manifold M on the same field then we can define class r actions. It is assumed to be smooth if not specified otherwise.

A manifold endowed with a right or left action is called a G-space.

The orbit of the action through p$\in E$ is the subset of E : $Gp = \{\lambda(g,p), g \in G\}$. The relation $q \in Gp$ is an equivalence relation between p,q denoted R_λ, the classes of equivalence form a partition of G called the orbits of the action.

The action is said to be :

transitive if $\forall p, q \in E, \exists g \in G : q = \lambda(g,p)$. : there is only one orbit.

free if : $\lambda(g,p) = p \Rightarrow g = 1$ (resp. $\rho(p,g) = p \Rightarrow g = 1$). Then each orbit is in bijective correspondance with G and the map : $\lambda(.,p) : G \to \lambda(G,p)$ is bijective.

effective if : $\forall p : \lambda(g,p) = \lambda(h,p) => g = h$ (resp. $\rho(p,g) = \rho(p,h) \Rightarrow g = h$)

Theorem 1720 *(Kolar p.44) If $\lambda : G \times M \to M$ is a continuous effective left action from a locally compact topological group G on a smooth manifold M, then G is a Lie group and the action is smooth*

A subset F of E is **invariant** by the action if : $\forall p \in F, \forall g \in G : \lambda(g,p) \in F$. F is invariant iff it is the union of a collection of orbits. The minimal non empty invariant sets are the orbits.

Theorem 1721 *(Duistermaat p.94) If $\lambda : G \times M \to M$ is a left action from a Lie group G on a smooth manifold M, then for any $p \in E$ the set $A(p) = \{g \in G : \lambda(g,p) = p\}$ is a closed Lie subgroup of G called the **isotropy subgroup** of p. The map $\lambda(.,p) : G \to M$ factors over the projection : $\pi : G \to G/A(p)$ to an injective immersion : $\imath : G/G/A(p) \to M$ which is G equivariant : $\lambda(g, \imath([x])) = \imath([\lambda(g,p)])$. The image of $\imath$ is the orbit through p.*

Proper actions

Definition 1722 *A left action $\lambda : G \times M \to M$ of a Lie group G on a manifold M is **proper** if the preimage of a compact of M is a compact of GxM*

If G and M are compact and Hausdorff, and λ continuous then it is proper (see topology)

Theorem 1723 *(Duistermaat p.98) A left action $\lambda : G \times M \to M$ of a Lie group G on a manifold M is proper if for any convergent sequences $p_n \to p, g_m \to g$ there is a subsequence (g_m, p_n) such that $\lambda(g_m, p_n) \to \lambda(g,p)$*

Theorem 1724 *(Duistermaat p.53) If the left action $\lambda : G \times M \to M$ of a Lie group G on a manifold M is proper and continuous then the quotient set M/R_λ whose elements are the orbits of the action, is Hausdorff with the quotient topology.*

If moreover M,G are finite dimensional and of class r, λ is free and of class r, then the quotient set M/R_λ has a unique structure of class r real manifold of dimension = dim M - dim G. M has a principal fiber bundle structure with group G.

That means the following :

The projection $\pi : M \to M/R_\lambda$ is a class r map;

$\forall p \in M/R_\lambda$ there is a neighborhood n(p) and a diffeomorphism $\tau : \pi^{-1}(n(p)) \to G \times n(p) :: \tau(m) = (\tau_1(m), \tau_2(m))$ such that $\forall g \in G, m \in \pi^{-1}(n(p)) : \tau(\lambda(g,p)) = (\lambda(g, \tau_1(m)), \pi(m))$

Identities

From the definition of an action of a group over a manifold one can deduce some identities which are useful.

1. As a consequence of the definition :
$\lambda\left(g^{-1}, p\right) = \lambda(g, p)^{-1} ; \rho\left(p, g^{-1}\right) = \rho(p, g)^{-1}$

2. By taking the derivative of $\lambda(h, \lambda(g,p)) = \lambda(hg, p)$ and putting successively $g = 1, h = 1, h = g^{-1}$

$\lambda'_p(1, p) = Id_{TM}$

$\lambda'_g(g, p) = \lambda'_g(1, \lambda(g, p)) R'_{g^{-1}}(g) = \lambda'_p(g, p) \lambda'_g(1, p) L'_{g^{-1}}(g)$

$\left(\lambda'_p(g, p)\right)^{-1} = \lambda'_p(g^{-1}, \lambda(g, p))$

Notice that $\lambda'_g(1, p) \in \mathcal{L}(T_1 G; T_p M)$ is not necessarily inversible.

3. Similarly :

$\rho'_p(p, 1) = Id_{TM}$

$\rho'_g(p, g) = \rho'_g(\rho(p, g), 1) L'_{g^{-1}}(g) = \rho'_p(p, g) \rho'_g(p, 1) R'_{g^{-1}}(g)$

$\left(\rho'_p(p, g)\right)^{-1} = \rho'_p(\rho(p, g), g^{-1})$

Fundamental vector fields

They are used in Principal bundles.

Definition 1725 *For a differentiable left action* $\lambda : G \times M \to M$ *of a Lie group G on a manifold M, the* **fundamental vector fields** *are the vectors fields on M generated by a vector of the Lie algebra of G:*
$\zeta_L : T_1 G \to TM :: \zeta_L(u)(p) = \lambda'_g(1, p) u$

We have similarly for a right action :
$\zeta_R : T_1 G \to TM :: \zeta_R(u)(p) = \rho'_g(p, 1) u$

Theorem 1726 *(Kolar p.46) For a differentiable action of a Lie group G on a manifold M, the* **fundamental vector fields** *have the following properties :*

i) the maps ζ_L, ζ_R *are linear*

ii) $[\zeta_L(u), \zeta_L(v)]_{\mathfrak{X}(TM)} = -\zeta_L\left([u, v]_{T_1 G}\right)$

$[\zeta_R(u), \zeta_R(v)]_{\mathfrak{X}(TM)} = \zeta_R\left([u, v]_{T_1 G}\right)$

iii) $\lambda'_p(x, q)|_{p=q} \zeta_L(u)(q) = \zeta_L(Ad_x u)(\lambda(x, q))$

$\rho'_p(q, x)|_{p=q} \zeta_R(u)(q) = \zeta_R(Ad_{x^{-1}} u)(\rho(q, x))$

iv) $\zeta_L(u) = \lambda_*(X_R(u), 0), \zeta_R(u) = \rho_*(X_L(u), 0)$ *with* $X_R(u) = R'_x(1) u, X_L(u) = L'_x(1) u$

v) the fundamental vector fields span an integrable distribution over M, whose leaves are the connected components of the orbits.

Theorem 1727 *The flow of the fundamental vector fields is :*
$$\Phi_{\zeta_L(u)}(t,p) = \lambda(\exp tu, p)$$
$$\Phi_{\zeta_L(u)}(t,p) = \rho(p, \exp tu)$$

Proof. use the relation : $f \circ \Phi_V = \Phi_{f_*V} \circ f$ with $\lambda\left(\Phi_{X_R(u)}(t,x),p\right) = \Phi_{\zeta_L(u)}(t,\lambda(x,p))$ and x=1 ∎

Equivariant mapping

Definition 1728 *A map $f : M \to N$ between the manifolds M,N is **equivariant** by the left actions of a Lie group G λ_1 on M, λ_2 on N, if : $\forall p \in M, \forall g \in G : f(\lambda_1(g,p)) = \lambda_2(g, f(p))$*

Theorem 1729 *(Kolar p.47) If G is connected then f is equivariant iff the fundamental vector fields ζ_{L1}, ζ_{L2} are f related :*
$$f'(p)(\zeta_{L1}(u)) = \zeta_{L2}(u)(f(p)) \Leftrightarrow f_*\zeta_{L1}(u) = \zeta_{L2}(u)$$

A special case is of bilinear symmetric maps, which are invariant under the action of a map. This includes the isometries.

Theorem 1730 *(Duistermaat p.105) If there is a class r>0 proper action of a finite dimensional Lie group G on a smooth finite dimensional Riemannian manifold M, then M has a G-invariant class r-1 Riemannian structure.*

Conversely if M is a smooth finite dimensional riemannian manifold (M,g) with finitely many connected components, and if g is a class k>1 map, then the group of isometries of M is equal to the group of automorphisms of (M;g), it is a finite dimensional Lie group, with finitely many connected components. Its action is proper and of class k+1.

22.2   Structure of Lie groups

22.2.1   Subgroups

The definition of a Lie subgroup requires more than the algebraic definition. In this subsection there are many theorems, which can be useful. Some of them are really necessary for the representation theory, but their utility is not obvious until we get there.

Topological groups

Definition 1731 *A subset H of a topological group G is a subgroup of G if:*
i) H is an algebraic subgroup of G
ii) H has itself the structure of a topologic group
iii) the injection map : $\imath : H \to G$ is continuous.

Explanation : Let Ω be the set of open subsets of G. Then H inherits the relative topology given by $\Omega \cap H$. But an open in H is not necessarily open in G. So we take another Ω_H and the continuity of the map : $\imath : H \to G$ is checked with respect to $(G, \Omega), (H, \Omega_H)$.

Theorem 1732 *If H is an algebraic subgroup of G and is a closed subset of a topological group G then it is a topological subgroup of G.*

But a topological subgroup of G is not necessarily closed.

Theorem 1733 *(Knapp p.84) For a topological group G, with a separable, metric topology :*
i) any open subgroup H is closed and G/H has the discrete topology
ii) the identity component G_0 is open if G is locally connected
iii) any discrete subgroup (meaning whose relative topology is the discrete topology) is closed
iv) if G is connected then any discrete normal subgroup lies in the center of G.

Lie groups

Definition 1734 *A subset H of a Lie group is a **Lie subgroup** of G if :*
i) H is an algebraic subgroup of G
ii) H is itself a Lie group
iii) the inclusion $\imath : H \to G$ is smooth. Then it is an immersion and a smooth morphism of Lie group $\imath \in \hom(H; G)$.

Notice that one can endows any algebraic subgroup with a Lie group structure, but it can be non separable (Kolar p.43), thus the restriction of iii).
The most useful theorem is the following (the demonstration is still valid for G infinite dimensional) :

Theorem 1735 *(Kolar p.42) An algebraic subgroup H of a lie group G which is topologicaly closed in G is a Lie subgroup of G.*

But the converse is not true : a Lie subgroup is not necessarily closed.
As a corollary :

Theorem 1736 *If G is a closed subgroup of matrices in GL(K,n), then it is a Lie subgroup (and a Lie group).*

For instance if M is some Lie group of matrices in GL(K,n), the subset of M such that detg=1 is closed, thus it is a Lie subgroup of M.

Theorem 1737 *(Kolar p.41) If H is a Lie subgroup of the Lie group G, then the Lie algebra T_1H is a Lie subalgebra of T_1G.*

Conversely :

Theorem 1738 *If h is a Lie subalgebra of the Lie algebra of the finite dimensional Lie group G there is a unique connected Lie subgroup H of G which has h as Lie algebra. H is generated by exp(h) (that is the product of elements of exp(h)).*

(Duistermaat p.42) The theorem is still true if G is infinite dimensional and h is a closed linear subspace of $T_1 G$.

Theorem 1739 *Yamabe (Kolar p.43) An arc wise connected algebraic subgroup of a Lie group is a connected Lie subgroup*

22.2.2 Centralizer

Reminder of algebra (see Groups):

The centralizer of a subset A of a group G is the set of elements of G which commute with the elements of A

The center of a group G is the subset of the elements which commute with all other elements.

The center of a topological group is a topological subgroup.

Theorem 1740 *(Kolar p.44) For a Lie group G and any subset A of G:*

i) the centralizer Z_A of A is a subgroup of G.

ii) If G is connected then the Lie algebra of Z_A is the subset : $T_1 Z_A = \{u \in T_1 G : \forall a \in Z_A : Ad_a u = u\}$

If A and G are connected then $T_1 Z_A = \{u \in T_1 G : \forall v \in T_1 Z_A : [u, v] = 0\}$

iii) the center Z_G of G is a Lie subgroup of G and its algebra is the center of $T_1 G$.

Theorem 1741 *(Knapp p.90) A connected Lie subgroup H of a connected Lie group G is contained in the center of G iff $T_1 H$ is contained in the center of $T_1 G$.*

22.2.3 Quotient spaces

Reminder of Algebra (Groups) If H is a subgroup of the group G :

The quotient set G/H is the set $G/ \sim$ of classes of equivalence : $x \sim y \Leftrightarrow \exists h \in H : x = y \cdot h$

The quotient set $H \backslash G$ is the set $G/ \sim$ of classes of equivalence : $x \sim y \Leftrightarrow \exists h \in H : x = h \cdot y$

Usually they are not groups.

The projections give the classes of equivalences denoted $[x]$:

$\pi_L : G \rightarrow G/H : \pi_L(x) = [x]_L = \{y \in G : \exists h \in H : x = y \cdot h\} = x \cdot H$

$\pi_R : G \rightarrow H \backslash G : \pi_R(x) = [x]_R = \{y \in G : \exists h \in H : x = h \cdot y\} = H \cdot x$

$x \in H \Rightarrow \pi_L(x) = \pi_R(x) = [x] = 1$

By choosing one element in each class, we have two maps :

For $G/H : \lambda : G/H \rightarrow G : x \neq y \Leftrightarrow \lambda(x) \neq \lambda(y)$

For $H \backslash G : \rho : H \backslash G \rightarrow G : x \neq y \Leftrightarrow \rho(x) \neq \rho(y)$

any $x \in G$ can be written as : $x = \lambda(x) \cdot h$ or $x = h' \cdot \rho(x)$ for unique h,h'$\in H$

G/H=H\G iff H is a normal subgroup. If so then G/H=H\G is a group. Then π_L is a morphism with kernel H.

Topological groups

Theorem 1742 *(Knapp p.83) If H is a closed subgroup of the separable, metrisable, topological group G, then :*

i) the projections : π_L, π_R are open maps

ii) G/H is a separable metrisable space

iii) if H and G/H (or H\G) are connected then G is connected

iv) if H and G/H (or H\G) are compact then G is compact

Lie groups

Theorem 1743 *(Kolar p.45, 88, Duistermaat p.56) If H is a closed Lie subgroup of the Lie group G then :*

i) the maps :

$\lambda : H \times G \to G :: \lambda(h, g) = L_h g = hg$

$\rho : G \times H \to G : \rho(g, h) = R_h g = gh$

are left (rigth) actions of H on G, which are smooth, proper and free.

*ii) There is a unique smooth manifold structure on G/H,H\G, called **homogeneous space**s of G.*

If G is finite dimensional then dimG/H=dimG - dimH.

iii) The projections π_L, π_R are submersions, so they are open maps and $\pi'_L(g), \pi'_L(g)$ are surjective

iv) G is a principal fiber bundle $G(G/H, H, \pi_L), G(H\G, H, \pi_R)$

v) The translation induces a smooth transitive right (left) action of G on H\G (G/H):

$\Lambda : G \times G/H \to G/H :: \Lambda(g, x) = \pi_L(g\lambda(x))$

$P : H\G \times G \to H\G : P(x, g) = \pi_R(\rho(x) g)$

vi) If H is a normal Lie subgroup then G/H=H\G=N is a Lie group (possibly finite) and the projection $G \to N$ is a Lie group morphism with kernel H.

The action is free so each orbit, that is each coset $[x]$, is in bijective correspondance with H

remark : if H is not closed and G/H is provided with a topology so that the projection is continuous then G/H is not Hausdorff.

Theorem 1744 *(Duistermaat p.58) For any Lie group morphism $f \in \hom(G, H)$:*

i) $K = \ker f = \{x \in G : f(x) = 1_H\}$ is a normal Lie subgroup of G with Lie algebra $\ker f'(1)$

ii) if $\pi : G \to G/K$ is the canonical projection, then the unique homomorphism $\phi : G/K \to H$ such that $f = \phi \circ \pi$ is a smooth immersion making $f(G) = \phi(G/K)$ into a Lie subgroup of H with Lie algebra $f\prime(1)T_1G$

iii) with this structure on $f(G)$, G is a principal fiber bundle with base $f(G)$ and group K.

iv) If G has only many countably components, and f is surjective then G is a principal fiber bundle with base H and group K.

Normal subgroups

A subgroup is normal if for all g in G, $gH = Hg \Leftrightarrow \forall x \in G : x \cdot H \cdot x^{-1} \in H$.

1. For a topological group:

Theorem 1745 *(Knapp p.84) The identity component of a topological group is a closed normal subgroup .*

2. For a Lie group :

Theorem 1746 *(Kolar p.44, Duistermaat p.57) A connected Lie subgroup H of a connected Lie group is normal iff its Lie algebra T_1H is an ideal in T_1G. Conversely : If h is an ideal of the Lie algebra of a Lie group G then the group H generated by exp(h) is a connected Lie subgroup of G, normal in the connected component G_0 of the identity and has h as Lie algebra.*

Theorem 1747 *(Duistermaat p.57) For a closed Lie subgroup H of Lie group G, and their connected component of the identity G_0, H_0 the following are equivalent :*

i) H_0 is normal in G_0

ii) $\forall x \in G_0, u \in T_1H : Ad_x u \in T_1H$

iii) T_1H is an ideal in T_1G

If H is normal in G then H_0 is normal in G_0

Theorem 1748 *(Duistermaat p.58) If f is a Lie group morphism between the Lie groups G,H then $K = \ker f = \{x \in G : f(x) = 1_H\}$ is a normal Lie subgroup of G with Lie algebra $\ker f'(1)$*

Theorem 1749 *(Kolar p.44) For any closed subset A of a Lie group G, the normalizer $N_A = \{x \in G : Conj_x(A) = A\}$ is a Lie subgroup of G. If A is a Lie subgroup of G, A and G connected, then $N_A = \{x \in G : \forall u \in T_1N_a : Ad_x u \in T_1N_A\}$ and $T_1N_A = \{u \in T_1G : \forall v \in T_1N_a ad(u)v \in T_1N_A\}$*

22.2.4   Connected component of the identity

Theorem 1750 *The connected component of the identity G_0 in a Lie group G:*

i) is a normal Lie subgoup of G, both closed and open in G. It is the only open connected subgroup of G.

ii) is arcwise connected

iii) is contained in any open algebraic subgroup of G

iv) is generated by $\{\exp u, u \in T_1G\}$

v) G/G_0 is a discrete group

The connected components of G are generated by xG_0 or $G_0 x$: so it suffices to know one element of each of the other connected components to generate G.

22.2.5 Semi-direct product of groups

Theorem 1751 *(Kolar p.47) If $\lambda : G \times K \to K$ is a left action of the Lie group G on the Lie group K, such that for each $g \in G : \lambda(g,.) : K \to K$ is a group morphism, the operation in KxG : $(k,g) \times (k',g') = (k\lambda(g,k'), gg')$ defines a Lie group $\widetilde{G}$,denoted $\widetilde{G} = K \propto_\lambda G$, called the **semi-direct product** of K and G.*

The second projection : $\pi_2 : \widetilde{G} \to G$ is a surjective smooth morphism with kernel Kx1

The insertion : $\imath : G \to \widetilde{G} : \imath(g) = (1,g)$ is a smooth morphism with $\pi_2 \circ \imath = Id_G$

The map : $\Lambda : G \to Aut(K)$ is an automorphism in the group Aut(K) of automorphisms on K

We have similar results with a right action and the multiplication :
$(k,g) \times (k',g') = (gg', \rho(k,g)k')$

Theorem 1752 *(Neeb p.35) If $\lambda : G \times K \to K$ is a left action of the Lie group G on the Lie group K, such that for each $g \in G : \lambda(g,.) : K \to K$ is a group morphism, a map $f : K \to G$ is a **1-cocycle** if $f(kk') = \lambda(f(k), k')$. Then the map : $(f, Id_K) : K \to K \propto_\lambda G$ is a group morphism, and conversely every group morphism is of this form.*

This is the starting point to another cohomology theory.

22.2.6 Third Lie's theorem

A Lie group has a Lie algebra, the third Lie's theorem adresses the converse : given a Lie algebra, can we build a Lie group ?

Theorem 1753 *(Kolar p.42, Duistermaat p.79) Let g be a finite dimensional real Lie algebra, then there is a simply connected Lie group with Lie algebra g. The restriction of the exponential mapping to the center Z of g induces an isomorphism from (Z,+) to the identity component of the center of G (the center Z of g and of G are abelian).*

Notice that that the group is not necessarily a group of matrices : there are finite dimensional Lie groups which are not isomorphic to a matrices group (meanwhile a real finite dimensional Lie algebra is isomorphic to a matrices algebra).

This theorem does not hold if g is infinite dimensional.

Theorem 1754 *Two simply connected Lie groups with isomorphic Lie algebras are Lie isomorphic.*

But this is generally untrue if they are not simply connected. However if we have a simply connected Lie group, we can deduce all the other Lie groups, simply connected or not, sharing the same Lie algebra, as quotient groups. This is the purpose of the next topic.

22.2.7 Covering group

See topology for the definition of covering spaces.

Theorem 1755 *(Knapp p.89) Let G be a connected Lie group, there is a unique connected, simply connected Lie group $\widetilde{G}$ and a smooth Lie group morphism : $\pi : \widetilde{G} \to G$ such that $\left(\widetilde{G}, \pi\right)$ is a universal covering of G. $\widetilde{G}$ and G have the same dimension, and same Lie algebra. G is Lie isomorphic to $\widetilde{G}/H$ where H is some discrete subgroup in the center of G. Any connected Lie group G' with the same Lie algebra as G is isomorphic to $\widetilde{G}/D$ for some discrete subgroup D in the center of G.*

So for any connected Lie group G, there is a unique simply connected Lie group $\widetilde{G}$ which has the same Lie algebra. And $\widetilde{G}$ is the direct product of G and some finite groups. The other theorems give results useful whith topological groups.

Theorem 1756 *(Knapp p.85) Let G be a connected, locally pathwise connected, separable topological metric group, H be a closed locally pathwise connected subgroup, H_0 the identity component of H. Then*
i) the quotient G/H is connected and pathwise connected
ii) if G/H is simply connected then H is connected
iii) the map $G/H_0 \to G/H$ is a covering map
iv) if H is discrete, then the quotient map $G \to G/H$ is a covering map
v) if H is connected ,G simply connected, G/H locally simply connected, then G/H is simply connected

Theorem 1757 *(Knapp p.88) Let G be a locally connected, pathwise connected, locally simply connected, separable topological metric group, $\left(\widetilde{G}, \pi : \widetilde{G} \to G\right)$ a simply connected covering of G with $e = \pi^{-1}(1)$. Then there is a unique multiplication in $\widetilde{G}$ such that it is a topological group and π is a group homeomorphism. $\widetilde{G}$ with this structure is called the **universal covering group** of G. It is unique up to isomorphism.*

Theorem 1758 *(Knapp p.88) Let G be a connected, locally pathwise connected, locally simply connected, separable topological metric group, H a closed subgroup, locally pathwise connected, locally simply connected. If G/H is simply connected then the fundamental group $\pi_1(G, 1)$ is isomorphic to a quotient group of $\pi_1(H, 1)$*

22.2.8 Complex structures

The nature of the field K matters only for Lie groups, where the manifold structure is involved. It does not matter for topological groups. All the previous results are valid for K=$\mathbb{R}, \mathbb{C}$ whenever it is not stated otherwise. So if G is a complex manifold its Lie algebra is a complex algebra and the exponential is a holomorphic map.

The converse (how a real Lie group can be made a complex Lie group) is less obvious, as usual. The group structure is not involved, so the problem is to define a complex manifold structure. The way to do it is through the Lie algebra.

Definition 1759 *A complex Lie group $G_{\mathbb{C}}$ is the* **complexification** *of a real Lie group G if G is a Lie subgroup of $G_{\mathbb{C}}$ and if the Lie algebra of $G_{\mathbb{C}}$ is the complexification of the Lie algebra of G.*

There are two ways to "complexify" G.

1. By a complex structure J on T_1G. Then T_1G and the group G stay the same as set. But there are compatibility conditions (the dimension of G must be even and J compatible with the bracket), moreover the exponential must be holomorphic (Knapp p.96) with this structure : $\frac{d}{dv}\exp J(v)\big|_{v=u} = J((\frac{d}{dv}\exp v)\big|_{v=u})$. We have a partial answer to this problem :

Theorem 1760 *(Knapp p.435) A semi simple real Lie group G whose Lie algebra has a complex structure admits uniquely the structure of a complex Lie group such that the exponential is holomorphic.*

2. By complexification of the Lie algebra. This is always possible, but the sets do not stay the same. The new complex algebra $g_{\mathbb{C}}$ can be the Lie algebra of some complex Lie group $G_{\mathbb{C}}$ with complex dimension equal to the real dimension of G. But the third's Lie theorem does not apply, and more restrictive conditions are imposed to G. If there is a complex Lie group $G_{\mathbb{C}}$ such that : its Lie algebra is $(T_1G)_{\mathbb{C}}$ and G is a subgroup of $G_{\mathbb{C}}$ then one says that $G_{\mathbb{C}}$ is the complexified of G. Complexified of a Lie group do not always exist, and they are usually not unique. Anyway then $G_{\mathbb{C}} \neq G$.

If G is a real semi simple finite dimensional Lie group, its Lie algebra is semi-simple and its complexified is still semi-simple, thus $G_{\mathbb{C}}$ must be a complex semi simple group, isomorphic to a Lie group of matrices, and so for G.

Theorem 1761 *(Knapp p.537) A compact finite dimensional real Lie group admits a unique complexification (up to isomorphism)*

22.2.9 Solvable, nilpotent Lie groups

Theorem 1762 *(Kolar p.130) The commutator of two elements of a group G is the operation : $K : G \times G \to G :: K(g,h) = ghg^{-1}g^{-1}$*

If G is a Lie group the map is continuous. If G_1, G_2 are two closed subgroup, then $K[G_1, G_2]$ generated by all the commutators $K(g_1, g_2)$ with $g_1 \in G_1, g_2 \in G_2$ is a closed subgroup, thus a Lie group.

From there one can build sequences similar to the sequences of brackets of Lie algebra :

$G^0 = G = G_0, G^n = K\left[G^{n-1}, G^{n-1}\right], G_n = K\left[G, G_{n-1}\right], G_n \subset G^n$

A Lie group is said to be solvable if $\exists n \in \mathbb{N} : G^n = 1$

A Lie group is said to be nilpotent if $\exists n \in \mathbb{N} : G_n = 1$

But the usual and most efficient way is to proceed through the Lie algebra.

Theorem 1763 *A Lie group is solvable (resp.nilpotent) if its Lie algebra is solvable (resp.nilpotent).*

Theorem 1764 *(Knapp p.106) If g is a finite dimensional, solvable, real, Lie algebra, then there is a simply connected Lie group G with Lie algebra g, and G if diffeomorphic to an euclidean space with coordinates of the second kind.*

If $(e_i)_{i=1}^n$ is a basis of g, then : $\forall g \in G, \exists t_1, ..t_n \in \mathbb{R} : g = \exp t_1 e_1 \times \exp t_2 e_2 ... \times \exp t_n e_n$

There is a sequence (G_p) of closed simply connected Lie subgroups of G such that :

$G = G_0 \supseteq G_1 ... \supseteq G_n = \{1\}$

$G_p = \mathbb{R}^p \propto G_{p+1}$

G_{p+1} *normal in* G_p

Theorem 1765 *(Knapp p.107) On a simply connected finite dimensional nilpotent real Lie group G the exponential map is a diffeomorphism from $T_1 G$ to G (it is surjective). Moreover any Lie subgroup of G is simply connected and closed.*

22.2.10   Abelian Lie groups

Abelian group = commutative group

Main result

Theorem 1766 *(Duistermaat p.59) A connected Lie group G is abelian iff its Lie algebra is abelian. Then the exponential map is onto and its kernel is a discrete (closed, zero dimensional) subgroup of $(T_1 G, +)$. The exponential induces an isomorphism of Lie groups : $T_1 G / \ker \exp \to G$*

That means that there are $0 < p \leq \dim T_1 G$ linearly independant vectors V_k of $T_1 G$ such that :

$\ker \exp = \sum_{k=1}^p z_k V_k, z_k \in \mathbb{Z}$

Such a subset is called a p dimensional integral lattice.

Any n dimensional abelian Lie group over the field K is isomorphic to the group (with addition) : $(K/\mathbb{Z})^p \times K^{n-p}$ with $p = \dim span \ker \exp$

Definition 1767 *A **torus** is a compact abelian topological group*

Theorem 1768 *Any torus which is a finite n dimensional Lie group on a field K is isomorphic to $((K/\mathbb{Z})^n, +)$*

A subgroup of $(\mathbb{R}, +)$ is of the form $G = \{ka, k \in \mathbb{Z}\}$ or is dense in $\mathbb{R}$

Examples :

the nxn diagonal matrices $\mathrm{diag}(\lambda_1, ... \lambda_n)$, $\lambda_k \neq 0 \in K$ is a commutative n dimensional Lie group isomorphic to K^n.

the nxn diagonal matrices $\mathrm{diag}(\exp(i\lambda_1), ... \exp(i\lambda_n))$, $\lambda_k \neq 0 \in \mathbb{R}$ is a commutative n dimensional Lie group which is a torus.

Pontryagin duality

This is a very useful concept which is used mainly to define the Fourier transform (see Functional analysis). It is defined for topological groups. Some of the concepts are similar to the linear multiplicative functionals of Normed algebras.

Definition 1769 *The "**Pontryagin dual**" $\widehat{G}$ of an abelian topological group G is : the set of continuous morphisms, called **characters**, $\chi : G \to T$ where T is the set of complex number of module 1 endowed with the product as internal operation : $T = (\{z \in \mathbb{C} : |z| = 1\}, \times)$. Endowed with the compact-open topology and the pointwise product as internal operation $\widehat{G}$ is a topological abelian group.*

$\chi \in \widehat{G}, g, h \in G : \chi(g + h) = \chi(g) \chi(h), \chi(-g) = \chi(g)^{-1} \chi(1) = 1$
$(\chi_1 \chi_2)(g) = \chi_1(g) \chi_2(g)$

The "double-dual" of G : $\left(\widehat{\widehat{G}}\right) : \theta : \widehat{G} \to T$

The map : $\tau : G \times \widehat{G} \to T :: \tau(g, \chi) = \chi(g)$ is well defined and depends only on G.

For any $g \in G$ the map : $\tau_g : \widehat{G} \to T :: \tau_g(\chi) = \tau(g, \chi) = \chi(g)$ is continuous and $\tau_g \in \left(\widehat{\widehat{G}}\right)$

The map, called Gel'fand transformation : $\widehat{} : G \to \left(\widehat{\widehat{G}}\right) :: \widehat{g} = \tau_g$ has the defining property : $\forall \chi \in \widehat{G} : \widehat{g}(\chi) = \chi(g)$

Theorem 1770 *Pontryagin-van Kampen theorem: If G is an abelian, locally compact topological group, then G is continuously isomorphic to its bidual $\left(\widehat{\widehat{G}}\right)$ through the Gel'fand transformation. Then if G is compact, its Pontryagin dual $\widehat{G}$ is discrete, and conversely if $\widehat{G}$ is discrete, then G is compact. If G is finite then $\widehat{G}$ is finite.*

If G is compact, then it is isomorphic to a closed subgroup of $T^{\widehat{G}}$.

Examples : $\widehat{\mathbb{Z}} = T, \widehat{T} = \mathbb{Z}, \widehat{\mathbb{R}} = \mathbb{R}, (\widehat{\mathbb{Z}/n\mathbb{Z}}) = \mathbb{Z}/n\mathbb{Z}$

A subset E of $\widehat{G}$ is said to separate G if : $\forall g, h \in G, g \neq h, \exists \chi \in E : \chi(g) \neq \chi(h)$

Any subset E which separates G is dense in $\widehat{G}$.

Theorem 1771 *Peter-Weyl: If G is an abelian, compact topological group, then its topological dual $\widehat{G}$ separates G:*
$\forall g, h \in G, g \neq h, \exists \chi \in \widehat{G} : \chi(g) \neq \chi(h)$

22.2.11 Compact groups

Compact groups are of special interest for physicists. They have many specific properties that we will find again in representation theory.

Main properties

A Lie algebra A is compact if the component of the identity of its group of automorphisms Int(A) is compact with the topology of $\mathcal{L}(A; A)$ (see Lie algebras)

Definition 1772 *A topological or Lie group is compact if it is compact with its topology.*

Theorem 1773 *The Lie algebra of a compact Lie group is compact.*
Any closed algebraic subgroup of a compact Lie group is a compact Lie subgroup. Its Lie algebra is compact.

Theorem 1774 *A compact Lie group is necessarily*
i) finite dimensional
ii) a torus if it is a connected complex Lie group
iii) a torus if it is abelian

Proof. i) a compact manifold is locally compact, thus it cannot be infinite dimensional

ii) the Lie algebra of a compact complex Lie group is a complex compact Lie algebra, thus an abelian algebra, and the Lie group is an abelian Lie group. The only abelian Lie groups are the product of torus and euclidean spaces, so a complex compact Lie group must be a torus. ∎

Theorem 1775 *(Duistermaat p.149) A real finite dimensional Lie group G :*
i) is compact iff the Killing form of its Lie algebra T_1G is negative semi-definite and its kernel is the center of T_1G
ii) is compact, semi-simple, iff the Killing form of its Lie algebra T_1G is negative definite (so it has zero center).

Theorem 1776 *(Knapp p.259) For any connected compact Lie group the exponential map is onto. Thus : $\exp : T_1G \to G$ is a diffeomorphism*

Theorem 1777 *Weyl's theorem (Knapp p.268): If G is a compact semi-simple real Lie group, then its fundamental group is finite, and its universal covering group is compact.*

8. For a compact real Lie group there is a Haar measure ϖ, $\left| \int_G \varpi \right| < \infty$ and an inner product invariant by the adjoint Ad on the Lie algebra, meaning a bilinear symmetric form $\langle \rangle$ such that : $\langle u, v \rangle = \langle Ad_x u, Ad_x v \rangle$

Take any bilinear symmetric form B on the real finite dimensional vector space T_1G and define :

$\langle u, v \rangle = \int_G B\left(Ad_x u, Ad_x v \right) \varpi$

460

Structure of compact real Lie groups

The study of the internal structure of a compact group proceeds along lines similar to the complex simple Lie algebras, the tori replacing the Cartan algebras. It mixes analysis at the algebra and group levels (Knapp IV.5 for more).

Let G be a compact, connected, real Lie group.

1. Torus:

A torus of G is an abelian Lie subgroup. It is said to be maximal if it is not contained in another torus. Maximal tori are conjugate from each others via Ad_g. Each element of G lies in some maximal torus and is conjugate to an element of any maximal torus. The center of G lies in all maximal tori.

So let T be a maximal torus, then : $\forall g \in G : \exists t \in T, x \in G : g = xtx^{-1}$. The relation : $x \sim y \Leftrightarrow \exists z : y = zxz^{-1}$ is an equivalence relation, thus we have a partition of G in classes of conjugacy, T is one class, pick up $(x_i)_{i \in I}$ in the other classes and G=$\left\{ x_i t x_i^{-1}, i \in I, t \in T \right\}$.

2. Root space decomposition:

Let T be a maximal torus, with Lie algebra t. If we take the complexified $(T_1 G)_C$ of the Lie algebra of G, and t_C of t, then t_C is a Cartan subalgebra of $(T_1 G)_C$ and we have a root-space decomposition in a similar fashion as a semi simple complex Lie algebra :

$(T_1 G)_C = t_C \oplus_\alpha g_\alpha$

where the root vectors $g_\alpha = \{ X \in (T_1 G)_C : \forall H \in t_C : [H, X] = \alpha (H) X \}$ are the unidimensional eigen spaces of ad over t_C, with eigen values $\alpha (H)$, which are the roots of $(T_1 G)_C$ with respect to t.

The set of roots $\Delta ((T_1 G)_C, t_C)$ has the properties of a roots system except that we do not have $t_C^* = span \Delta$.

For any $H \in t : \alpha (H) \in i\mathbb{R}$: the roots are purely imaginary.

3. For any $\lambda \in t_c^* : \lambda$ is said to be analytically integral if it meets one of the following properties :

i) $\forall H \in t : \exp H = 1 \Rightarrow \exists k \in \mathbb{Z} : \lambda (H) = 2i\pi k$

ii) there is a continuous homomorphism ξ from T to the complex numbers of modulus 1 (called a multiplicative character) such that : $\forall H \in t : \exp \lambda (H) = \xi (\exp H)$

then λ is real valued on t. All roots have these properties.

remark : $\lambda \in t_c^*$ is said to be algebraically integral if $\frac{2 \langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle} \in \mathbb{Z}$ with some inner product on the Lie algebra as above.

22.2.12 Semi simple Lie groups

Definition 1778 *A Lie group is :*

*   ***simple** if the only normal subgroups are 1 and G itself.*
*   ***semi-simple** if its Lie algebra is semi-simple (it has no non zero solvable ideal).*

The simplest criterium is that the Killing form of a semi-simple Lie group is non degenerate. The center of a connected semi-simple Lie group is just 1.

Any real semi-simple finite dimensional Lie algebra A has a Cartan decomposition that is a pair of subvector spaces l_0, p_0 of A such that : $A = l_0 \oplus p_0$, l_0 is a subalgebra of A, and an involution $\theta(l_0) = l_0, \theta(p_0) = -p_0$

We have something similar at the group level, which is both powerful and useful because semi-simple Lie groups are common.

Theorem 1779 *(Knapp p.362) For any real, finite dimensional, semi-simple Lie group G, with the subgroup L corresponding to $l_0 \in T_1G$:*

i) There is a Lie group automorphism Θ on G such that $\Theta'(g)|_{g=1} = \theta$

ii) L is invariant by Θ

iii) the maps : $L \times p_0 \to G :: g = l \exp p$ and $p_0 \times L \to G :: g = (\exp p)l$ are diffeomorphisms onto.

iv) L is closed

v) L contains the center Z of G

vi) L is compact iff Z is finite

vii) when Z is finite then L is a maximal compact subgroup of G.

So any element of G can be written as : $g = l \exp p$ or equivalently as $g = (\exp X)l$. Moreover if L is compact the exponential is onto : $l = \exp \lambda, \lambda \in l_0$

Θ is called the **global Cartan involution**

The decomposition $g = (\exp p)l$ is the **global Cartan decomposition**.

Warning ! usually the set $\{\exp p, p \in p_0\}$ is not a group.

As an application :

Theorem 1780 *(Knapp p.436) For a complex semi simple finite dimensional Lie group G:*

i) its algebra is complex semi simple, and has a real form u_0 which is a compact semi simple real Lie algebra and the Lie algebra can be written as the real vector space $T_1G = u_0 \oplus iu_0$

ii) G has necessarily a finite center.

iii) G is Lie complex isomorphic to a complex Lie group of matrices. And the same is true for its universal covering group (which has the same algebra).

Remark : while semi simple Lie algebras can be realized as matrices algebras, semi simple *real* Lie groups need not to be realizable as group of matrices : there are examples of such groups which have no linear faithful representation (ex : the universal covering group of SL(2,$\mathbb{R}$)).

22.2.13   Classification of Lie groups

The isomorphisms classes of finite dimensional :

i) simply connected compact semi simple real Lie groups

ii) complex semi simple Lie algebras

iii) compact semi simple real Lie algebras

iv) reduced abstract roots system

v) abstract Cartan matrices and their associated Dynkin diagrams

are in one one correspondance, by passage from a Lie group to its Lie algebra, then to its complexification and eventually to the roots system.

So the list of all simply connected compact semi simple real Lie groups is deduced from the list of Dynkin diagrams given in the Lie algebra section, and we go from the Lie algebra to the Lie group by the exponential.

22.3 Integration on a group

The integral can be defined on any measured set, and so on topological groups, and we start with this case which is the most general. The properties of integration on Lie groups are similar, even if they proceed from a different approach.

22.3.1 Integration on a topological group

Haar Radon measure

The integration on a topological group is based upon Radon measure on a topological group. A Radon measure is a Borel, locally finite, regular,signed measure on a topological Hausdorff locally compact space (see Measure). So if the group is also a Lie group it must be finite dimensional.

Definition 1781 *(Neeb p.46) A left (right) Haar Radon measure on a locally compact topological group G is a positive Radon measure μ such that : $\forall f \in C_{0c}(G;\mathbb{C}), \forall g \in G : \ell(f) = \int_G f(gx)\mu(x) = \int_G f(x)\mu(x)$*

and for right invariant : $\forall f \in C_{0c}(G;\mathbb{C}), \forall g \in G : \ell(f) = \int_G f(xg)\mu(x) = \int_G f(x)\mu(x)$

Theorem 1782 *(Neeb p.46) Any locally compact topological group has Haar Radon measures and they are proportional.*

The Lebesgue measure is a Haar measure on $(\mathbb{R}^m,+)$ so any Haar measure on $(\mathbb{R}^m,+)$ is proportional to the Lebesgue measure.

If G is a discrete group a Haar Radon measure is just a map : $\int_G f\mu = \sum_{g \in G} f(g)\mu(g), \mu(g) \in \mathbb{R}_+$

On the circle group T=$\{\exp it, t \in \mathbb{R}\} : \ell(f) = \frac{1}{2\pi}\int_0^{2\pi} f(\exp it)\,dt$

Theorem 1783 *All connected, locally compact groups G are σ-finite under Haar measure.*

Modular function

Theorem 1784 *For any left Haar Radon measure μ_L on the group G there is a continuous homomorphism, called the **modular function** $\Delta : G \to \mathbb{R}_+$ such that : $\forall a \in G : R_a^* \mu_L = \Delta(a)^{-1}\mu_L$. It does not depend on the choice of μ_L.*

Definition 1785 *If the group G is such that $\Delta(a) = 1$ then G is said to be* **unimodular** *and then any left invariant Haar Radon measure is also right invariant, and called a Haar Radon measure.*

Are unimodular the topological, locally compact, goups which are either : compact, abelian, or for which the commutator group (G,G) is dense.

Remark : usually affine groups are not unimodular.

Spaces of functions

1. If G is a topological space endowed with a Haar Radon measure (or even a left invariant or right invariant measure) μ, then one can implement the classical definitions of spaces of integrable functions on G (Neeb p.32). See the part Functional Analysis for the definitions.

$L^p(G, S, \mu, \mathbb{C})$ is a Banach vector space with the norm: $\|f\|_p = \left(\int_E |f|^p \mu\right)^{1/p}$

$L^2(G, S, \mu, \mathbb{C})$ is a Hilbert vector space with the scalar product : $\langle f, g \rangle = \int_E \overline{f} g \mu$

$\mathcal{L}^\infty(G, S, \mu, \mathbb{C}) = \{f : G \to \mathbb{C} : \exists C \in \mathbb{R} : |f(x)| < C\}$

$L^\infty(G, S, \mu, \mathbb{C})$ is a C*-algebra (with pointwise multiplication and the norm $\|f\|_\infty = \inf(C \in \mathbb{R} : \mu(\{|\mu(f)| > C\}) = 0)$

2. If H is a separable Hilbert space the definition can be extended to maps valued in H (Knapp p.567)..

One use the fact that :

if $(e_i)_{\in I}$ is a Hilbert basis, then for any measurable maps $G \to H$.

$(\varphi(g), \psi(g)) = \sum_{i \in I} \langle \varphi(g) e_i, e_i \rangle \langle e_i, \psi(g) e_i \rangle$ is a measurable map : $G \to \mathbb{C}$

The scalar product is defined as : $(\varphi, \psi) = \int_G (\varphi(g), \psi(g)) \mu$.

Then we can define the spaces $L^p(G, \mu, H)$ as above and $L^2(G, \mu, H)$ is a separable Hilbert space

Convolution

Definition 1786 *(Neeb p.134) Let μ_L be a left Haar measure on the locally compact topological group G. The* **convolution** *on $L^1(G, S, \mu_L, \mathbb{C})$ is defined as the map :*

$*: L^1(G, S, \mu_L, \mathbb{C}) \times L^1(G, S, \mu_L, \mathbb{C})$

$\to L^1(G, S, \mu_L, \mathbb{C}) :: \varphi * \psi(g) = \int_G \varphi(x) \psi(x^{-1}g) \mu_L(x) = \int_G \varphi(gx) \psi(x^{-1}) \mu(x)$

Definition 1787 *(Neeb p.134) Let μ_L be a left Haar measure on the locally compact topological group G. On the space $L^1(G, S, \mu_L, \mathbb{C})$ the involution is defined as : $\varphi^*(g) = (\Delta(g))^{-1} \overline{\varphi(g^{-1})}$ so it will be $\varphi^*(g) = \overline{\varphi(g^{-1})}$ if G is unimodular.*

$\text{Supp}(\varphi * \psi) \subset Supp(\varphi) Supp(\psi)$

convolution is associative

$\|\varphi * \psi\|_1 \le \|\varphi\|_1 \|\psi\|_1$

$\|\varphi^*\|_1 = \|\varphi\|_1$

$(\varphi * \psi)^* = \psi^* * \varphi^*$

If G is abelian the convolution is commutative (Neeb p.161)

With the left and right actions of G on $L^1(G, S, \mu_L, \mathbb{C})$:

$\Lambda(g)\varphi(x) = \varphi(g^{-1}x)$

$P(g)\varphi(x) = \varphi(xg)$

then :

$\Lambda(g)(\varphi * \psi) = (\Lambda(g)\varphi) * \psi$

$P(g)(\varphi * \psi) = \varphi * (P(g)\psi)$

$(P(g)\varphi) * \psi = \varphi * (\Delta(g))^{-1}\Lambda(g^{-1})\psi$

$(\Lambda(g)\varphi)^* = \Delta(g)(P(g)\varphi^*)$

$(P(g)\varphi)^* = (\Delta(g))^{-1}(\Lambda(g)\varphi^*)$

$\|\Lambda(g)\varphi\|_1 = \|\varphi\|_1$

$\|P(g)\varphi\|_1 = (\Delta(g))^{-1}\|\varphi\|_1$

Theorem 1788 *With convolution as internal operation $L^1(G, S, \mu_L, \mathbb{C})$ is a complex Banach *-algebra and $\Lambda(g), \Delta(g)P(g)$ are isometries.*

22.3.2 Integration on a Lie group

The definition of a Haar measure on a Lie group proceeds differently since it is the integral of a n-form.

Haar measure

Definition 1789 *A **left Haar measure** on a real n dimensional Lie group is a n-form ϖ on TG such that : $\forall a \in G : L_a^* \varpi = \varpi$*

*A **right Haar measure** on a real n dimensional Lie group is a n-form ϖ on TG such that : $\forall a \in G : R_a^* \varpi = \varpi$*

that is :$\forall u_1, .. u_r \in T_x G : \varpi(ax)(L_a'(x)u_1, ..., L_a'(x)u_r) = \varpi(x)(u_1, ...u_r)$

Theorem 1790 *Any real finite dimensional Lie group has left and right Haar measures, which are volume forms on TG*

Proof. Take the dual basis $(e^i)_{i=1}^n$ of $T_1 G$ and its pull back in x :

$e^i(x)(u_x) = e^i(L_{x^{-1}}'(x)) \Rightarrow e^i(x)(e_j(x)) = e^i(L_{x^{-1}}'(x)e_j(x)) = \delta_j^i$

Then $\varpi_r(x) = e^1(x) \wedge ... \wedge e^r(x)$ is left invariant :

$\varpi(ax)(L_a'(x)u_1, ..., L_a'(x)u_r) = \sum_{(i_1...i_r)} \epsilon(i_1, ..i_r)e^{i_1}(ax)(L_a'(x)u_1)...e^{i_r}(ax)(L_a'(x)u_r)$

$= \sum_{(i_1...i_r)} \epsilon(i_1, ..i_r)e^{i_1}\left(L_{(ax)^{-1}}'(ax)L_a'(x)u_1\right)...e^{i_r}\left(L_{(ax)^{-1}}'(ax)L_a'(x)u_r\right)$

$L_{(ax)^{-1}}'(ax) = L_{x^{-1}}'(x)(L_a'(x))^{-1} \Rightarrow e^{i_1}\left(L_{(ax)^{-1}}'(ax)L_a'(x)u_1\right) = e^{i_1}\left(L_{x^{-1}}'(x)u_1\right) = e^{i_1}(x)(u_1)$

Such a n form is never null, so $\varpi_n(x) = e^1(x) \wedge ... \wedge e^n(x)$ defines a left invariant volume form on G. And G is orientable. ∎

All left (rigth) Haar measures are proportionnal. A particularity of Haar measures is that any open non empty subset of G has a non null measure : indeed if there was such a subset by translation we could always cover any compact, which would have measure 0.

Remark : Haar measure is a bit a misnomer. Indeed it is a volume form, the measure itself is defined through charts on G (see Integral on manifolds). The use of notations such that d_L, d_R for Haar measures is just confusing.

Modular function

Theorem 1791 *For any left Haar measure ϖ_L on a finite dimensional Lie group G there is some non null function $\Delta(a)$, called a **modular function**, such that : $R_a^* \varpi_L = \Delta(a)^{-1} \varpi_L$*

Proof. $R_a^* \varpi_L = R_a^* (L_b^* \varpi_L) = (L_b R_a)^* \varpi_L = (R_a L_b)^* \varpi_L = L_b^* R_a^* \varpi_L = L_b^* (R_a^* \varpi_L)$ thus $R_a^* \varpi_L$ is still a left invariant measure, and because all left invariant measure are proportionnal there is some non null function $\Delta(a)$ such that : $R_a^* \varpi_L = \Delta(a)^{-1} \varpi_L$ ∎

Theorem 1792 *(Knapp p.532) The modular function on a finite dimensional Lie group G, with ϖ_L, ϖ_R left, rigth Haar measure, has the following properties :*

i) its value is given by : $\Delta(a) = |\det Ad_a|$

ii) $\Delta : G \to \mathbb{R}_+$ is a smooth group homomorphism

iii) if $a \in H$ and H is a compact or semi-simple Lie subgroup of G then $\Delta(a) = 1$

iv) $\Im^* \varpi_L = \Delta(x) \varpi_L$ are rigth Haar measures (with $\Im$ = inverse map)

v) $\Im^* \varpi_R = \Delta(x)^{-1} \varpi_R$ are left Haar measures

vi) $L_a \varpi_R = \Delta(a) \varpi_R$

Definition 1793 *A Lie group is said to be **unimodular** if any left Haar measure is a right measure (and vice versa). Then we say that any right or left invariant volume form is a **Haar measure**.*

A Lie group is unimodular iff $\forall a \in g : \Delta(a) = 1$.

Are unimodular the following Lie groups : abelian, compact, semi simple, nilpotent, reductive.

Decomposition of Haar measure

We have something similar to the Fubini theorem for Haar measures.

Theorem 1794 *(Knapp p.535) If S,T are closed Lie subgroups of the finite dimensional real Lie group G, such that : $S \cap T$ is compact, then the multiplication $M : S \times T \to G$ is an open map, the products ST exhausts the whole of G except for a null subset. Let Δ_S, Δ_T be the modular functions on S,T. Then any left Haar measure ϖ_L on S, T and G can be normalized so that :*

$\forall f \in C(G; \mathbb{R}):$

$\int_G f \varpi_L = \int_{S \times T} M^* \left(f \frac{\Delta_T}{\Delta_S} (\varpi_L)_S \otimes (\varpi_L)_T \right) = \int_S \varpi_L(s) \int_T \frac{\Delta_T(t)}{\Delta_S(s)} f(st) \varpi_L(t)$

Theorem 1795 *(Knapp p.538) If H is a closed Lie subgroup of a finite dimensional real Lie group G, Δ_G, Δ_H are the modular functions on G,H, if and only if the restriction of Δ_G to H is equal to Δ_H there is a volume form μ on G/H invariant with respect to the right action. Then it is unique up to a scalar and can be normalized such that :*

$\forall f \in C_{0c}(G; \mathbb{C}) : \int_G f \varpi_L = \int_{G/H} \mu(x) \int_H f(xh) \varpi_L(h)$ *with* $\pi_L : G \to G/H :: \pi_L(g) h = g \Leftrightarrow g^{-1} \pi_L(g) \in H$

G/H is not a group (if H is not normal) but can be endowed with a manifold structure, and the left action of G on G/H is continuous.

Comments

A Haar measure on a Lie group is a volume form so it is a Lebesgue measure on the manifold G, and is necessarily absolutely continuous. A Haar measure on a topological Group is a Radon measure, without any reference to charts, and can have a discrete part.

22.4 CLASSICAL LINEAR LIE GROUPS AND ALGEBRAS

What is usually called "classical Lie groups", or "classical linear groups" are Lie groups and Lie algebras of matrices. They are of constant use in the practical study of Lie groups and their properties have been extensively studied and referenced. So it is useful to sum up all the results.

22.5 General results

22.5.1 Notations

We denote here (they are traditional notations) :

The field K is either $\mathbb{R}$ or $\mathbb{C}$.

L(K,n) the set of square nxn matrices on a field K,

GL(K,n) the set of inversible square nxn matrices on a field K,

SL(K,n) the set of inversible square nxn matrices on a field K with determinant=1

L(K,n) is a Lie algebra on the field K with the bracket : $[X, Y] = [X][Y] - [Y][X]$

GL(K,n) is a Lie group on the field K with Lie algebra L(K,n) and group operations left and right multiplications of matrices.

The identity matrix is $I_n = diag(1, ..., 1) \in GL(K, n)$

The Lie groups of matrices are usually (the most convenient) denoted as $AB(K, n) \subset GL(K, n)$ where the letters AB specifie the group, and $SB(K, n) \subset SL(K, n)$ denotes the special group, subgroup of AB(K,n).

The Lie algebras of matrices are usually denoted as $ab(K, n) \subset L(K, n)$ where the lower case letters ab correspond of the Lie group which has the same Lie algebra.

It is common to consider matrices on the division ring of quaternions denoted H. This is not a field (it is not commutative), thus there are some difficulties (and so these matrices should be avoided).

A quaternion reads (see the Algebra part) : : $x = a + bi + cj + dk$ with $a, b, c, d \in \mathbb{R}$ and

$i^2 = j^2 = k^2 = -1, ij = k = -ji, jk = i = -kj, ki = j = -ik$

So we can write : $x = z_1 + j(c - id) = z_1 + jz_2, z_1, z_2 \in \mathbb{C}$

$xx' = z_1 z_1' + j^2 z_2 z_2' + j(z_1 z_2' + z_1' z_2) = z_1 z_1' - z_2 z_2' + j(z_1 z_2' + z_1' z_2) = x'x$

and a matrix on L(H,n) reads : $M = M_1 + jM_2, M_1, M_2 \in L(\mathbb{C}, n)$ so it can be considered as a couple of complex matrices with the multiplication rule :

$[MM']_q^p = \sum_r [M_1]_r^p [M_1']_q^r - [M_2]_r^p [M_2']_q^r + j \sum_r [M_1]_r^p [M_2']_q^r + [M_1']_r^p [M_2]_q^r$

that is : $MM' = M_1 M_1' - M_2 M_2' + j(M_1 M_2' + M_1' M_2)$

The identity is : $(I_n, 0)$

L(K,n) is a Banach finite dimensional vector space with the norm : $\|M\| = Tr(MM^*) = Tr(M^*M)$ where M* is the transpose conjugate.

22.5.2 Lie algebras

Definition

A Lie algebra of matrices on the field K is a subspace L of L(K,n) for some n, such that :

$\forall X, Y \in L, \forall r, r' \in K : r[X] + r'[Y] \in L, [X][Y] - [Y][X] \in L.$

The dimension m of L is usually different from n. A basis of L is a set of matrices $[e_i]_{i=1}^m$ which is a basis of L as vector space, and so the structure coefficients are given by :

$C_{jk}^i [e_i] = [e_j][e_k] - [e_k][e_j]$

with the Jacobi identies : $\forall i, j, k, p : \sum_{l=1}^m \left(C_{jk}^l C_{il}^p + C_{ki}^l C_{jl}^p + C_{ij}^l C_{kl}^p \right) = 0$

To be consistant with our definitions, L(K,n) is a Lie algebra on the field K. So $L(\mathbb{C}, n)$ is a Lie algebra on the field $\mathbb{C}$ but also on the field $\mathbb{R}$, by restriction of multiplication by a scalar to real scalars. Thus we have *real Lie algebras of complex matrices* : the only condition is that $\forall X, Y \in L, \forall r, r' \in \mathbb{R} : r[X] + r'[Y] \in L, [X][Y] - [Y][X] \in L.$

Complex and real structure

If L is a real Lie algebra, its complexified is just the set : $\{X + iY, X, Y \in L\} \subset L(\mathbb{C}, n)$ which is a complex Lie algebra with Lie bracket : $[X + iY, X' + iY']_{L_\mathbb{C}} = [X, X']_L - [Y, Y']_L + i([X, Y']_L + [X', Y]_L)$

If L is a complex Lie algebra, the obvious real structure is just : $[X] = [x] + i[y], [x], [y] \in L(\mathbb{R}, n)$

We have a real Lie algebra $L_\mathbb{R} = L_{0.} \oplus iL_0$, comprised of *couples* of real matrices, with the bracket above, with two isomorphic real subalgebras $L_{0.}, iL_0$.

L_0 is a real form of L.

If L is a even dimensional real Lie algebra endowed with a complex structure, meaning a linear map $J \in L(L; L)$ such that $J^2 = -Id_L$ and $J \circ ad = ad \circ J$ then take a basis of L $:(e_j)_{j=1}^{2m}$ with p=1...m : $J(e_j) = e_{j+m}, J(e_{j+m}) = -e_j$

The complex Lie algebra reads :

$L_\mathbb{C} = \sum_{p=1}^m (x^p + iy^p)[e_p] = \sum_{p=1}^m (x^p[e_p] + y^p[e_{p+m}])$

22.5.3 Lie groups

Definition

Let G be an algebraic subgroup G of GL(K,n) for some n. The group operations are always smooth and GL(K,n) is a Lie group. So G is a Lie subgroup if it is a submanifold of L(K,n).

There are several criteria to make a Lie group :

i) if G is closed in GL(K,n) it is a Lie group

ii) if G is a finite group (it is then open and closed) it is a Lie group

It is common to have a group of matrices defined as solutions of some equation. Let $F : L(K, n) \rightarrow L(K, n)$ be a differentiable map on the manifold L(K,n) and define G as a group whose matrices are solutions of F(M)=0. Then following the theorem of constant rank (see Manifolds) if F' has a constant rank r in

L(K,n) the set $F^{-1}(0)$ is a closed n^2-r submanifold of L(K,n) and thus a Lie subgroup. The map F can involve the matrix M and possibly its transpose and its conjugate. We know that a map on a complex Banach is not C-differentiable if its involve the conjugate. So usually a group of complex matrices defined by an equation involving the conjugate is not a complex Lie group, but can be a real Lie group (example : U(n) see below).

Remark : if F is continuous then the set $F^{-1}(0)$ is closed in L(K,n), but this is not sufficient : it should be closed in the Lie group GL(K,n), which is an open subset of L(K,n).

If G,H are Lie group of matrices, then $G \cap H$ is a Lie group of matrices with Lie algebra $T_1 G \cap T_1 H$

Connectedness

The connected component of the identity G_0 is a normal Lie subgroup of G, with the same dimension as G. The quotient set G/G_0 is a finite group. The other components of G can be written as : $g = g_L x = y g_R$ where x,y are in one of the other connected components, and g_R, g_L run over G_0.

If G is connected there is always a universal covering group $\widetilde{G}$ which is a Lie group. It is a compact group if G is compact. It is a group of matrices if G is a a complex semi simple Lie group, but otherwise $\widetilde{G}$ is not necessarily a group of matrices (ex : $GL(\mathbb{R}, n)$).

If G is not connected we can consider the covering group of the connected component of the identity G_0.

Translations

The translations are $L_A(M) = [A][M], R_A(M) = [M][A]$

The conjugation is : $Conj_A M = [A][M][A]^{-1}$ and the derivatives :

$\forall X \in L(K,n) : L'_A(M)(X) = [A][X], R'_A(M) = [X][A], (\Im(M))'(X) = -M^{-1}XM^{-1}$

So : $Ad_M X = Conj_M X = [M][X][M]^{-1}$

Lie algebra

The Lie algebra is a subalgebra of L(K,n)

If the Lie group is defined through a matrix equation involving $M, M^*, M^t, \overline{M}$: $P\left(M, M^*, M^t, \overline{M}\right) = 0$

Take a path : $M : \mathbb{R} \to G$ such that M(0)=I. Then $X = M'(0) \in T_1 G$ satisfies the polynomial equation :

$\left(\frac{\partial P}{\partial M} X + \frac{\partial P}{\partial M^*} X^* + \frac{\partial P}{\partial M^t} X^t + \frac{\partial P}{\partial \overline{M}} \overline{X}\right)|_{M=I} = 0$

Then a left invariant vector field is $X_L(M) = MX$

The exponential is computed as the exponential of a matrix : $\exp tX = \sum_{p=0}^{\infty} \frac{t^p}{p!}[X]^p$

Complex structures

Theorem 1796 *(Knapp p.442) For any real compact connected Lie group of matrices G there is a unique (up to isomorphism) closed complex Lie group of matrices whose Lie algebra is the complexified $T_1G \oplus iT_1G$ of T_1G.*

Cartan decomposition

Theorem 1797 *(Knapp p.445) Let be the maps :*
$\Theta : GL(K,n) \to GL(K,n) :: \Theta(M) = \left(M^{-1}\right)^*$,
$\theta : L(K,n) \to L(K,n) :: \theta(X) = -X^*$
If G is a connected closed semi simple Lie group of matrices in GL(K,n), invariant under Θ , then :
i) its Lie algebra is invariant under θ,
ii) $T_1G = l_0 \oplus p_0$ where l_0, p_0 are the eigenspaces corresponding to the eigen values +1,-1 of θ
iii) the map : $K \times p_0 \to G :: g = k \exp X$ where $K = \{x \in G : \Theta x = x\}$ is a diffeomorphism onto.

Groups of tori
A group of tori is defined through a family $[e_k]_{k=1}^m$ of commuting matrices in L(K,n) which is a basis of the abelian algebra. Then the group G is generated by : $[g]_k = \exp t\,[e_k] = \sum_{p=0}^\infty \frac{t^p}{p!}[e_k]^p, t \in K$
A group of diagonal matrices is a group of tori, but they are not the only ones.
The only compact complex Lie group of matrices are groups of tori.

22.6 List of classical Lie groups and algebras

See also Knapp (annex C) for all detailed information related to the roots systems and the exceptional groups.

22.6.1 GL(K,n)

$K=\mathbb{R},\mathbb{C}$
If n=1 we have the trivial group G={1} so we assume that n>1
The square nxn matrices on K which are inversible are a closed Lie group GL(K,n) of dimension n^2 over K, with Lie algebra L(K,n).
The center of GL(K,n) is comprised of scalar matrices $k\,[I]$
GL(K,n) is not semi simple, not compact, not connected.
$GL(\mathbb{C}, n)$ is the complexified of $GL(\mathbb{R}, n)$

SL(K,n) SL(K,n) is the Lie subgroup of GL(K,n) comprised of matrices such that detM=1. It has the dimension n^2-1 over K, and its Lie algebra is: sl(K,n)={ X∈L(K,n):Trace(X)=0}.
They are connected, semi-simple, not compact groups.
$SL(\mathbb{C}, n)$ is simply connected, and simple for n>1

$SL(\mathbb{R}, n)$ is not simply connected. For n>1 the universal covering group of $SL(\mathbb{R}, n)$ is not a group of matrices.

The complexified of $sl(\mathbb{R}, n)$ is $sl(\mathbb{C}, n)$, and $SL(\mathbb{C}, n)$ is the complexified of $SL(\mathbb{R}, n)$

The Cartan algebra of $sl(\mathbb{C}, n)$ is the subset of diagonal matrices.

The simple root system of $sl(\mathbb{C}, n)$ is $A_{n-1}, n \geq 2$:

$V = \sum_{k=1}^{n} x_k e_k, \sum_{k=1}^{n} x_k = 0$

$\Delta = e_i - e_j, i \neq j$

$\Pi = \{e_1 - e_2, e_2 - e_3, ..e_{n-1} - e_n\}$

22.6.2 Orthogonal groups O(K,n)

K=$\mathbb{R}, \mathbb{C}$

If n=1 we have the trivial group G={1} so we assume that n>1

O(K,n) is the Lie subgroup of GL(K,n) comprised of matrices such that $M^t M = I_n$ .Its dimension is n(n-1)/2 over K, and its Lie algebra is: o(K,n)={ X∈L(K,n):X+X^t=0}.

O(K,n) is semi-simple for n>2

O(K,n) is compact.

O(K,n) has two connected components, with detM=+1 and detM=-1. The connected components are not simply connected.

SO(K,n)

SO(K,n) is the Lie subgroup of O(K,n) comprised of matrices such that $\det M = 1$. Its dimension is n(n-1)/2 over K, and its Lie algebra is : o(K,n)={ X∈L(K,n):X+X^t=0}.

SO(K,n) is semi-simple for n>2

SO(K,n) is compact.

SO(K,n) is the connected component of the identity of O(K,n)

SO(K,n) is not simply connected. The universal covering group of SO(K,n) is the Spin group Spin(K,n) (see below) which is a double cover.

so$(\mathbb{C}, n)$ is the complexified of so$(\mathbb{R}, n)$, SO$(\mathbb{C}, n)$ is the complexified of SO$(\mathbb{R}, n)$

Roots system of $so(\mathbb{C}, n)$: it depends upon the parity of n

For $so(\mathbb{C}, 2n + 1), n \geq 1 : B_n$ system:

$V = \mathbb{R}^n$

$\Delta = \{\pm e_i \pm e_j, i < j\} \cup \{\pm e_k\}$

$\Pi = \{e_1 - e_2, e_2 - e_3, ..e_{n-1} - e_n, e_n\}$

For $so(\mathbb{C}, 2n), n \geq 2 : D_n$ system:

$V = \mathbb{R}^n$

$\Delta = \{\pm e_i \pm e_j, i < j\}$

$\Pi = \{e_1 - e_2, e_2 - e_3, ..e_{n-1} - e_n, e_{n-1} + e_n\}$

SO$(\mathbb{R}, 3)$

This is the group of rotations in the euclidean space. As it is used very often in physics it is good to give more details and some useful computational results.

1. The algebra $o(\mathbb{R}; 3)$ is comprised of 3x3 skewsymmetric matrices.

Take as basis for $o(\mathbb{R};3)$ the matrices :

$$\varepsilon_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} ; \varepsilon_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} ; \varepsilon_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

then a matrix of o(3) reads with the operator :

$$j : \mathbb{R}(3,1) \to o(\mathbb{R};3) :: j\left(\begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}\right) = \begin{bmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{bmatrix}$$

which has some nice properties :
$j(r)^t = -j(r) = j(-r)$
$j(x)y = -j(y)x = x \times y$
(this is just the "vector product $\times$" of elementary geometry)
$y^t j(x) = -x^t j(y)$
$j(x)y = 0 \Leftrightarrow \exists k \in R : y = kx$
$j(x)j(y) = yx^t - (y^t x) I$
$j(j(x)y) = yx^t - xy^t = j(x)j(y) - j(y)j(x)$
$j(x)j(y)j(x) = -(y^t x) j(x)$
$M \in L(\mathbb{R},3) : M^t j(Mx)M = (\det M) j(x)$
$M \in O(\mathbb{R},3) : j(Mx)My = Mj(x)y \Leftrightarrow Mx \times My = M(x \times y)$
$k > 0 : j(r)^{2k} = (-r^t r)^{k-1} j(r)j(r)$
$k \geq 0 : J(r)^{2k+1} = (-r^t r)^k j(r)$
2. The group $SO(\mathbb{R}, 3)$ is compact, thus the exponential is onto and any matrix can be written as :
$\exp(j(r)) = I_3 + j(r)\frac{\sin\sqrt{r^t r}}{\sqrt{r^t r}} + j(r)j(r)\frac{1-\cos\sqrt{r^t r}}{r^r r}$

The eigen values of g=expj(r) are $\left(1; \exp i\sqrt{r_1^2 + r_2^2 + r_3^2}; \exp\left(-i\sqrt{r_1^2 + r_2^2 + r_3^2}\right)\right)$

The vector r of components r_i is the axis of the rotation in $\mathbb{R}^3$ whose matrix is g in an orthonoral basis.
3. The universal covering group of $SO(\mathbb{R}, 3)$ is Spin$(\mathbb{R}, 3)$ isomorphic to SU(2).
If we take as basis of su(2) the matrices :
$e_1 = \frac{1}{2}\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}, e_2 = \frac{1}{2}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, e_3 = \frac{1}{2}\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$
then the cover is :
$\psi : su(2) \to o(\mathbb{R}, 3) :: \psi(r_1 e_1 + r_2 e_2 + r_3 e_3) = j(r)$
$\Psi : SU(2) \to SO(\mathbb{R}, 3) :: \Psi(\pm(r_1 e_1 + r_2 e_2 + r_3 e_3)) = \exp j(r)$

22.6.3 Unitary groups U(n)

The matrices of GL($\mathbb{C}$,n) such that $M^* M = I_n$ are a *real* Lie subgroup of GL($\mathbb{C}$,n) denoted U(n), with dimension n^2 over $\mathbb{R}$, and Lie algebra : u(n)={ X$\in$L($\mathbb{C}$,n):X+X*=0}. So U(n) is not a complex Lie group, even if its elements

comprise complex matrices. The algebra u(n) is a *real* Lie algebra and not a complex Lie algebra.

U(n) is not semi simple. The center of U(n) are the purely imaginary scalar matrices kiI_n

U(n) is compact.

U(n) is connected but not simply connected. Its universal covering group is $T \times SU(n) = \{e^{it}, t \in \mathbb{R}\} \times SU(n)$ with $\pi : T \to U(n) :: \pi \left(\left(e^{it} \times [g]\right)\right) = it\,[g]$ so for n=1 the universal cover is $(\mathbb{R}, +)$.

The matrices of $U(n) \cap GL(\mathbb{R}, n)$ comprised of real elements are just $O(\mathbb{R}, n)$.

SU(n)

The matrices of U(n) such that detM=1 are a *real* Lie subgroup of U(n) denoted SU(n), with dimension n^2-1 over $\mathbb{R}$, and Lie algebra : su(n)={ X∈L($\mathbb{C}$,n):X+X*=0, TrX=0}. So SU(n) is not a complex Lie group, even if its elements comprise complex matrices. The algebra su(n) is a *real* Lie algebra and not a complex Lie algebra.

SU(n) is not semi simple.

SU(n) is compact.

SU(n) is connected and simply connected.

The complexified of the Lie algebra $sl\,(\mathbb{C}, n)_{\mathbb{C}} = sl\,(\mathbb{C}, n)$ and the complexified $SU\,(n)_{\mathbb{C}} = SL\,(\mathbb{C}, n)$.

22.6.4   Special orthogonal groups O(K,p,q)

K=$\mathbb{R}, \mathbb{C}$

p>0,q>0, p+q=n>1

Let $I_{p,q}$ be the matrix : $I_{p,q} = Diag(+1, ... + 1, -1, ... - 1)$ (p + and q -)

The matrices of GL(K,n) such that $M^t I_{p,q} M = I_{p,q}$ are a Lie subgroup of GL(K,n) denoted O(K,p,q), with dimension n(n-1)/2 over K, and Lie algebra : $o(K, p, q) = \{X \in L(K, n) : I_{p,q}X + X^t I_{p,q} = 0\}$.

For K=$\mathbb{C}$ the group $O(\mathbb{C}, p, q)$ is isomorphic to $O(\mathbb{C}, p + q)$

$O(K, p, q), O\,(K, q, p)$ are identical : indeed $I_{p,q} = -I_{q,p}$

O(K,p,q) has four connected components, and each component is not simply connected.

O(K,p,q) is semi-simple for n>2

O(K,p,q) is not compact. The maximal compact subgroup is O(K,p)xO(K,q).

$O(\mathbb{C}, p + q)$ is the complexified of $O(\mathbb{R}, p, q)$.

SO(K,p,q)

The matrices of O(K,p,q) such that detM=1 are a Lie subgroup of O(K,p,q) denoted SO(K,p,q), with dimension n(n-1)/2 over K, and Lie algebra : $o(K, p, q) = \{X \in L(K, n) : I_{p,q}X + X^t I_{p,q} = 0\}$.

For K=$\mathbb{C}$ the group $SO(\mathbb{C}, p, q)$ is isomorphic to $SO(\mathbb{C}, p + q)$

$SO(\mathbb{R},p,q)$ is not connected, and has two connected components. Usually one considers the connected component of the identity $SO_0(\mathbb{R},p,q)$. The universal covering group of $SO_0(\mathbb{R},p,q)$ is Spin($\mathbb{R}$,p,q).

SO($\mathbb{R}$,p,q) is semi-simple for n>2

SO($\mathbb{R}$,p,q) is not compact. The maximal compact subgroup is SO(K,p)xSO(K,q).

SO($\mathbb{C}$, $p + q$) is the complexified of $SO(\mathbb{R}, p, q)$.

$SO(\mathbb{R}, p, q)$

$O(\mathbb{R}, p, q)$ is invariant by transpose, and admits a Cartan decomposition :

$$o(\mathbb{R}, p, q) = l_0 \oplus p_0 \text{ with} : l_0 = \left\{ l = \begin{bmatrix} M_{p \times p} & 0 \\ 0 & N_{q \times q} \end{bmatrix} \right\}, p_0 = \left\{ p = \begin{bmatrix} 0 & P_{p \times q} \\ P_{q \times p}^t & 0 \end{bmatrix} \right\}$$

$[l_0, l_0] \subset l_0, [l_0, p_0] \subset p_0, [p_0, p_0] \subset l_0$

So the maps :

$\lambda : l_0 \times p_0 \to SO(\mathbb{R}, p, q) :: \lambda(l, p) = (\exp l)(\exp p) ;$

$\rho : p_0 \times l_0 \to SO(\mathbb{R}, p, q) :: \rho(p, l) = (\exp p)(\exp l) ;$

are diffeomorphisms;

It can be proven (see Algebra - Matrices) that :

i) the Killing form is $B(X, Y) = \frac{n}{2} Tr(XY)$

ii) $\exp p = \begin{bmatrix} I_p + H(\cosh D - I_q) H^t & H(\sinh D) U^t \\ U(\sinh D) H^t & U(\cosh D) U^t \end{bmatrix}$ with $H_{p \times q}$ such that :

$H^t H = I_q, P = HDU^t$ where D is a real diagonal qxq matrix and U is a qxq real orthogonal matrix. The powers of exp(p) can be easily deduced.

$\mathbf{SO}(\mathbb{R}, 3, 1)$

This is the group of rotations in the Minkovski space (one considers also $SO(\mathbb{R}, 1, 3)$ which is the same).

1. If we take as basis of the algebra the matrices :

$$l_0 : \varepsilon_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} ; \varepsilon_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} ; \varepsilon_3 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$p_0 : \varepsilon_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} ; \varepsilon_5 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} ; \varepsilon_6 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

It is easy to show that the map j of SO(3) extends to a map :

$$J : \mathbb{R}(3, 1) \to o(\mathbb{R}; 3, 1) :: J \left(\begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} \right) = \begin{bmatrix} 0 & -r_3 & r_2 & 0 \\ r_3 & 0 & -r_1 & 0 \\ -r_2 & r_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

with the same identities as above with j.

We have similarly :

$$K : \mathbb{R}(3, 1) \to o(\mathbb{R}; 3, 1) :: K \left(\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \right) = \begin{bmatrix} 0 & 0 & 0 & v_1 \\ 0 & 0 & 0 & v_2 \\ 0 & 0 & 0 & v_3 \\ v_1 & v_2 & v_3 & 0 \end{bmatrix}$$

And $\forall X \in O\left(\mathbb{R}, 3, 1\right) : \exists r, v \in \mathbb{R}\left(3, 1\right) : X = J(r) + K(v)$

The identities above read:

$$\exp K(v) = \begin{bmatrix} I_3 + \left(\cosh \sqrt{v^t v} - 1\right) \frac{v v^t}{v^t v} & \frac{v}{\sqrt{v^t v}} \left(\sinh \sqrt{v^t v}\right) \\ \left(\sinh \sqrt{v^t v}\right) \frac{v^t}{\sqrt{v^t v}} & \cosh \sqrt{v^t v} \end{bmatrix}$$

that is :

$\exp K(v) = I_4 + \frac{\sinh \sqrt{v^t v}}{\sqrt{v^t v}} K(v) + \frac{\cosh \sqrt{v^t v} - 1}{v^t v} K(v) K(v)$

Similarly :

$\exp J(r) = I_4 + \frac{\sin \sqrt{r^t r}}{\sqrt{r^t r}} J(r) + \frac{1 - \cos \sqrt{r^t r}}{r^r r} J(r) J(r)$

2. The universal covering group of $SO_0(\mathbb{R},3,1)$ is $Spin(\mathbb{R},3,1)$ which is isomorphic to $SL(\mathbb{C}, 2)$. It is more explicit with the map :

$\phi : \mathbb{C}^3 \to sl\left(\mathbb{C}, 2\right) :: \phi\left(z\right) = \frac{1}{2} \begin{bmatrix} z_3 & z_1 + i z_2 \\ z_1 - i z_2 & -z_3 \end{bmatrix}$ which is an isomorphism

between vector spaces.

Then the map : $\psi : sl(2, C) \to o(3, 1) : \psi\left(\phi\left(z\right)\right) = K(\mathrm{Re}\, z) + J(\mathrm{Im}\, z)$ is a Lie algebra real isomorphism and the map :

$\Psi : SL(\mathbb{C}, 2) \to SO_0\left(\mathbb{R}, 3, 1\right) ::$

$\Psi\left(\pm\left(\exp i \phi\left(\mathrm{Im}\, z\right)\right) \exp \phi\left(\mathrm{Re}\, z\right)\right) = \left(\exp J\left(\mathrm{Im}\, z\right)\right)\left(\exp K\left(\mathrm{Re}\, z\right)\right)$

is a double cover of $SO_0\left(\mathbb{R}, 3, 1\right)$ and $\phi = \Psi'(1)$

22.6.5 Special unitary groups U(p,q)

p>0,q>0, p+q=n>1

Let $I_{p,q}$ be the matrix : $I_{p,q} = Diag(+1, ... +1, -1, ... -1)$ (p + and q -)

The matrices of GL($\mathbb{C}$,n) such that $M^* I_{p,q} M = I_{p,q}$ are a *real* Lie subgroup of GL(p,q), with dimension n² over $\mathbb{R}$, and Lie algebra : u(p,q)={ X∈L($\mathbb{C}$,n):$I_{p,q} X + X^* I_{p,q}$=0}. So U(n) is not a complex Lie group, even if its elements comprise complex matrices. The algebra u(p,q) is a *real* Lie algebra and not a complex Lie algebra.

$U(p, q), U\left(q, p\right)$ are identical : indeed $I_{p,q} = -I_{q,p}$

U(p,q) is semi-simple

U(p,q) is not compact.

U(p,q) has two connected components.

SU(p,q)

The matrices of U(p,q) such that detM=1 are a *real* Lie subgroup of U(n) denoted SU(p,q), with dimension n² over $\mathbb{R}$, and Lie algebra : u(p,q)={ X∈L($\mathbb{C}$,n):$I_{p,q} X + X^* I_{p,q}$=0}. So SU(n) is not a complex Lie group, even if its elements comprise complex matrices.

$SU(p, q), SU\left(q, p\right)$ are identical.

SU(p,q) is a semi-simple, connected, non compact group

SU(p,q) is the connected component of the identity of U(p,q).

22.6.6 Symplectic groups Sp(K,n)

K=$\mathbb{R}, \mathbb{C}$

Let J_n be the 2nx2n matrix : $J_n = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$

The matrices of GL(K,2n) such that $M^t J_n M = J$ are a Lie subgroup of GL(K,2n) denoted Sp(K,n) over K, and Lie algebra : $sp(K, p, q) = \{X \in L(K, n) : JX + X^t J = 0\}$. Notice that we have either real or complex Lie groups.

Sp(K,n) is a semi-simple, connected, non compact group.

Root system for $sp(\mathbb{C}, n), n \geq 3 : C_n$

$V = \mathbb{R}^n$

$\Delta = \{\pm e_i \pm e_j, i < j\} \cup \{\pm 2e_k\}$

$\Pi = \{e_1 - e_2, e_2 - e_3, ..e_{n-1} - e_n, 2e_n\}$

22.6.7 Pin and Spin groups

These groups are defined starting from Clifford algebra over a finite dimensional vector space F on a field K, endowed with a bilinear symmetric form g (valued in K) (see the Algebra part).

1. All Clifford algebras on vector spaces with the same dimension on the same field, with bilinear form of the same signature are isomorphic. So we can speak of Pin(K,p,q), Spin(K,p,q).

There are groups morphism :

$\mathbf{Ad} : Pin(K, p, q) \rightarrow O(K, p, q)$

$\mathbf{Ad} : Spin(K, p, q) \rightarrow SO(K, p, q)$

with a double cover (see below) : for any $g \in O(K, p, q)$ (or SO(K,p,q) there are two elements $\pm w$ of Pin(K,p,q) (or Spin(K,p,q) such that : $\mathbf{Ad}_w = h$.

They have the same Lie algebra : o(K,p,q) is the Lie algebra of Pin(K,p,q) and so(K,p,q) is the Lie algebra of Spin(K,p,q).

2. The situation with respect to the cover is a bit complicated. We have always two elements of the Pin or Spin group for one element of the orthogonal group, and they are a cover as a manifold, but not necessarily the universal cover as Lie group, which has been defined only for connected Lie groups. O(K,p,q) or SO(K,p,q) are Lie groups, but not necessarily connected.

When they are connected, they have a unique universal cover as a topological space, which has a unique Lie group structure $\widetilde{G}$ which is a group of matrices, which can be identified to Pin(K,p,q) or Spin(K,p,q) respectively.

When they are disconnected, the same result is valid for their connected component, which is a Lie subgroup. One can take the same cover as topological spaces on the other connected components (which are principal homogeneous spaces for the identity component) but the group structure on the other components is not uniquely determined in general.

For all practical purposes one consider usually only the connected component, from which some simple construct can be found for the others when necessary.

3. If K=$\mathbb{C}$:

Pin$(\mathbb{C}, p, q) \simeq Pin(\mathbb{C}, p+q)$

Spin$(\mathbb{C}, p, q) \simeq Spin(\mathbb{C}, p+q)$

Spin$(\mathbb{C}, n)$ is connected, simply connected.

SO$(\mathbb{C}, n) \simeq SO(\mathbb{C}, p, q)$ is a semi simple, complex Lie group, thus its universal covering group is a group of matrices which can be identified with Spin$(\mathbb{C}, n)$.

Spin$(\mathbb{C}, n)$ and SO$(\mathbb{C}, n)$ have the same Lie algebra which is compact, thus Spin$(\mathbb{C}, n)$ is compact.

We have the isomorphisms :

$SO(\mathbb{C}, n,) \simeq Spin(\mathbb{C}, n)/U(1)$

$Spin(\mathbb{C}, 2) \simeq \mathbb{C}$

$Spin(\mathbb{C}, 3) \simeq SL(\mathbb{C}, 2)$

$Spin(\mathbb{C}, 4) \simeq SL(\mathbb{C}, 2) \times SL(\mathbb{C}, 2)$

$Spin(\mathbb{C}, 5) \simeq Sp(\mathbb{C}, 4)$

$Spin(\mathbb{C}, 6) \simeq SL(\mathbb{C}, 4)$

5. If K=$\mathbb{R}$

Pin($\mathbb{R}$,p,q),Pin($\mathbb{R}$,q,p) are not isomorphic if $p \neq q$

Pin($\mathbb{R}$,p,q) is not connected, it maps to O($\mathbb{R}$,p,q) but the map is not surjective and it is not a cover of O($\mathbb{R}$,p,q)

Spin($\mathbb{R}$,p,q) and Spin($\mathbb{R}$,q,p) are isomorphic, and simply connected if p+q>2

Spin($\mathbb{R}$,0,n) and Spin($\mathbb{R}$,n,0) are equal to Spin($\mathbb{R}$,n)

For n>2 Spin($\mathbb{R}$,n) is connected, simply connected and is the universal cover of SO($\mathbb{R}$,n) and has the same Lie algebra, so it is compact.

If p+q>2 Spin($\mathbb{R}$,p,q) is connected, simply connected and is a double cover of SO$_0$($\mathbb{R}$,p,q) and has the same Lie algebra, so it is not compact.

We have the isomorphisms :

Spin($\mathbb{R}$,1) $\simeq$ O($\mathbb{R}$,1)

Spin($\mathbb{R}$,2) $\simeq$ U(1) $\simeq$ SO($\mathbb{R}$,2)

Spin($\mathbb{R}$,3) $\simeq$ Sp(1) $\simeq$ SU(2)

Spin($\mathbb{R}$,4) $\simeq$ Sp(1) x Sp(1)

Spin($\mathbb{R}$,5) $\simeq$ Sp(2)

Spin($\mathbb{R}$,6) $\simeq$ SU(4)

Spin($\mathbb{R}$,1,1) $\simeq$ $\mathbb{R}$

Spin($\mathbb{R}$,2,1) = SL(2,R)

Spin($\mathbb{R}$,3,1) = SL($\mathbb{C}$,2)

Spin($\mathbb{R}$,2,2) = SL($\mathbb{R}$,2) x SL($\mathbb{R}$,2)

Spin($\mathbb{R}$,4,1) = Sp(1,1)

Spin($\mathbb{R}$,3,2) = Sp(4)

Spin($\mathbb{R}$,4,2) = SU(2,2)

22.7 Heisenberg group

The Heisenberg group is met in quantum theory.

22.7.1 Definition

Let be E a symplectic vector space, meaning a real finite n-dimensional vector space endowed with a non degenerate 2-form $h \in \Lambda_2 E$. So n must be even : n=2m

Take the set $E \times \mathbb{R}$, endowed with its natural structure of vector space $E \oplus \mathbb{R}$ and the internal product $\cdot$:

$\forall u, v \in E, x, y \in \mathbb{R} : (u, x) \cdot (v, y) = \left(u + v, x + y + \frac{1}{2} h(u, v)\right)$

The product is associative

the identity element is (0,0) and each element has an inverse :

$(u, x)^{-1} = (-u, -x)$

So it has a group structure. $E \times \mathbb{R}$ with this structure is called the **Heisenberg group** H(E,h).

As all symplectic vector spaces with the same dimension are isomorphic, all Heisenberg group for dimension n are isomorphic and the common structure is denoted H(n).

22.7.2 Properties

The Heisenberg group is a connected, simply-connected Lie group. It is isomorphic (as a group) to the matrices of $GL(\mathbb{R}, n + 1)$ which read :

$$\begin{bmatrix} 1 & [p]_{1 \times n} & [c]_{1 \times 1} \\ 0 & I_n & [q]_{n \times 1} \\ 0 & 0 & 1 \end{bmatrix}$$

$E \times \mathbb{R}$ is a vector space and a Lie group. Its Lie algebra denoted also H(E,h) is the set $E \times \mathbb{R}$ itself with the bracket:

$[(u, x), (v, y)] = (u, x) \cdot (v, y) - (v, y) \cdot (u, x) = (0, h(u, v))$

Take a canonical basis of E : $(e_i, f_i)_{i=1}^m$ then the structure coefficients of the Lie algebra H(n) are :

$[(e_i, 1), (f_j, 1)] = (0, \delta_{ij})$ all the others are null

It is isomorphic (as Lie algebra) to the matrices of $L(\mathbb{R}, n + 1)$:

$$\begin{bmatrix} 0 & [p]_{1 \times n} & [c]_{1 \times 1} \\ 0 & [0]_{n \times n} & [q]_{n \times 1} \\ 0 & 0 & 0 \end{bmatrix}$$

There is a complex structure on E defined from a canonical basis $(\varepsilon_i, \varphi_i)_{i=1}^m$ by taking a complex basis $(\varepsilon_j, i\varphi_j)_{j=1}^m$ with complex components. Define the new complex basis :

$a_k = \frac{1}{\sqrt{2}}(\varepsilon_k - i\varphi_k), a_k^\dagger = \frac{1}{\sqrt{2}}(\varepsilon_k + i\varphi_k)$

and the commutation relations becomes : $[a_j, a_k] = \left[a_j^\dagger, a_k^\dagger\right] = 0; \left[a_j, a_k^\dagger\right] = \delta_{jk}$ called CAR

22.7.3 Representations

All irreducible finite dimensional linear representation of H(n) are 1-dimensional. The character is :

$\chi_{ab}(x, y, t) = e^{-2i\pi(ax+by)}$ where $(a, b) \in \mathbb{R}^n \times \mathbb{R}^n$

The only unitary representations are infinite dimensional over $F = L^2(\mathbb{R}^n)$

$\lambda \neq 0 \in r, f \in L^2(\mathbb{R}^n), (x, y, t) \in H(n)$

$\rightarrow \rho_\lambda(x, y, t) f(s) = (\exp(-2i\pi\lambda t - i\pi\lambda \langle x, y \rangle + 2i\pi\lambda \langle s, y \rangle)) f(s - x)$

Two representations ρ_λ, ρ_μ are equivalent iff $\lambda = \mu$. Then they are unitary equivalent.

22.7.4 Heisenberg group on Hilbert space

There is a generalization of the previous definition for complex Hilbert spaces H (possibly infinite dimensional) (Neeb p.102)

The Heisenberg group "Heis" is $\mathbb{R} \times H$ endowed with the product :

$(t, v)(s, w) = \left(t + s - \frac{1}{2}\operatorname{Im}\langle v, w \rangle, v + w\right)$

$\sigma(t, v)(u) = u + v$ defines a continuous action of Heis on H

The Heisenberg group is a topological group and has unitary representations on the Fock space.

23 REPRESENTATION THEORY

23.1 Definitions and general results

Let E be a vector space on a field K. Then the set L(E;E) of linear endomorphisms of E is : an algebra, a Lie algebra and its subset GL(E;E) of inversible elements is a Lie group. Thus it is possible to consider morphisms between algebra, Lie algebra or groups and L(E;E) or GL(E;E). Moreover if E is a Banach vector space then we can consider continuous, differentiable or smooth morphisms.

23.1.1 Definitions

Definition 1798 *A linear* **representation of the Banach algebra** *(A,·) over a field K is a couple (E,f) of a Banach vector space E over K and a smooth map* $f : A \to \mathcal{L}(E; E)$ *which is an algebra morphism :*

$\forall X, Y \in A, k, k' \in K : f(kX + k'Y) = kf(X) + k'f(Y)$
$\forall X, Y \in A, : f(X \cdot Y) = f(X) \circ f(Y)$
$f \in \mathcal{L}(A; \mathcal{L}(E; E))$
If A is unital (there is a unit element I) then we require that f(I)=Id$_E$
Notice that f must be K linear.

The representation is over an algebra of linear maps, so this is a geometrical representation (usually called linear representation). A Clifford algebra is an algebra, so it enters the present topic, but morphisms of Clifford algebras have some specificities which are addressed in the Algebra part.

The representation of Banach algebras has been addressed in the Analysis part. So here we just recall some basic definitions.

Definition 1799 *A linear* **representation of the Banach Lie algebra** *(A,[]) over a field K is a couple (E,f) of a Banach vector space E over K and a smooth map* $f : A \to \mathcal{L}(E; E)$ *which is a Lie algebra morphism:*

$\forall X, Y \in A, k, k' \in K : f(kX + k'Y) = kf(X) + k'f(Y)$
$\forall X, Y \in A, : f([X, Y]) = f(X) \circ f(Y) - f(Y) \circ f(X) = [f(X), f(Y)]_{\mathcal{L}(E;E)}$
$f \in \mathcal{L}(A; \mathcal{L}(E; E))$
Notice that f must be K linear.
If E is a Hilbert space H then $\mathcal{L}(H; H)$ is a C*-algebra.

Definition 1800 *A linear* **representation of the topological group G** *is a couple (E,f) of a topological vector space E and a continuous map* $f : G \to G\mathcal{L}(E; E)$ *which is a continuous group morphism :*

$\forall g, h \in G : f(gh) = f(g) \circ f(h), f(g^{-1}) = (f(g))^{-1} ; f(1) = Id_E$
$f \in C_0(G; G\mathcal{L}(E; E))$
f is usually not linear but f(g) must be inversible. E can be over any field.

Definition 1801 *A linear **representation of the Lie group G** over a field K is a couple (E,f) of a Banach vector space E over the field K and a differentiable class $r \geq 1$ map $f : G \rightarrow G\mathcal{L}(E; E)$ which is a Lie group morphism :*

$\forall g, h \in G : f(gh) = f(g) \circ f(h), f(g^{-1}) = (f(g))^{-1} ; f(1) = Id_E$
$f \in C_r(G; G\mathcal{L}(E; E))$
a continuous Lie group morphism is necessarily smooth

Definition 1802 *A representation (E,f) is **faithful** if the map f is bijective.*

Then we have an isomorphism, and conversely if we have an isomorphism with some subset of linear maps over a vector space we have a representation.

Definition 1803 *The **trivial representation** (E,f) of an algebra or a Lie algebra is $f(X) = 0$ with any E.*
The trivial representation (E,f) of a group is $f(g) = Id$ with any E.

Definition 1804 *The **standard representation** of a Lie algebra of matrices (resp.of a Lie group of matrices) in L(K,n) is (K^n, f) where f(X) (resp.f(g)) is just the linear map in $L(K^n; K^n)$ whose matrix is X (resp.g) in the canonical basis .*

Matrix representation : any representation (E,f) on a finite dimensional vector space E becomes a representation on a set of matrices by choosing a basis. But a representation is not necessarily faithful, and the algebra ot the group may not to be isomorphic to a set of matrices. The matrix representation depends on the choice of a basis, which can be specific (usually orthonormal).

Definition 1805 *An **interwiner** between two representations $(E_1, f_1), (E_2, f_2)$ of a set X is a morphism : $\phi \in \mathcal{L}(E_1; E_2)$ such that :*
$\forall x \in X : \phi \circ f_1(x) = f_2(x) \circ \phi$
*If ϕ is an isomorphism then the two representations are said to be **equivalent**.*

Conversely, if there is an isomorphism $\phi \in G\mathcal{L}(E_1; E_2)$ between two vector spaces E_1, E_2 and if (E_1, f_1) is a representation, then with $f_2(x) = \phi \circ f_1(x) \circ \phi^{-1}$, (E_2, f_2) is an equivalent representation. Two vector spaces on the same field and same finite dimension are isomorphic so, representation wise, the nature of E does not matter, and for a finite dimensional representation we can take $E = K^n$. But we can wish to endow E with some additional structure (usually an internal product) and then it becomes more complicated.

Definition 1806 *An **invariant vector space** in a representation (E,f) is a vector subspace F of E such that: $\forall x \in X, \forall u \in F : f(x)(u) \in F$*

If V is a closed vector subspace of E, invariant in the representation, then $(V, f_{|V})$ where f_V is the restriction of f to $\mathcal{L}(V;V)$, is still a representation.

Definition 1807 *A representation (E,f) is **irreducible** if the only closed invariant vector subspaces are 0 and E.*

Irreducible representations are useful because many representations can be built from simpler irreducible representations.

Definition 1808 *If H is a Lie subgroup of G, (E,f) a representation of G, then (E, f_H) where f_H is the restriction of f to H, is a representation of H, called a **subrepresentation**.*

Similarly If B is a Lie subalgebra of A, (E,f) a representation of A, then (E, f_B) where f_B is the restriction of f to B, is a representation of B.

23.1.2 Complex and real representations

These results seem obvious but are very useful, as many classical representations can be derived from each other by using simultaneously complex and real representations.

1. For any complex representation (E,f), with E a complex vector space and f a $\mathbb{C}$-differentiable map, $(E_\mathbb{R} \oplus iE_\mathbb{R}, f)$ is a real representation with any real structure on E.

So if (E,f) is a complex representation of a complex Lie algebra or Lie group we have easily a representation of any of their real forms.

2. There is a bijective correspondance between the real representations of a real Lie algebra A and the complex representations of its complexified $A_\mathbb{C} = A \oplus iA$. And one representation is irreducible iff the other is irreducible.

A real representation (E,f) of a Lie algebra A can be extended uniquely in a complex representation $(E_\mathbb{C}, f_\mathbb{C})$ of $A_\mathbb{C}$ by $f_\mathbb{C}(X + iY) = f(X) + if(Y)$

Conversely if A is a real form of the complex Lie algebra B=$A \oplus iA$, any complex representation (E,f) of B gives a representation of A by taking the restriction of f to the vector subspace A.

3. If (E,f) is a complex representation of a Lie group G and σ a real structure on E, there is always a conjugate complex vector space $\overline{E}$ and a bijective antilinear map : $\sigma : E \to \overline{E}$. To any $f(g) \in L(E;E)$ we can associate a unique conjugate map :

$\overline{f}(g) = \sigma \circ f(g) \circ \sigma^{-1} \in \underline{L}(E;E) :: \overline{f}(g)u = \overline{f(g)\overline{u}}$

If f is a real map then $\overline{f}(g) = f(g)$. If not they are different maps, and we have the **conjugate representation** (also called contragrediente) : $(E; \overline{f})$ which is not equivalent to (E,f).

23.1.3 Sum and product of representations

Given a representation (E,f) we can define infinitely many other representations, and in physics finding the right representation is often a big issue (exemple in the standard model).

Lie algebras - Sum of representations

Definition 1809 *The **sum** of the representations $(E_i, f_i)_{i=1}^r$ of the Lie algebra $(A, [])$, is the representation $(\oplus_{i=1}^r E_i, \oplus_{i=1}^r f_i)$*

For two representations :
$f_1 \oplus f_2 : A \to \mathcal{L}(E_1 \oplus E_2; E_1 \oplus E_2) :: (f_1 \oplus f_2)(X) = f_1(X) + f_2(X)$
So : $(f_1 \oplus f_2)(X)(u_1 \oplus u_2) = f_1(X) u_1 + f_2(X) u_2$
The bracket on $E_1 \oplus E_2$ is $[u_1 + u_2, v_1 + v_2] = [u_1, v_1] + [u_2, v_2]$
The bracket on $\mathcal{L}(E_1 \oplus E_2; E_1 \oplus E_2)$ is $[\varphi_1 \oplus \varphi_2, \psi_1 \oplus \psi_2] = [\varphi_1, \psi_1] + [\varphi_2, \psi_2] = \varphi_1 \circ \psi_1 - \psi_1 \circ \varphi_1 + \varphi_2 \circ \psi_2 - \psi_2 \circ \varphi_2$

The direct sum of representations is not irreducible. Conversely a representation is said to **completely reducible** if it can be expressed as the direct sum of irreducible representations.

Lie algebras - tensorial product of representations

Definition 1810 *The **tensorial product of the representations** $(E_i, f_i)_{i=1}^r$ of the Lie algebra $(A, [])$ is a representation $(E = E_1 \otimes E_2 ... \otimes E_r, D)$ with the morphism D defined as follows : for any $X \in A$, $D(X)$ is the unique extension of $\phi(X) = \sum_{k=1}^r Id \otimes ... \otimes f(X) \otimes ... \otimes Id \in L^r(E_1, ..E_r; E)$ to a map $L(E; E)$ such that : $\phi(X) = D(X) \circ \imath$ with the canonical map : $\imath : \prod_{k=1}^r E^r \to E$*

As $\phi(X)$ is a multilinear map such an extension always exist.
So for two representations :
$D(X)(u_1 \otimes u_2) = (f_1(X) u_1) \otimes u_2 + u_1 \otimes (f_2(X) u_2)$
The bracket on $L(E_1 \otimes E_2; E_1 \otimes E_2)$ is
$[f_1(X) \otimes Id_2 + Id_1 \otimes f_2(X), f_1(Y) \otimes Id_2 + Id_1 \otimes f_2(Y)]$
$= (f_1(X) \otimes Id_2 + Id_1 \otimes f_2(X)) \circ (f_1(Y) \otimes Id_2 + Id_1 \otimes f_2(Y)) - (f_1(Y) \otimes Id_2 + Id_1 \otimes f_2(Y)) \circ (f_1(X) \otimes Id_2 + Id_1 \otimes f_2(X))$
$= [f_1(X), f_1(Y)] \otimes Id_2 + Id_1 \otimes [f_2(X), f_2(Y)] = (f_1 \times f_2)([X, Y])$
If all representations are irreducible, then their tensorial product is irreducible.
If (E,f) is a representation the procedure gives a representation $(\otimes^r E, D^r f)$

Definition 1811 *If (E,f) is a representation of the Lie algebra $(A, [])$ the representation $(\wedge^r E; , D_A^r f)$ is defined by extending the antisymmetric map : $\phi_A(X) \in L^r(E^r; \wedge^r E) :: \phi_A(X) = \sum_{k=1}^r Id \wedge ... \wedge f(X) \wedge ... \wedge Id$*

Definition 1812 *If (E,f) is a representation of the Lie algebra $(A, [])$ the representation $(\odot^r E; , D_S^r f)$ is defined by extending the symmetric map $\phi_S(X) \in L^r(E^r; S^r(E)) :: \phi_S(X) = \sum_{k=1}^r Id \odot ... \odot f(X) \odot ... \odot Id$*

Remarks :

i) $\odot^r E, \wedge^r E \subset \otimes^r E$ as vector subspaces.

ii) If a vector subspace F of E is invariant by f, then $\otimes^r F$ is invariant by $D^r f$, as $\odot^r E, \wedge^r F$ for $D^r_s f, D^r_a f$.

iii) If all representations are irreducible, then their tensorial product is irreducible.

Groups - sum of representations

Definition 1813 *The **sum of the representations** $(E_i, f_i)^r_{i=1}$ of the group G, is a representation $(\oplus^r_{i=1} E_i, \oplus^r_{i=1} f_i)$*

For two representations : $(f_1 \oplus f_2)(g)(u_1 \oplus u_2) = f_1(g)u_1 + f_2(g)u_2$

The direct sum of representations is not irreducible. Conversely a representation is said to **completely reducible** if it can be expressed as the direct sum of irreducible representations.

Groups - tensorial product of representations

Definition 1814 *The **tensorial product of the representations** $(E_i, f_i)^r_{i=1}$ of the group G is a representation*

$(E = E_1 \otimes E_2 ... \otimes E_r, D)$ with the morphism D defined as follows : for any $g \in G$, $D(g)$ is the unique extension of $\phi(g)(u_1,..,u_r) = f_1(g)u_1 \otimes ... \otimes f_r(g)u_r \in \mathcal{L}^r(E_1,..,E_r;E)$ to $D(g) \in \mathcal{L}(E;E)$ such that $\phi(g) = D(g) \circ \imath$ with the canonical map : $\imath : \prod\limits_{k=1}^{r} E^r \to E$

As $\phi(g)$ is a multilinear map such an extension always exist.

If (E,f) is a representation the procedure gives a representation $(\otimes^r E, D^r f)$

Definition 1815 *If (E,f) is a representation of the group G the representation $(\wedge^r E;, D^r_A f)$ is defined by extending the antisymmetric map : $\phi_A(g) \in L^r(E^r; \wedge^r E) :: \phi_A(g) = \sum_{k=1}^{r} f(g) \wedge ... \wedge f(g) \wedge ... \wedge f(g)$*

Definition 1816 *If (E,f) is a representation of the group G the representation $(\odot^r E;, D^r_S f)$ is defined by extending the symmetric map $\phi_S(g) \in L^r(E^r; S^r(E)) :: \phi_S(g) = \sum_{k=1}^{r} f(g) \odot ... \odot f(g) \odot ... \odot f(g)$*

Remarks :

i) $\odot^r E, \wedge^r E \subset \otimes^r E$ as vector subspaces.

ii) If a vector subspace F of E is invariant by f, then $\otimes^r F$ is invariant by $D^r f$, as $\odot^r E, \wedge^r F$ for $D^r_s f, D^r_a f$.

iii) If all representations are irreducible, then their tensorial product is irreducible.

23.1.4 Representation of a Lie group and its Lie algebra

From the Lie group to the Lie algebra

Theorem 1817 *If (E,f) is a representation of the Lie group G then (E,f'(1)) is a representation of T_1G and $\forall u \in T_1G : f\left(\exp_G u\right) = \exp_{G\mathcal{L}(E;E)} f'(1)u$*

Proof. f is a smooth morphism $f \in C_\infty\left(G; G\mathcal{L}\left(E;E\right)\right)$ and its derivative f'(1) is a morphism : $f'(1) \in \mathcal{L}\left(T_1G; \mathcal{L}\left(E;E\right)\right)$ ■
   The exponential on the right side is computed by the usual series.

Theorem 1818 *If $(E_1, f_1), (E_2, f_2)$ are two equivalent representations of the Lie group G, then $(E_1, f'_1(1)), (E_2, f'_2(1))$ are two equivalent representations of T_1G*

Theorem 1819 *If the closed vector subspace F is invariant in the representation (E,f) of the Lie group G, then F is invariant in the representation (E,f'(1)) of T_1G .*

Proof. (F,f) is a representation is a representation of G, so is (F,f'(1)) ■

Theorem 1820 *If the Lie group G is connected, the representation (E,f) of G is irreducible (resp.completely reducible) iff the representation (E,f'(1)) of its Lie algebra is irreducible (resp.completely reducible)*

   Remark :
   If $(E_1, f_1), (E_2, f_2)$ are two representations of the Lie group G, the derivative of the product of the representations is :
   $(f_1 \otimes f_2)'(1) : T_1G \to \mathcal{L}\left(E_1 \otimes E_2; E_1 \otimes E_2\right) :: (f_1 \otimes f_2)'(1)(u_1 \otimes u_2) = (f'_1(1)u_1) \otimes u_2 + u_1 \otimes (f'_2(1)u_2)$
   that is the product $(E_1 \otimes E_2, f'_1(1) \times f'_2(1))$ of the representations $(E_1, f'_1(1)), (E_2, f'_2(1))$ of T_1G
   We have similar results for the sum of representations.

From the Lie algebra to the Lie group
   The converse is more restrictive.

Theorem 1821 *If (E,f) is a representation of the Lie algebra T_1G of a connected finite dimensional Lie group, $\widetilde{G}$ a universal covering group of G with the smooth morphism :$\pi : \widetilde{G} \to G$, there is a smooth Lie group morphism $F \in C_\infty\left(\widetilde{G}; G\mathcal{L}\left(E;E\right)\right)$ such that F'(1)=f and (E,F) is a representation of $\widetilde{G}$, (E,F $\circ$ π) is a representation of G.*

Proof. G and $\widetilde{G}$ have the same Lie algebra, f is a Lie algebra morphism $T_1\widetilde{G} \to \mathcal{L}\left(E;E\right)$ which can be extended globally to $\widetilde{G}$ because it is simply connected. As a product of Lie group morphisms $F \circ \pi$ is still a smooth morphism in $C_\infty\left(G; G\mathcal{L}\left(E;E\right)\right)$ ■
   F can be computed from the exponential : $\forall u \in T_1G : \exp f'(1)u = F(\exp_{\widetilde{G}} u) = F\left(\widetilde{g}\right)$.

Weyl's unitary trick

(Knapp p.444)

It allows to go from different representations involving a semi simple Lie group. The context is the following :

Let G be a semi simple, finite dimensional, real Lie group. This is the case if G is a group of matrices closed under negative conjugate transpose.There is a Cartan decomposition : $T_1G = l_0 \oplus p_0$. If $l_0 \cap ip_0 = 0$ the real Lie algebra $u_0 = l_0 \oplus ip_0$ is a compact real form of the complexified $(T_1G)_\mathbb{C}$. So there is a compact, simply connected real Lie group U with Lie algebra u_0.

Assume that there is a complex Lie group $G_\mathbb{C}$ with Lie algebra $(T_1G)_\mathbb{C}$ which is the complexified of G. Then $G_\mathbb{C}$ is simply connected, semi simple and G,U are Lie subgroup of $G_\mathbb{C}$.

We have the identities : $(T_1G)_\mathbb{C} = (T_1G) \oplus i(T_1G) = u_0 \oplus iu_0$

Then we have the trick :

1. If (E,f) is a complex representation of $G_\mathbb{C}$ we get real representations (E,f) of G,U by restriction to the subgroups.

2. If (E,f) is a representation of G we have a the representation (E,f'(1)) of T_1G or u_0.

3. If (E,f) is a representation of U we have a the representation (E,f'(1)) of T_1G or u_0

4. If (E,f) is a representation of T_1G or u_0 we have a the representation (E,f'(1)) of $(T_1G)_\mathbb{C}$

5. A representation (E,f) of $(T_1G)_\mathbb{C}$ lifts to a representation of $G_\mathbb{C}$

6. Moreover in all these steps the invariant subspaces and the equivalences of representations are preserved.

23.1.5   Universal envelopping algebra

Principle

Theorem 1822 *(Knapp p.216) : The representations (E,f) of the Lie algebra A are in bijective correspondance with the representations (E,F) of its universal envelopping algebra U(A) by : $f = F \circ \imath$ where $\imath : A \to U(A)$ is the canonical injection.*

If V is an invariant closed vector subspace in the representation (E,f) of the Lie algebra A, then V is invariant for (E,F)

So if (E,F) is irreducible iff (E,f) is irreducible.

Components expressions

If $(e_i)_{i \in I}$ is a basis of A then a basis of U(A) is given by monomials :

$(\imath(e_{i_1}))^{n_1} (\imath(e_{i_2}))^{n_2} .... (\imath(e_{i_p}))^{n_p} , i_1 < i_2... < i_p \in I, n_1, ...n_p \in \mathbb{N}$

and F reads :

$F\left(((e_{i_1}))^{n_1} ((e_{i_2}))^{n_2} .... ((e_{i_p}))^{n_p}\right) = (f((e_{i_1})))^{n_1} \circ (f((e_{i_2})))^{n_2} .... \circ (f((e_{i_p})))^{n_p}$

On the right hand side the powers are for the iterates of f.

$F(1_K) = Id_E \Rightarrow \forall k \in K : F(k)(U) = kU$

If the representation (E,f) is given by matrices $[f(X)]$ then F reads as a product of matrices :

$F\left((\imath(e_{i_1}))^{n_1}(\imath(e_{i_2}))^{n_2}....\left(\imath\left(e_{i_p}\right)\right)^{n_p}\right) = [f(\imath(e_{i_1}))]^{n_1}[f(\imath(e_{i_2}))]^{n_2}....\left[f\left(\imath\left(e_{i_p}\right)\right)\right]^{n_p}$

Casimir elements

The Casimir elements of U(A) are defined as :

$\Omega_r = \sum_{(i_1,...i_r)=1}^{n} Tr\left(\rho\left(e_{i_1} \cdot e_{i_2}... \cdot e_{i_r}\right)\right) \imath(E_{i_1})...\imath(E_{i_r}) \in U(A)$

where the basis $(E_i)_{i=1}^{n}$ is a basis of A such that $B(E_i, e_j) = \delta_{ij}$ with the Killing form B.

They do not depend on the choice of a basis and belongs to the center of U(A).

They have for images in the representation (E,F) of U(A) :

$F(\Omega_r) = \sum_{(i_1,...i_r)=1}^{n} Tr\left(f(e_{i_1}) \circ f(e_{i_2})... \circ f(e_{i_r})\right) f(E_{i_1})...f(E_{i_r}) \in \mathcal{L}(E;E)$

As $F(\Omega_r)$ commutes with each f(X), they acts by scalars in any irreducible representation of A which can be used to label the irreducible representations.

Infinitesimal character

If (E,F) is an irreducible representation of U(A) there is a function χ, called the infinitesimal character of the representation, such that : $\chi : Z(U(A)) \rightarrow K :: F(U) = \chi(U) Id_E$ where Z(U(A)) is the center of U(A).

U is in the center of U(A) iff $\forall X \in A : XU = UX$ or $\exp(ad(X))(U) = U$.

Hilbertian representations

U(A) is a Banach C*-algebra with the involution : $U^* = U^t$ such that : $\imath(X)^* = -\imath(X)$

If (H,f) is a representation of the Lie algebra over a Hilbert space H, then $\mathcal{L}(H;H)$ is a C*-algebra.

(H,f) is a representation of the Banach C*-algebra U(A) if $\forall U \in U(A)$: $F(U^*) = F(U)^*$ and this condition is met if : $f(X)^* = -f(X)$: the representation of A must be anti-hermitian.

23.1.6 Adjoint representations

Lie algebras

Theorem 1823 *For any Lie algebra A (A, ad) is a representation of A on itself.*

This representation is extended to representations $(U_n(A), f_n)$ of A on its universal enveloping algebra U(A):

$U_n(A)$ is the subspace of homogeneous elements of U(A) of order n

$f_n : A \rightarrow L(U_n(A); U_n(A)) :: f_n(X)u = Xu - uX$ is a Lie algebra morphism.

For n=1 we have the adjoint represention.

Theorem 1824 *(Knapp p.291) If A is a Banach algebra there is a representation (U(A),f) of the component of identity Int(A) of GL(A;A)*

If A is a Banach algebra, then $G\mathcal{L}$(A;A) is a Lie group with Lie algebra $\mathcal{L}$(A;A), and it is the same for its component of the identity Int(A). With any automorphism $g \in Int(A)$ and the canonical map : $\imath : A \to U(A)$ the map : $\imath \circ g : A \to U(A)$ is such that : $i \circ g(X) \imath \circ g(Y) - i \circ g(Y) \imath \circ g(X) = \imath \circ g[X,Y]$ and $\imath \circ g$ can be extended uniquely to an algebra morphism $f(g)$ such that : $f(g) : U(A) \to U(A) : \imath \circ g = f(g) \circ \imath$ .Each f(g)$\in G\mathcal{L}(U(A);U(A))$ is an algebra automorphism of U(A) and each $U_n(A)$ is invariant.

The map : $f : Int(A) \to \mathcal{L}(U(A);U(A))$ is smooth and we have : $f(g) \circ f(h) = f(g \circ h)$ so (U(A),f) is a representation of Int(A).

Lie groups

Theorem 1825 *For any Lie group G the **adjoint representation** is the representation (T_1G, Ad) of G on its Lie algebra*

The map : $Ad : G \to G\mathcal{L}(T_1G; T_1G)$ is a smooth Lie group homorphism

This representation is not necessarily faithful.

It is irreducible iff G has no normal subgroup other than 1. The adjoint representation is faithful for simple Lie groups but not for semi-simple Lie groups.

It can be extended to a representation on the universal envelopping algebra $U(T_1G)$.

There is a representation (U(A),f) of the component of identity Int(A) of $G\mathcal{L}$(A;A)). $Ad_g \in Int(T_1G)$ so it gives a family of representations $(U_n(T_1G), Ad)$ of G on the universal envelopping algebra.

23.1.7 Unitary and orthogonal representations

Definition

Unitary or orthogonal representations are considered when there is some scalar product on E. It has been required that E is a Banach space, so the norm on E should come from the scalar product, and E must be a real or complex Hilbert space H, finite or infinite dimensional. So we will assume that H is a complex Hilbert space (the definitions and results are easily adjusted for the real case) with scalar product $\langle \rangle$, antilinear in the first variable.

Each operator X in $\mathcal{L}$(H;H) (or at least defined on a dense domain of H) has an adjoint X* in $\mathcal{L}$(H;H) such that :

$\langle Xu, v \rangle = \langle u, X^*v \rangle$

The map *:$\mathcal{L}$(H;H)→$\mathcal{L}$(H;H) is an involution, antilinear, bijective, continuous, isometric and if X is invertible, then X* is invertible and $\left(X^{-1}\right)^* = \left(X^*\right)^{-1}$ .With this involution $\mathcal{L}$(H;H) is a C*-algebra.

Definition 1826 *A **unitary representation** (H,f) of a group G is a representation on a Hilbert space H such that $\forall g \in G : f(g)^* f(g) = f(g) f(g)^* = I$*

If H is finite dimensional then f(g) is represented *in a Hilbert basis* by a unitary matrix. In the real case it is represented by an orthogonal matrix.

This is equivalent to : f preserves the scalar product : $\forall g \in G, u, v \in H$: $\langle f(g) u, f(g) v \rangle = \langle u, v \rangle$

Remark : if f there is a dense subsace E of H such that :$\forall u, v \in E$ the map $G \to K :: \langle u, f(g) v \rangle$ is continuous then f is continuous.

Sum of unitary representations of a group

Theorem 1827 *(Neeb p.24) The Hilbert sum of the unitary representations* $(H_i, f_i)_{i \in I}$ *is a unitary representation* (H, f) *where :*
$H = \oplus_{i \in I} H_i$ *the Hilbert sum of the spaces*
$f : G \to \mathcal{L}(H; H) :: f\left(\sum_{i \in I} u_i\right) = \sum_{i \in I} f_i(u_i)$

This definition generalizes the sum for any set I for a hilbert space.

Representation of the Lie algebra

Theorem 1828 *If (H,f) is a unitary representation of the Lie group G, then* $(H, f'(1))$ *is an anti-hermitian representation of* $T_1 G$

Proof. $(H, f'(1))$ is a representation of $T_1 G$. The scalar product is a continuous form so it is differentiable and :
$\forall X \in T_1 G, u, v \in H : \langle f'(1)(X) u, v \rangle + \langle u, f'(1)(X) v \rangle = 0 \Leftrightarrow (f'(1)(X))^* = -f'(1)(X)$ ■
$(H, f'(1))$ is a representation of the C*-algebra U(A).

Dual representation

Theorem 1829 *If (H,f) is a unitary representation of the Lie group G, then there is a unitary representation* $(H', \widetilde{f})$ *of G*

Proof. The dual H' of H is also Hilbert. There is a continuous anti-isomorphism $\tau : H' \to H$ such that :
$\forall \lambda \in H', \forall u \in H : \langle \tau(\varphi), u \rangle = \varphi(u)$
$\widetilde{f}$ is defined by : $\widetilde{f}(g) \varphi = \tau^{-1}(f(g) \tau(\varphi)) \Leftrightarrow \widetilde{f}(g) = \tau^{-1} \circ f(g) \circ \tau$
Which is $\mathbb{C}$ linear. If (H,f) is unitary then $\left(H', \widetilde{f}\right)$ is unitary:
$\left\langle \widetilde{f}(g) \varphi, \widetilde{f}(g) \psi \right\rangle_{H'} = \left\langle \tau \circ \widetilde{f}(g) \varphi, \tau \circ \widetilde{f}(g) \psi \right\rangle_H = \langle f(g) \circ \tau \varphi, f(g) \circ \tau \psi \rangle_H = \langle \tau \varphi, \tau \psi \rangle_H = \langle \varphi, \psi \rangle_{H^*}$ ■
The dual representation is also called the contragredient representation.

23.2   Representation of Lie groups

The representation of a group is a rich mathematical structure, so there are many theorems which gives a good insight of what one could expect, and are useful tools when one deals with non specified representations.

23.2.1   Action of the group

A representation (E,f) of G can be seen as a left (or right) action of G on E:

$\rho : E \times G \to E :: \rho(u,g) = f(g)u$

$\lambda : G \times E \to E :: \lambda(g,u) = f(g)u$

Theorem 1830 *The action is smooth and proper*

Proof. As f:G$\to \mathcal{L}(E;E)$ is assumed to be continuous, the map $\phi : \mathcal{L}(E;E) \times E \to E$ is bilinear, continuous with norm 1, so $\lambda(g,u) = \phi(f(g),u)$ is continuous.

The set GxE has a trivial manifold structure, and group structure. This is a Lie group. The maps λ is a continuous Lie group morphism, so it is smooth and a diffeomorphism. The inverse if continuous, and λ is proper. ∎

An invariant vector space is the union of orbits.

The representation is irreducible iff the action is transitive.

The representation is faithful iff the action is effective.

The map : $\mathbb{R} \to G\mathcal{L}(E;E) :: f(\exp tX)$ is a diffeomorphism in a neighborhood of 0, thus f'(1) is inversible.

We have the identities :

$\forall g \in G, X \in T_1G :$

$f'(g) = f(g) \circ f'(1) \circ L'_{g^{-1}}g$

$f'(g)(R'_g 1)X = f'(1)(X) \circ f(g)$

$Ad_{f(g)}f'(1) = f'(1) Ad_g$

The fundamental vector fields are :

$\zeta_L : T_1G \to \mathcal{L}(E;E) :: \zeta_L(X) = f'(1)X$

23.2.2   Functional representations

A functional representation (E,f) is a representaion where E is a space of functions (or even maps). Then the action of the group is usually on the variable of the function. Functional representations are the paradigm of infinite dimensional representations of a group. They exist for any group, and there are "standard" functional representations which have nice properties.

Right and left representations

Definition 1831 *The **left representation** of a topological group G on a Hilbert space of maps $H \subset C(E;F)$ is defined, with a continuous left action λ of G on the topological space E by : $\Lambda : G \to \mathcal{L}(H;H) :: \Lambda(g)\varphi = \lambda^*_{g^{-1}}\varphi$ with $\lambda_{g^{-1}} = \lambda(g^{-1},.)$*

So G acts on the variable inside φ : $\Lambda\left(g\right)\varphi\left(x\right) = \varphi\left(\lambda\left(g^{-1}, x\right)\right)$

Proof. Λ is a morphism:

For g fixed in G consider the map : $H \to H :: \varphi\left(x\right) \to \varphi\left(\lambda\left(g^{-1}, x\right)\right)$

$\Lambda\left(gh\right)\varphi = \lambda^*_{(gh)^{-1}}\varphi = \left(\lambda_{h^{-1}} \circ \lambda_{g^{-1}}\right)^* \varphi = \left(\lambda^*_{g^{-1}} \circ \lambda^*_{h^{-1}}\right)\varphi = \left(\Lambda\left(g\right) \circ \Lambda\left(h\right)\right)\varphi$

$\Lambda\left(1\right)\varphi = \varphi$

$\Lambda\left(g^{-1}\right)\varphi = \lambda^*_g\varphi = \left(\lambda^*_{g^{-1}}\right)^{-1}\varphi$ ∎

We have similarly the **right representation** with a right action :

$H \to H :: \varphi\left(x\right) \to \varphi\left(\rho\left(x, g\right)\right)$

$P : G \to \mathcal{L}\left(H; H\right) :: P\left(g\right)\varphi = \rho^*_g\varphi$

$P\left(g\right)\varphi\left(x\right) = \varphi\left(\rho\left(x, g\right)\right)$

Remark : some authors call right the left representation and vice versa.

Theorem 1832 *If there is a finite Haar Radon measure μ on the topological group G any left representation is unitary*

Proof. as H is a Hilbert space there is a scalar product denoted $\langle\varphi, \psi\rangle$

$\forall g \in G, \varphi \in H : \Lambda\left(g\right)\varphi \in H$ so $\langle\Lambda\left(g\right)\varphi, \Lambda\left(g\right)\psi\rangle$ is well defined

$(\varphi, \psi) = \int_G \langle\Lambda\left(g\right)\varphi, \Lambda\left(g\right)\psi\rangle\mu$ is well defined and $<\infty$. This is a scalar product (it has all the properties of $\langle\rangle$) over H

It is invariant by the action of G, thus with this scalar product the representation is unitary. ∎

Theorem 1833 *A left representation (H, Λ) of a Lie group G on a Hilbert space of differentiables maps $H \subset C_1\left(M; F\right)$, with a differentiable left action λ of G on the manifold M, induces a representation $(H, \Lambda'\left(1\right))$ of the Lie algebra $T_1 G$ where $T_1 G$ acts by differential operators.*

Proof. $(H, \Lambda'\left(1\right))$ is a representation of $T_1 G$

By the differentiation of : $\Lambda\left(g\right)\varphi\left(x\right) = \varphi\left(\lambda\left(g^{-1}, x\right)\right)$

$\Lambda'\left(g\right)\varphi\left(x\right)|_{g=1} = \varphi'\left(\lambda\left(g^{-1}, x\right)\right)|_{g=1}\lambda'_g\left(g^{-1}, x\right)|_{g=1}\left(-R'_{g^{-1}}\left(1\right) \circ L'_{g^{-1}}\left(g\right)\right)|_{g=1}$

$\Lambda'\left(1\right)\varphi\left(x\right) = -\varphi'\left(x\right)\lambda'_g\left(1, x\right)$

$X \in T_1 G : \Lambda'\left(1\right)\varphi\left(x\right)X = -\varphi'\left(x\right)\lambda'_g\left(1, x\right)X$

$\Lambda'\left(1\right)\varphi\left(x\right)X$ is a local differential operator (x does not change) ∎

Similarly : $P'\left(1\right)\varphi\left(x\right) = \varphi'\left(x\right)\rho'_g\left(x, 1\right)$

It is usual to write these operators as :

$\Lambda'\left(1\right)\varphi\left(x\right)X = \frac{d}{dt}\varphi\left(\lambda\left(\left(\exp\left(-tX\right)\right), x\right)\right)|_{t=0}$

$P'\left(1\right)\varphi\left(x\right)X = \frac{d}{dt}\varphi\left(\rho\left(x, \exp\left(tX\right)\right)\right)|_{t=0}$

These representations can be extended to representations of the universal envelopping algebra $U\left(T_1 G\right)$. We have differential operators on H of any order. These operators have an algebra structure, isomorphic to $U\left(T_1 G\right)$.

Polynomial representations

If G is a set of matrices in K(n) and λ the action of G on $E = K^n$ associated to the standard representation of G, then for any Hilbert space of functions of n variables on K we have the left representation (H, Λ) :

$\Lambda : G \to \mathcal{L}(H; H) :: \Lambda(g) \varphi(x_1, x_2, ..x_n) = \varphi(y_1, ..y_n)$ with $[Y] = [g]^{-1}[X]$

The set $K_p[x_1, ...x_n]$ of polynomials of degree p with n variables over a field K has the structure of a finite dimensional vector space, which is a Hilbert vector space with a norm on K^{p+1}. Thus with $H = K_p[x_1, ...x_n]$ we have a finite dimensional left representation of G.

The tensorial product of two polynomial representations :

$(K_p[x_1, ...x_p], \Lambda_p), (K_q[y_1, ...y_q], \Lambda_q)$

is given by :

the tensorial product of the vector spaces, which is : $K_{p+q}[x_1, ...x_p, y_1, ...y_q]$ represented in the canonical basis as the product of the polynomials

the morphism : $(\Lambda_p \otimes \Lambda_q)(g)(\varphi_p(X) \otimes \varphi_q(Y)) = \varphi_p\left([g]^{-1}[X]\right) \varphi_q\left([g]^{-1}[Y]\right)$

Representations on L^p spaces

See Functional analysis for the properties of these spaces.

The direct application of the previous theorems gives :

Theorem 1834 *(Neeb p.45) If G is a topological group, E a topological locally compact space, $\lambda : G \times E \to E$ a continuous left action of G on E, μ a G invariant Radon measure on E, then the left representation $\left(L^2(E, \mu, \mathbb{C}), f\right)$ with $(g)\varphi(x) = \varphi\left(\lambda\left(g^{-1}, x\right)\right)$ is an unitary representation of G.*

Theorem 1835 *(Neeb p.49) For any locally compact, topological group G, left invariant Haar Radon measure μ_L on G*

*the **left regular representation** of G is $\left(L^2(G, \mu_L, \mathbb{C}), \Lambda\right)$ with : $\Lambda(g)\varphi(x) = \varphi\left(g^{-1}x\right)$*

*the **right regular representation** of G is $\left(L^2(G, \mu_R, \mathbb{C}), P\right)$ with : $P(g)\varphi(x) = \sqrt{\Delta(g)}\varphi(xg)$*

which are both unitary.

The left regular representation is injective.

So any locally compact, topological group has a least one faithful unitary representation (usually infinite dimensional).

Averaging

With a unitary representation (H,f) of a group it is possible to define a representation of the *-algebra $L^1(G, S, \mu, \mathbb{C})$ of integrable functions on G. So this is different from the previous cases where we build a representation of G itself on spaces of maps on any set E.

Theorem 1836 *(Knapp p.557, Neeb p.134,143) If (H,f) is a unitary representation of a locally compact topological group G endowed with a finite Radon*

Haar measure μ, and H a Hilbert space, then the map : $F : L^1(G, S, \mu, \mathbb{C}) \to \mathcal{L}(H; H) :: F(\varphi) = \int_G \varphi(g) f(g) \mu(g)$ *gives a representation (H,F) of the Banach *-algebra* $L^1(G, S, \mu, \mathbb{C})$ *with convolution as internal product. The repesentations (H,f),(H,F) have the same invariant subspaces, (H,F) is irreducible iff (H,f) is irreducible.*

Conversely for each non degenerate Banach *-algebra representation (H,F) of $L^1(G, S, \mu, \mathbb{C})$ *there is a unique unitary continuous representation (H,f) of G such that :* $f(g) F(\varphi) = F(\Lambda(g)\varphi)$ *where* Λ *is the left regular action :* $\Lambda(g)\varphi(x) = \varphi(g^{-1}x)$.

F is defined as follows :

For any $\varphi \in L^1(G, S, \mu, \mathbb{C})$ fixed, the map : $B : H \times H \to \mathbb{C} :: B(u, v) = \int_G \langle u, \varphi(g) f(g) v \rangle \mu$ is sesquilinear and bounded because $|B(u, v)| \leq \|f(g)\| \|u\| \|v\| \int_G |\varphi(g)| \mu$ and there is a unique map : $A \in \mathcal{L}(H; H) : \forall u, v \in H : B(u, v) = \langle u, Av \rangle$. We put $A = F(\varphi)$. It is linear continuous and $\|F(\varphi)\| \leq \|\varphi\|$

$F(\varphi) \in \mathcal{L}(H; H)$ and can be seen as the integral of f(g) "averaged" by φ.

For convolution see Integration in Lie groups.

F has the following properties :

$F(\varphi)^* = F(\varphi^*)$ with $\varphi^*(g) = \overline{\varphi(g^{-1})}$

$F(\varphi * \psi) = F(\varphi) \circ F(\psi)$

$\|F(\varphi)\|_{\mathcal{L}(H;H)} \leq \|\varphi\|_{L^1}$

$F(g)F(\varphi)(x) = F(f(gx))$

$F(\varphi) F(g)(x) = \Delta_G(g) F(f(xg))$

For the commutants : $\left(F\left(L^1(G, S, \mu, \mathbb{C}) \right) \right)' = (f(G))'$

Representations given by kernels

(Neeb p.97)

1. Let (H, Λ) be a left representation of the topological group G, with H a Hilbert space of functions on a topological space, valued in a field K and a left action $\lambda : G \times E \to E$

H can be defined uniquely by a definite positive kernel $N : E \times E \to K$. (see Hilbert spaces).

So let J be a map (which is a cocycle) : $J : G \times E \to K^E$ such that :

$J(gh, x) = J(g, x) J\left(h, \lambda\left(g^{-1}, x \right) \right)$

Then (H, f) with the morphism : $f(g)(\varphi)(x) = J(g, x) \varphi\left(\lambda\left(g^{-1}, x \right) \right)$ is a unitary representation of G iff :

$N(\lambda(g, x), \lambda(g, y)) = J(g, \lambda(g, x)) N(x, y) \overline{J(g, \lambda(g, y))}$

If J, N, λ are continuous, then the representation is continuous.

Any G invariant closed subspace $A \sqsubseteq H_N$ has for reproducing kernel P which satisfies :

$P(\lambda(g, x), \lambda(g, y)) = J(g, \lambda(g, x)) P(x, y) \overline{J(g, \lambda(g, y))}$

Remarks :

i) if N is G invariant then take J=1

ii) if N(x,x)$\neq 0$ by normalization $Q(x,y) = \frac{N(x,y)}{\sqrt{|N(x,x)||N(y,y)|}}, J_Q(g,x) = \frac{J(g,x)}{|J(g,x)|}$, we have an equivalent representation where all the maps of H_Q are valued in the circle T.

2. Example : the Heisenberg group Heis(H) has the continuous unitary representation on the Fock space given by :

$f(t,v)\varphi(u) = \exp\left(it + \langle u,v\rangle - \frac{1}{2}\langle v,v\rangle\right)\varphi(u-v)$

23.2.3 Irreducible representations

General theorems

The most important theorems are the following:

Theorem 1837 *Schur's lemna : An interwiner $\phi \in \mathcal{L}(E_1; E_2)$ of two irreducible representations $(E_1, f_1), (E_2, f_2)$ of a group G is either 0 or an isomorphism.*

Proof. From the theorems below:

ker ϕ is either 0,and then ϕ is injective, or E_1 and then $\phi = 0$

Im ϕ is either 0, and then $\phi = 0$, or E_2 and then ϕ is surjective

Thus ϕ is either 0 or bijective, and then the representations are isomorphic :

$\forall g \in G : f_1(g) = \phi^{-1} \circ f_2(g) \circ \phi$ ∎

Therefore for any two irreducible representations either they are not equivalent, or they are isomorphic, and we can define **classes of irreducible representations**. If a representation(E,f) is reducible, we can define the number of occurences of a given class j of irreducible representation, which is called the **multiplicity** d_j of the class of representations j in (E,f).

Theorem 1838 *If $(E, f_1), (E, f_2)$ are two irreducible equivalent representations of a Lie group G on the same complex space then $\exists\lambda \in \mathbb{C}$ and an interwiner $\phi = \lambda Id$*

Proof. There is a bijective interwiner ϕ because the representations are equivalent. The spectrum of $\phi \in GL(E; E)$ is a compact subset of $\mathbb{C}$ with at least a non zero element λ, thus $\phi - \lambda Id$ is not injective in $\mathcal{L}$(E;E) but continuous, it is an interwiner of $(E, f_1), (E, f_2)$, thus it must be zero. ∎

Theorem 1839 *(Kolar p.131) If F is an invariant vector subspace in the finite dimensional representation (E,f) of a group G, then any tensorial product of (E,f) is completely reducible.*

Theorem 1840 *If (E,f) is a representation of the group G and F an invariant subspace, then : $\forall u \in F, \exists g \in G, v \in F : u = f(g)v$*

and $\left(E/F, \widehat{f}\right)$ is a representation of G, with : $\widehat{f} : G \to GL(E/F; E/F) ::$
$\widehat{f}(g)([u]) = [f(g)u]$

Proof. $\forall v \in F, \forall g \in G : f(g)v \in F \Rightarrow v = f(g)\left(f(g^{-1})v\right) = f(g)w$ with $w = \left(f(g^{-1})v\right)$

u$\sim v \Leftrightarrow u - v = w \in F \Rightarrow f(g)u - f(g)v = f(g)w \in F$

and if F is a closed vector subspace of E, and E a Banach, then E/F is still a Banach space. ■

Theorem 1841 *If $\phi \in \mathcal{L}(E_1; E_2)$ is an interwiner of the representations $(E_1, f_1), (E_2, f_2)$ of a group G then : $\ker \phi, \operatorname{Im} \phi$ are invariant subspaces of E_1, E_2 respectively*

Proof. $\forall g \in G : \phi \circ f_1(g) = f_2(g) \circ \phi$

$u \in \ker \phi \Rightarrow \phi \circ f_1(g)u = f_2(g) \circ \phi u = 0 \Rightarrow f_1(g)u \in \ker \phi \Leftrightarrow \ker \phi$ is invariant for f_1

$v \in \operatorname{Im} \phi \Rightarrow \exists u \in E_1 : v = \phi u \Rightarrow \phi(f_1(g)u) = f_2(g) \circ \phi u = f_2(g)v \Rightarrow f_2(g)v \in \operatorname{Im} \phi \Leftrightarrow \operatorname{Im} \phi$ is invariant for f_2 ■

Theorems for unitary representations

The most important results are the following:

Theorem 1842 *(Neeb p.77) If (H,f) is a unitary representation of the topological group G, H_d the closed vector subspace generated by all the irreducible subrepresentations in (H,f), then :*

i) H_d is invariant by G, and (H_d ,f) is a unitary representation of G which is a direct sum of irreducible representations

ii) the orthogonal complement $H_d^{\perp}$ does not contain any irreducible representation.

So : a unitary representation of a topological group can be written as the direct sum (possibly infinite) of subrepresentations :

$H = (\oplus_j d_j H_j) \oplus H_c$

each H_j is a class of irreducible representation, and d_j their multiplicity in the representation (H,f)

H_c does not contain any irreducible representation.

the components are mutually orthogonal : $H_j \perp H_k$ for $j \neq k, H_j \perp H_c$

the representation (H,f) is completely reducible iff $H_c = 0$

Are completely reducible in this manner :

- the continuous unitary representations of a topological finite or compact group;

- the continuous unitary finite dimensional representations of a topological group

Moreover we have the important result for compact groups :

Theorem 1843 *(Knapp p.559) Any irreducible unitary representation of a compact group is finite dimensional. Any compact Lie group has a faithful finite dimensional representation, and thus is isomorphic to a closed group of matrices.*

Thus for a compact group any continuous unitary representation is completely reducible in the direct sum of orthogonal finite dimensional irreducible unitary representations.

The tensorial product of irreducible representations is not necessarily irreducible. But we have the following result :

Theorem 1844 *(Neeb p.91) If $(H_1, f_1), (H_2, f_2)$ are two irreducible unitary infinite dimensional representations of G, then $(H_1 \otimes H_2, f_1 \otimes f_2)$ is an irreducible representation of G.*

This is untrue if the representations are not unitary or infinite dimensional.

Theorem 1845 *(Neeb p.76) A unitary representation (H,f) of the topological group G on a Hilbert space over the field K is irreducible if the commutant S' of the subset S={f (g) , g \in G} of $\mathcal{L}$(H;H) is trivial : S'=KxId*

Theorem 1846 *If E is an invariant vector subspace in a unitary representation (H,f) of the topological group G, then its orthogonal complement $E^\intercal$ is still a closed invariant vector subspace.*

Proof. The orthogonal complement $E^\intercal$ is a closed vector subspace, and also a Hilbert space and $H = E \oplus E^\intercal$

Let be $u \in E, v \in E^\intercal$, then $\langle u, v \rangle = 0, \forall g \in G : f(g) u \in E$

$\langle f(g) u, v \rangle = 0 = \langle u, f(g)^* v \rangle = \langle u, f(g)^{-1} v \rangle = \langle u, f(g^{-1}) v \rangle \Rightarrow \forall g \in G :$ $f(g) u \in E^\perp$ ∎

Definition 1847 *A unitary representation (H,f) of the topological group G is **cyclic** if there is a vector u in H such that $F(u) = \{f(g)u, g \in G\}$ is dense in H.*

Theorem 1848 *(Neeb p.117) If (H,f,u),(H',f',u') are two continuous unitary cyclic representations of the topological group G there is a unitary interwining operator F with u'=F(u) iff $\forall g : \langle u, f(g) u \rangle_H = \langle u', f'(g) u' \rangle_{H'}$*

Theorem 1849 *If F is a closed invariant vector subspace in the unitary representation (H,f) of the topological group G, then each vector of F is cyclic in F, meaning that $\forall u \neq 0 \in F : F(u) = \{f(g)u, g \in G\}$ is dense in F*

Proof. Let S={f(g), g \in G} $\sqsubset GL(H;H)$. We have S=S* because f(g) is unitary, so f(g)*=f(g^{-1}) \in S.

F is a closed vector subspace in H, thus a Hilbert space, and is invariant by S. Thus (see Hilbert spaces) :

$\forall u \neq 0 \in F : F(u) = \{f(g)u, g \in G\}$ is dense in F and the orthogonal complement F'(u) of F(u) in F is 0. ∎

Theorem 1850 *(Neeb p.77) If $(H_1, f), (H_2, f)$ are two inequivalent irreducible subrepresentations of the unitary representation (H,f) of the topological group G, then $H_1 \perp H_2$.*

Theorem 1851 *(Neeb p.24) Any unitary representation (H,f) of a topological group G is equivalent to the Hilbert sum of mutually orthogonal cyclic subrepresentations:* $(H, f) = \oplus_{i \in I} (H_i, f|_{H_i})$

23.2.4 Character

Definition 1852 *The **character** of a finite dimensional representation (E,f) of the topological group G is the function :*
$$\chi_f : G \to K :: \chi_f(g) = Tr(f(g))$$

The trace of any endomorphism always exists if E is finite dimensional. If E is an infinite dimensional Hilbert space H there is another definition, but a unitary operator is never trace class, so the definition does not hold any more.

The character reads in any orthonormal basis : $\chi_f(g) = \sum_{i \in I} \langle e_i, f(g) e_i \rangle$

Properties for a unitary representation

Theorem 1853 *(Knapp p.242) The character χ of the unitary finite dimensional representation (H,f) of the group G has the following properties :*
$$\chi_f(1) = \dim E$$
$$\forall g, h \in G : \chi_f(ghg^{-1}) = \chi_f(h)$$
$$\chi_{f^*}(g) = \chi_f(g^{-1})$$

Theorem 1854 *(Knapp p.243) If $(H_1, f_1), (H_2, f_2)$ are unitary finite dimensional representations of the group G :*

For the sum $(E_1 \oplus E_2, f = f_1 \oplus f_2)$ of the representations : $\chi_f = \chi_{f_1} + \chi_{f_2}$

For the tensorial product $(E_1 \otimes E_2, f = f_1 \otimes f_2)$ of the representations : $\chi_{f_1 \otimes f_2} = \chi_{f_1} \chi_{f_2}$

If the two representations $(E_1, f_1), (E_2, f_2)$ are equivalent then : $\chi_{f_1} = \chi_{f_2}$

So if (H,f) is the direct sum of $(H_j, f_j)_{j=1}^p$: $\chi_f = \sum_{q=1}^r d_q \chi_{f_q}$ where χ_{f_q} is for a class of equivalent representations, and d_q, called the **multiplicity**, is the number of representations in the family $(H_j, f_j)_{j=1}^p$ which are equivalent to (H_q, f_q).

If G is a compact connected Lie group, then there is a maximal torus T and any element of G is conjugate to an element of T : $\forall g \in G, \exists x \in G, t \in T : g = xtx^{-1}$ thus : $\chi_f(g) = \chi_f(t)$. So all the characters of the representation can be obtained by taking the characters of a maximal torus.

Compact Lie groups

Theorem 1855 *Schur's orthogonality relations (Knapp p.239) : Let G be a compact Lie group, endowed with a Radon Haar measure μ.*

i) If the unitary finite dimensional representation (H,f) is irreducible :
$$\forall u_1, v_1, u_2, v_2 \in H : \int_G \langle u_1, f(g) v_1 \rangle \overline{\langle u_2, f(g) v_2 \rangle} \mu = \frac{1}{\dim H} \langle u_1, v_1 \rangle \langle u_2, v_2 \rangle$$

$\chi_f \in \mathcal{L}^2(G, \mu, \mathbb{C})$ *and* $\|\chi_f\| = 1$

ii) If $(H_1, f_1), (H_2, f_2)$ *are two inequivalent irreducible unitary finite dimensional representations of* G :

$\forall u_1, v_1 \in H_1, u_2, v_2 \in H_2 : \int_G \langle u_1, f(g) v_1 \rangle \overline{\langle u_2, f(g) v_2 \rangle} \mu = 0$

$\int_G \chi_{f_1} \overline{\chi_{f_2}} \mu = 0$

$\chi_{f_1} * \chi_{f_2} = 0$ *with the definition of convolution above.*

iii) If $(H_1, f_1), (H_2, f_2)$ *are two equivalent irreducible unitary finite dimensional representations of* G : $\chi_{f_1} * \chi_{f_2} = d_{f_1}^{-1} \chi_{f_1}$ *where* d_{f_1} *is the multiplicity of the class of representations of both* $(H_1, f_1), (H_2, f_2)$.

Theorem 1856 *Peter-Weyl theorem (Knapp p.245) : For a compact Lie group* G *the linear span of all matrix coefficients for all finite dimensional irreducible unitary representation of* G *is dense in* $L^2(G, S, \mu, \mathbb{C})$

If (H,f) is a unitary representation of G, u,v$\in$ H, a matrix coefficient is a map : $G \rightarrow C(G; \mathbb{C}) : m(g) = \langle u, f(g) v \rangle$

Thus if $(H_j, f_j)_{j=1}^r$ is a family of mutually orthogonal and inequivalent unitary, finite dimensional representations, take an orthonormal basis (ε_α) in each H_j and define : $\varphi_{j\alpha\beta}(g) = \langle \varepsilon_\alpha, f_j(g) \varepsilon_\beta \rangle$, then the family of functions $\left(\sqrt{d_j} \varphi_{j\alpha\beta} \right)_{j,\alpha,\beta}$ is an orthonormal basis of $L^2(G, S, \mu, \mathbb{C})$.

23.2.5 Abelian groups

Fundamental theorem

The main feature of representations of abelian groups is that irreducible representations are unidimensional. More precisely :

Theorem 1857 *(Neeb p.76) Any irreducible unitary representation of a topological abelian group is one dimensional.*

Theorem 1858 *(Duistermaat p.213) Any finite dimensional irreducible representation of an abelian group is one dimensional.*

A one dimensional representation (E,f) takes the form $v \in E : f(g) v = \chi(g) u$ where u is some vector and $\chi : G \rightarrow K :: \chi(g + h) = \chi(g)\chi(h), \chi(1) = 1$ is a homomorphism from G to the abelian group (K,x). The character is just $\chi_f = \chi$. Thus the classes of irreducible representations of an abelian group are given by the characters.

If the irreducible representation is unitary, then $\chi : G \rightarrow T$ where T is the torus $\{z \in \mathbb{C} : |z| = 1\}$ and $\chi \in \widehat{G}$ the Pontryagin dual of G.

Any unitary representation takes the form :

$\left(H = \left(\oplus_{\chi \in \widehat{G}} d_\chi e_\chi \right) \oplus H_c, f = \left(\oplus_{\chi \in \widehat{G}} d_\chi \chi e_\chi \right) \oplus f_c \right)$

where the vectors e_j are orthonormal, and orthogonal to H_c. The irreducible representations are given with their multiplicity by the e_χ, indexed by the characters, and H_c does not contain any irreducible representation. It can happen that H_c is non trivial, so a unitary representation of an abelian group is not necessarily completely reducible.

So the problem is to define precisely such sums of representations. Roughly we can see that a representation will be on the space generated by the eigenvectors of f. The spectral theory is well suited for that.

Fourier transform

If (G,+) is a locally compact, abelian, topological group we have more results. G is isomorphic to its bidual by the Gelf'and transform (see Group structure) :

$$\tau : G \times \widehat{G} \to T :: \tau(g, \chi) = \chi(g)$$

$$\tau_g : \widehat{G} \to T :: \tau_g(\chi) = \chi(g) \Rightarrow \tau_g \in \widehat{\widehat{G}} \simeq G$$

$$\widehat{\ } : G \to \widehat{\widehat{G}} \simeq G :: \widehat{g} = \tau_g$$

$$\forall \chi \in \widehat{G} : \widehat{g}(\chi) = \chi(g)$$

There is a Haar measure μ, which is both left and right invariant and is a Radon measure, so the spaces $L^p(G, \mu, \mathbb{C}), 1 \leq p \leq \infty$ are well defined (see Functional analysis) and are Banach vector spaces, L^2 is a Hilbert space.

Theorem 1859 *(Neeb p.150-163) If G is a locally compact, abelian, topological group, endowed with a Haar measure μ , $\widehat{G}$ its Pontryagin dual, then we define the **Fourier transform** of a function $\varphi \in L^1(G, \mu, \mathbb{C})$ by : $\widehat{\varphi}(\chi) = \int_G \overline{\chi(g)} \varphi(g) \mu$*

i) $\widehat{\varphi}$ is well defined, continuous and vanishes at infinity : $\widehat{\varphi} \in C_{0\nu}\left(\widehat{G}; \mathbb{C}\right)$

ii) the map : $F : \widehat{G} \to L^1(G, S, \mu, \mathbb{C})' :: F(\chi)(\varphi) = \widehat{\varphi}(\chi)$ is bijective

iii) If P is a regular spectral measure on $\widehat{G}$, valued in $\mathcal{L}(H;H)$ for a Hilbert space H, (H,f) is a unitary representation of G with : $f(g) = P \circ \widehat{g}$ where $\widehat{g} : G \to \widehat{G} :: \widehat{g}(\chi) = \chi(g)$

iv) Conversely, for any unitary continuous representation (H,f) of G there is a unique regular spectral measure P such that :

$$P : S \to \mathcal{L}(H; H) :: P(\chi) = P^2(\chi) = P^*(\chi)$$

$$f(g) = P(\widehat{g}) \text{ where } \widehat{g} : G \to \widehat{G} :: \widehat{g}(\chi) = \chi(g)$$

Moreover any operator in $\mathcal{L}(H;H)$ commutes with f iff it commutes with each $P(\chi)$.

v) $L^1(G, S, \mu, \mathbb{C})$ is a commutative C*-algebra with convolution product as internal operation

Remarks :

i) If we come back to the fundamentals (see Set theory), a map χ is a subset $\chi(G, T) = \{(g, \chi(g)), g \in G\}$ of $(G \times T)$

A regular spectral measure P on $\widehat{G}$ is defined on the $\sigma-$algebra $\sigma\left(\widehat{G}\right)$ of $\widehat{G}$

. So $\varpi \in \sigma\left(\widehat{G}\right) = \{(g, \chi(g))\} \subset G \times T$

f(g) is the set $\varpi \cap \{g\}$

ii) the theorem establishes a bijective correspondance between unitary representations of G and spectral measures on its dual.

For any fixed character χ, we have the projection : $\chi(g) P(\chi) : H \to H_\chi$ where H_χ is an eigenspace of f(g) with eigenvalue $\chi(g)$. For any unitary vector

u in H, $\langle u, P(\chi) u \rangle = \|P(\chi) u\|^2$ is a probability on $\widehat{G}$, which gives a weigth to each character χ, specific to f. Notice that the measure P does not need to be absolutely continuous : it can have discrete components.

Abelian Lie groups are isomorphic to the products of vector spaces and tori. It is useful to see what happens in each of these cases.

Compact abelian Groups

Theorem 1860 *(Neeb p.79) For any compact abelian group G there is a Haar measure μ such that $\mu(G) = 1$ and a family $(\chi_n)_{n \in \mathbb{N}}$ of functions which constitute a Hermitian basis of $L^2(G, \mathbb{C}, \mu)$ so that every function $\varphi \in L^2(G, \mathbb{C}, \mu)$ is represented by the **Fourier Series** : $\varphi = \sum_{n \in \mathbb{N}} \left(\int_G \overline{\chi_n(g)} \varphi(g) \mu \right) \chi_n$*

Proof. If G is compact (it is isomorphic to a torus) then it has a finite Haar measure μ, which can be normalized to $1 = \int_G \mu$. And any continuous unitary representation is completely reducible in the direct sum of orthogonal one dimensional irreducible representations.

$\widehat{G} \subset C_{0b}(G; \mathbb{C})$ so $\forall p : 1 \le p \le \infty : \widehat{G} \subset L^p(G, \mathbb{C}, \mu)$

Moreover $\widehat{G}$ is an orthonormal subset of $L^2(G, \mathbb{C}, \mu)$ with the inner product $\langle \chi_1, \chi_2 \rangle = \int_G \overline{\chi_1(g)} \chi_2(g) \mu = \delta_{\chi_1, \chi_2}$

$\widehat{G}$ is discrete so : $\widehat{G} = \{\chi_n, n \in \mathbb{N}\}$ and we can take χ_n as a Hilbert basis of $L^2(G, \mathbb{C}, \mu)$

The Fourier transform reads : $\widehat{\varphi}_n = \int_G \overline{\chi_n(g)} \varphi(g) \mu$

$\varphi \in L^2(G, S, \mathbb{C}, \mu) : \varphi = \sum_{n \in \mathbb{N}} \langle \chi_n, \varphi \rangle \chi_n = \sum_{n \in \mathbb{N}} \left(\int_G \overline{\chi_n(g)} \varphi(g) \mu \right) \chi_n = \sum_{n \in \mathbb{N}} \widehat{\varphi}_n \chi_n$

which is the representation of φ by a Fourier series. ∎

For any unitary representation (H,f) of G, then the spectral measure is :

$P : S = \sigma(\mathbb{N}) \to \mathcal{L}(H; H) :: P_n = \int_G \overline{\chi_n(g)} f(g) \mu$

Finite dimensional vector space

Theorem 1861 *A unitary representation (H,f) of an abelian topological group G, isomorphic to a m dimensional vector space E reads :*

Theorem 1862 $f(g) = \int_{E^*} (\exp ip(g)) P(p)$ *where P is a spectral measure on the dual E*.*

Proof. There is an isomorphism between the dual E* of E and $\widehat{G}$:

$\Phi : E^* \to \widehat{G} :: \chi(g) = \exp ip(g)$

So any unitary representation (H,f) of G can be written :

$f(g) = \int_{E^*} (\exp ip(g)) P(p)$

where $P : \sigma(E^*) \to \mathcal{L}(H; H)$ is a spectral measure and $\sigma(E^*)$ the Borel $\sigma-$algebra of the dual E*, ∎

P is such that:

i) $\forall \varpi \in \sigma(E^*) : P(\varpi) = P(\varpi)^* = P(\varpi)^2 : P(\varpi)$ is a projection

ii) P(E*)=I

iii) $\forall u \in H$ the map : $\varpi \to \langle P(\varpi)u, u\rangle = \|P(\varpi)u\|^2 \in \mathbb{R}_+$ is a finite measure on (E*,$\sigma(E^*)$). Thus if $\langle u, u\rangle = 1$ $\|P(\varpi)u\|^2$ is a probability

If H is finite dimensional then the representation can be decomposed in a sum of orthogonal irreducible one dimensional representations and we have with a basis $(e_k)_{k=1}^n$ of H : $P(p) = \sum_{k=1}^n \pi_k(p)e_k$ where π_k is a measure on E* and $f(g) = \sum_{k=1}^n \left(\int_p (\exp ip(g))) \pi_k(p)\right) e_k$.

23.2.6   Induced representations

Induced representations are representations of a subgroup H of Lie group G which are extended to G.

Theorem 1863 *(Knapp p.564) A unitary representation (H,f) of a closed Lie subgroup S of a Lie group G endowed with a left invariant Haar measule ϖ_L , can be extended to a unitary representation (W, Λ_W) of G where W is a subset of $L^2(G; \varpi_L; S)$ and Λ_W the left regular representation on W.*

The set $L^2(G; \varpi_L; H)$ is a Hilbert space is H is separable.

W is the set of continuous maps φ in $L^2(G; \varpi_L; H)$ such that :

$W = \left\{ \varphi \in L^2(G; \varpi_L; H) \cap C_0(G; H) : \varphi(gs) = f\left(s^{-1}\right) \varphi(g) \right\}$

$\Lambda_W : G \to L(W; W) :: \Lambda_W(g)(\varphi)(g') = \varphi\left(g^{-1}g'\right)$

23.3   Representation of Lie algebras

Lie algebras are classified, therefore it is possible to exhibit almost all their representations, and this is the base for the classification of the representation of groups.

23.3.1   Irreducible representations

The results are similar to the results for groups.

Theorem 1864 *(Schur's lemna) Any interwiner $\phi \in \mathcal{L}(E_1; E_2)$ between the irreducible representations $(E_1, f_1), (E_2, f_2)$ of the Lie algebra A are either 0 or an isomorphism.*

Proof. with the use of the theorem below

$\ker \phi$ is either 0,and then ϕ is injective, or E_1 and then $\phi = 0$

$\text{Im} \phi$ is either 0, and then $\phi = 0$, or E_2 and then ϕ is surjective

Thus ϕ is either 0 or bijective, and then the representations are isomorphic :

$\forall X \in X : f_1(X) = \phi^{-1} \circ f_2(X) \circ \phi$ ∎

Theorem 1865 *If $(E, f_1), (E, f_2)$ are two irreducible equivalent representations of a Lie algebra A on the same complex space then $\exists \lambda \in \mathbb{C}$ and an interwiner $\phi = \lambda Id$*

Proof. The spectrum of $\phi \in GL(E;E)$ is a compact subset of $\mathbb{C}$ with at least a non zero element λ, thus $\phi - \lambda Id$ is not injective in $\mathcal{L}(E;E)$ but continuous, it is an interwiner of $(E, f_1), (E, f_2)$, thus it must be zero. ∎

Therefore for any two irreducible representations either they are not equivalent, or they are isomorphic, and we can define **classes of irreducible representations**. If a representation (E,f) is reducible, we can define the number of occurences of a given class j of irreducible representation, which is called the **multiplicity** d_j of the class of representations j in (E,f).

Theorem 1866 *(Knapp p.296) Any finite dimensional representation (E,f) of a complex semi-simple finite dimensional Lie algebra A is completely reducible:*

$E = \oplus_{k=1}^{p} E_k, (E_k, f|_{E_k})$ is an irreducible representation of A

Theorem 1867 *If $\phi \in \mathcal{L}(E_1; E_2)$ is an interwiner between the representations $(E_1, f_1), (E_2, f_2)$ of a Lie algebra A, then : $\ker \phi, \operatorname{Im} \phi$ are invariant subspaces of E_1, E_2 respectively*

Proof. $u \in \ker \phi \Rightarrow \phi \circ f_1(X) u = f_2(X) \circ \phi u = 0 \Rightarrow f_1(X) u \in \ker \phi \Leftrightarrow \ker \phi$ is invariant for f_1

$v \in \operatorname{Im} \phi \Rightarrow \exists u \in E_1 : v = \phi u \Rightarrow \phi(f_1(X) u) = f_2(X) \circ \phi u = f_2(X) v \Rightarrow f_2(X) v \in \operatorname{Im} \phi \Leftrightarrow \operatorname{Im} \phi$ is invariant for f_2 ∎

Theorem 1868 *(Knapp p.250) Any 1 dimensional representation of a semi simple Lie algebra is trivial (=0). Any 1-dimensional representation of a connected semi simple Lie group is trivial (=1).*

23.3.2 Classification of representations

Any finite dimensional Lie algebra has a representation as a matrix space over a finite dimensional vector space. All finite dimensional Lie algebras are classified, according to the abstract roots system. In a similar way one can build any irreducible representation from such a system. The theory is very technical (see Knapp) and the construction is by itself of little practical use, because all common Lie algebras are classified and well documented with their representations. However it is useful as a way to classify the representations, and decompose reducible representations into irreducible representations. The procedure starts with semi-simple complex Lie algebras, into which any finite dimensional Lie algebra can be decomposed.

Representations of a complex semi-simple Lie algebra
   (Knapp p.278, Fulton p.202)
   Let A be a complex semi-simple n dimensional Lie algebra., B its Killing form and B* the form on the dual A*. There is a Cartan subalgebra h of dimension r, the rank of A. Let h* be its dual.

The key point is that in any representation (E,f) of A, for any element H of h, f(H) acts diagonally with eigen values which are linear functional of the

$H \in h$. As the root-space decomposition of the algebra is just the representation (A,ad) of A on itself, we have many similarities.

i) So there are forms $\lambda \in h^*$, called **weights**, and eigenspaces, called **weigth spaces**, denoted E_λ such that :

$E_\lambda = \{u \in E : \forall H \in h : f(H)u = \lambda(H)u\}$

we have similarly :

$A_\alpha = \{X \in A : \forall H \in h : ad(H)X = \alpha(H)X\}$

The set of weights is denoted $\Delta(\lambda) \in h^*$ as the set of roots $\Delta(\alpha)$

Whereas the A_α are one dimensional, the E_λ can have any dimension $\leq n$ called the multiplicity of the weight

ii) E is the direct sum of all the weight spaces :

$E = \oplus_\lambda E_\lambda$

we have on the other hand : $A = h \oplus_\lambda A_\lambda$ because 0 is a common eigen value for h.

iii) every weight λ is real valued on h_0 and algebraically integral in the meaning that :

$\forall \alpha \in \Delta : 2\frac{B^*(\lambda,\alpha)}{B^*(\alpha,\alpha)} \in \mathbb{Z}$

where $h_0 = \sum_{\alpha \in \Delta} k^\alpha H_\alpha$ is the real vector space generated by H_α the vectors of A, dual of each root α with respect to the Killing form : $\forall H \in h : B(H, H_\alpha) = \alpha(H)$

iv) for any weigth λ : $\forall \alpha \in \Delta : f(H_\alpha)E_\lambda \sqsubseteq E_{\lambda+\alpha}$

As seen previously it is possible to introduce an ordering *of the roots* and compute a simple system of roots : $\Pi(\alpha) = \Pi(\alpha_1, ...\alpha_l)$ and distinguish positive roots $\Delta^+(\alpha)$ and negative roots $\Delta^-(\alpha)$

Theorem 1869 *Theorem of the highest weigth* *: If (E,f) is an irreducible finite dimensional representation of a complex semi-simple n dimensional Lie algebra A then there is a unique vector $V \in E$, called the highest weigth vector, such that :*

i) $V \in E_\mu$ for some $\mu \in \Delta(\lambda)$ called the highest weigth

ii) E_μ is one dimensional

iii) up to a scalar, V is the only vector such that : $\forall \alpha \in \Delta^+(\alpha), \forall H \in A_\alpha :$ $f(H)V = 0$

Then :

i) successive applications of $\forall \beta \in \Delta^-(\alpha)$ to V generates E:

$E = Span\left(f(H_{\beta_1})f(H_{\beta_2})...f(H_{\beta_p})V, \beta_k \in \Delta^-(\alpha)\right)$

ii) all the weights λ of the representation are of the form : $\lambda = \mu - \sum_{k=1}^l n_k \alpha_k$ with $n_k \in \mathbb{N}$ and $|\lambda| \leq |\mu|$

iii) μ depends on the simple system $\Pi(\alpha)$ and not the ordering

Conversely :

Let A be a finite dimensional complex semi simple Lie algebra. We can use the root-space decomposition to get $\Delta(\alpha)$. We know that a weight for any representation is real valued on h_0 and algebraically integral, that is : $2\frac{B^*(\lambda,\alpha)}{B^*(\alpha,\alpha)} \in \mathbb{Z}$.

So choose an ordering on the roots, and on a simple system $\Pi\,(\alpha_1, ...\alpha_l)$ define the **fundamental weights** : $(w_i)_{i=1}^l$ by : $2\frac{B^*(w_i, \alpha_j)}{B^*(\alpha_i, \alpha_i)} = \delta_{ij}$. Then any highest weight will be of the form : $\mu = \sum_{k=1}^l n_k w_k$ with $n_k \in \mathbb{N}$

The converse of the previous theorem is that, for any choice of such highest weight, there is a unique irreducible representation, up to isomorphism.

The irreducible representation (E_i, f_i) related to a fundamental weight w_i is called a **fundamental representation**. The dimension p_i of E_i is not a parameter : it is fixed by the choice of w_i.

To build this representation the procedure, which is complicated, is, starting with any vector V which will be the highest weight vector, compute successively other vectors by successive applications of $f\,(\beta)$ and prove that we get a set of independant vectors which consequently generates E. The dimension p of E is not fixed before the process, it is a result. As we have noticed, the choice of the vector space itself is irrelevant with regard to the representation problem, what matters is the matrix of f in some basis of E.

From fundamental representations one can build other irreducible representations, using the following result

Theorem 1870 *(Knapp p.341) If $(E_1, f_1), (E_2, f_2)$ are irreducible representations of the same algebra A, associated to the highest weights μ_1, μ_2. then the tensorial product of the representations, $(E_1 \otimes E_2, f_1 \times f_2)$ is an irreducible representation of A, associated to the highest weight $\mu_1 + \mu_2$.*

Notation 1871 *(E_i, f_i) denotes in the following the fundamental representation corresponding to the fundamental weight w_i*

Practically

Any finite dimensional semi simple Lie algebra belongs to one the 4 general families, or one of the 5 exceptional algebras. And, for each of them, the fundamental weigths $(w_i)_{i=1}^l$ (expressed as linear combinations of the roots) and the corresponding fundamental representations (E_i, f_i), have been computed and are documented. They are given in the next subsection with all the necessary comments.

Any finite dimensional irreducible representation of a complex semi simple Lie algebra of rank l can be labelled by l integers $(n_i)_{i=1}^l$: $\Gamma_{n_1...n_l}$ identifies the representation given by the highest weight : $w = \sum_{k=1}^l n_k w_k$. It is given by the tensorial product of fundamental representations. As the vector spaces E_i are distinct, we have the isomorphism $E_1 \otimes E_2 \simeq E_2 \otimes E_1$ and we can collect together the tensorial products related to the same vector space. So the irreducible representation labelled by $w = \sum_{k=1}^l n_k w_k$ is :

$\Gamma_{n_1...n_l} = \left(\otimes_{i=1}^l \left(\otimes_{k=1}^{n_i} E_i \right), \times_{i=1}^l \left(\times_{k=1}^{n_i} f_i \right) \right)$

And any irreducible representation is of this kind.

Each E_i has its specific dimension, thus if we want an irreducible representation on a vector space with a given dimension n, we have usually to patch together several representations through tensorial products.

Any representation is a combination of irreducible representations, however the decomposition is not unique. When we have a direct sum of such irreducible representations, it is possible to find equivalent representations, with a different sum of irreducible representations. The coefficients involved in these decompositions are called the Clebsh-Jordan coefficients. There are softwares which manage most of the operations.

Representation of compact Lie algebras

Compact complex Lie algebras are abelian, so only real compact Lie algebras are concerned. A root-space decomposition on a compact real Lie algebra A can be done (see Compact Lie groups) by : choosing a maximal Cartan subalgebra t (for a Lie group it comes from a torus, which is abelian, so its Lie subalgebra is also abelian), taking the complexified $A_{\mathbb{C}}, t_{\mathbb{C}}$ of A and t , and the roots α are elements of $t_{\mathbb{C}}^*$ such that :

$A_\alpha = \{X \in A_C : \forall H \in t_C : ad(H)X = \alpha(H)X\}$

$A_C = t_C \oplus_\alpha A_\alpha$

For any $H \in t : \alpha(H) \in i\mathbb{R}$: the roots are purely imaginary.

If we have a finite dimensional representation (E,f) of A, then we have weights with the same properties as above (except that t replaces h): there are forms $\lambda \in t_C^*$, called weights, and eigenspaces, called weigth spaces, denoted E_λ such that :

$E_\lambda = \{u \in E : \forall H \in t_C : f(H)u = \lambda(H)u\}$

The set of weights is denoted $\Delta(\lambda) \in t_C^*$.

The theorem of the highest weight extends in the same terms. In addition the result stands for the irreducible representations of compact connected Lie group, which are in bijective correspondance with the representations of the highest weight of their algebra.

23.4   Representation of classical groups

23.4.1   General rules

We recall the main general results :

- any irreducible representation of an abelian group is unidimensional (see the dedicated subsection)

- any continuous unitary representation of a compact or a finite group is completely reducible in the direct sum of orthogonal finite dimensional irreducible unitary representations.

- any continuous unitary finite dimensional representation of a topological group is completely reducible

- any 1 dimensional representation of a semi simple Lie algebra is trivial (=0). Any 1-dimensional representation of a connected semi simple Lie group is trivial (=1).

- any topological, locally compact, group has a least one faithful unitary representation (usually infinite dimensional) : the left (right) regular representations on the spaces $L^2(G, \mu_L, \mathbb{C})$.

- any Lie group has the adjoint representations over its Lie algebra and its universal envelopping algebra

- any group of matrices in K(n) has the standard representation over K^n where the matrices act by multiplication the usual way.

- there is a bijective correspondance between representations of real Lie algebras (resp real groups) and its complexified. And one representation is irreducible iff the other is irreducible.

- any Lie algebra has the adjoint representations over itself and its universal envelopping algebra

- any finite dimensional Lie algebra has a representation as a matrix group over a finite dimensional vector space.

- the finite dimensional representations of finite dimensional semi-simple complex Lie algebras are computed from the fundamental representations, which are documented

- the finite dimensional representations of finite dimensional real compact Lie algebras are computed from the fundamental representations, which are documented.

- Whenever we have a representation (E_1, f_1) and $\phi : E_1 \to E_2$ is an isomorphism we have an equivalent representation (E_2, f_2) with $f_2(g) = \phi \circ f_1(g) \circ \phi^{-1}$. So for finite dimensional representations we can take $K^n = E$.

23.4.2 Finite groups

The cas of finite groups, meaning groups with a finite number of elements (which are not usual Lie groups) has not been studied so far. We denote #G the number of its elements (its cardinality).

Standard representation

Definition 1872 *The standard representation (E,f) of the finite group G is :*
E is any #G dimensional vector space (such as $K^{\#G}$) on any field K
$f : G \to L(E; E) :: f(g)e_h = e_{gh}$ with any basis of $E : (e_g)_{g \in G}$

$f(g) \left(\sum_{h \in G} x_h e_h \right) = \sum_{h \in G} x_h e_{gh}$
$f(1) = I,$
$f(gh) u = \sum_{k \in G} x_k e_{ghk} = \sum_{k \in G} x_k f(g) \circ f(h) e_k = f(g) \circ f(h)(u)$

Unitary representation

Theorem 1873 *For any representation (E,f) of the finite group G, and any hermitian sequilinear form $\langle \rangle$ on E, the representation (E,f) is unitary with the scalar product : $(u, v) = \frac{1}{\#G} \sum_{g \in G} \langle f(g) u, f(g) v \rangle$*

Endowed with the discrete topology G is a compact group. So we can use the general theorem :

Theorem 1874 *Any representation of the finite group G is completely reducible in the direct sum of orthogonal finite dimensional irreducible unitary representations.*

Theorem 1875 *(Kosmann p.35) The number N of irreducible representations of the finite group G is equal to the number of conjugacy classes of G*

So there is a family $(E_i, f_i)_{i=1}^N$ of irreducible representations from which is deduced any other representation of G and conversely a given representation can be reduced to a sum and tensorial products of thse irreducible representations.

Irreducible representations

The irreducible representations $(E_i, f_i)_{i=1}^N$ are deduced from the standard representation, in some ways. A class of conjugacy is a subset G_k of G such that all elements of G_k commute with each other, and the G_k form a partition of G. Any irreducible representation (E_i, f_i) of G gives an irreducible subrepresentation of each G_k which is necessarily one dimensional because G_k is abelian. A representation (E_i, f_i) is built by patching together these one dimensional representations. There are many examples of these irreducible representations for the permutations group (Kosmann, Fulton).

Characters

The characters $\chi_f : G \to \mathbb{C} :: \chi_f(g) = Tr(f(g))$ are well defined for any representation. They are represented as a set of # scalars.

They are the same for equivalent representations. Moreover for any two irreducible representations $(E_p, f_p), (E_q, f_q) : \frac{1}{\#G} \sum_{g \in G} \overline{\chi_{f_p}(g)} \chi_{f_q}(g) = \delta_{pq}$

The table of characters of G is built as the matrix : $\left[\chi_{f_p}(g_q)\right]_{q=1..N}^{p=1...N}$ where g_q is a representative of each class of conjugacy of G. It is an orthonormal system with the previous relations. So it is helpful in patching together the one dimensional representations of the class of conjugacy.

For any representation (E,f) which is the direct sum of $(E_i, f_i)_{i=1}^N$, each with a multiplicity d_j. : $\chi_f = \sum_{q \in I} d_q \chi_{f_q}, \chi_f(1) = \dim E$

Functionnal representations

Theorem 1876 *The set $\mathbb{C}^G$ of functions $\varphi : G \to \mathbb{C}$ on the finite group G can be identified with the vector space $\mathbb{C}^{\#G}$.*

Proof. Any map is fully defined by #G complex numbers : $\{\varphi(g) = a_g, g \in G\}$ so it is a vector space on $\mathbb{C}$ with dimension is #G ∎

The natural basis of $\mathbb{C}^G$ is :$(e_g)_{g \in G} :: e_g(h) = \delta_{gh}$

It is is orthonormal with the scalar product : $\langle \varphi, \psi \rangle = \sum_{g \in G} \overline{\varphi(h)} \psi(h) \mu = \sum_{g \in G} \frac{1}{\#G} \overline{\varphi(g)} \psi(g)$

The Haar measure over G has for $\sigma-$algebra the set 2^G of subsets of G and for values:

$\varpi \in 2^G : \mu(\varpi) = \frac{1}{\#G}\delta(g) :: \mu(\varpi) = 0$ if $g \notin \varpi, : \mu(\varpi) = \frac{1}{\#G}$ if $g \in \varpi$

The left regular representation $(\mathbb{C}^G, \Lambda) : \varphi(x) \to \Lambda(g)(\varphi)(x) = \varphi(g^{-1}x)$ is unitary , finite dimensional, and $\Lambda(g)(e_h) = e_{gh}$. The characters in this representation are : $\chi_\Lambda(g) = Tr(\Lambda(g)) = (\#G)\delta_{1g}$

This representation is reducible : it is the sum of all the irreducible representations $(E_p, f_p)_{p=1}^N$ of G, each with a multiplicty equal to its dimension :
$(\mathbb{C}^G, \Lambda) = \sum_{p=1}^N (\otimes^{\dim E_p}(\mathbb{C}^G), \otimes^{\dim E_p} f_p)$

23.4.3   Representations of GL(K,n)

GL(K,n) is not semi simple, not compact, not connected. Its irreducible representations are tensorial products of the standard representation.

GL$(\mathbb{C}, n), SL(\mathbb{C}, n)$ are respectively the complexified of GL$(\mathbb{R}, n), SL(\mathbb{R}, n)$ so the representations of the latter are restrictions of the representations of the former. For more details about the representations of SL$(\mathbb{R}, n)$ see Knapp 1986.

Representations of GL$(\mathbb{C}, n)$

Theorem 1877 *All finite dimensional irreducible representations of GL$(\mathbb{C}, n)$ are alternate tensorial product of $(\wedge^k\mathbb{C}^n, D_A^k f)$ of the standard representation.$(\mathbb{C}^n, f)$.*

$\dim \wedge^k\mathbb{C}^n = C_n^k.$
For k=n we have the one dimensional representation : $(\mathbb{C}, \det)$.
The infinite dimensional representations are functional representations

Representations of SL$(\mathbb{C},n)$
1. SL(K,n) are connected, semi-simple, not compact groups. $SL(\mathbb{C}, n)$ is simply connected, and simple for n>1.
2. The algebras $sl(n, \mathbb{C}), n \geq 1$ form the A_{n-1} family in the classification of semi-simple complex Lie algebras.
The fundamental weights are : $w_l = \sum_{k=1}^l e_k, 1 \leq l \leq n-1$

Theorem 1878 *(Knapp p.340, Fulton p.221) If n>2 the fundamental representation (E_l, f_l) for the fundamental weight w_l is the alternate tensorial product $(\Lambda^l\mathbb{C}^n, D_A^l f)$ of the standard representation $(\mathbb{C}^n, f)$. The $\mathbb{C}-$dimension of $\Lambda^l\mathbb{C}^n$ is C_n^l*

The irreducible finite dimensional representations are then tensorial products of the fundamental representations.

Reprresentations of SL$(\mathbb{C}, 2)$
The representations of SL$(\mathbb{C}, 2)$ are of special importance (in many ways they are the "mother of all representations"). In short :
- for any n>0, Sl(C,2) has a unique (up to isomorphism) irreducible representation $(\mathbb{C}^n, f_n)$ on a complex n dimensional vector space, which is *not* unitary.

f_n is defined by computing its matrix on any n dimensional complex vector space E.

- SL(C,2) has several unitary representations which are not finite dimensional

Basis of sl(C,2)   The Lie algebra sl(C,2) is the algebra of 2x2 complex matrices with null trace. Its most usual basis in physics comprises the 3 matrices :

$$J_3 = \tfrac{1}{2}\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, J_+ = J_1 + iJ_2 = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix}, J_- = J_1 - iJ_2 = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix};$$

and we denote $J_\epsilon = J_+, J_-$ with $\epsilon = \pm 1$

So the bracket has the value : $[J_+, J_-] = 2J_3, [J_3, J_\epsilon] = \epsilon J_\varepsilon$

The exponential is not surjective in SL(C,2): $\forall g \in G : \exists X \in sl(C,2) : g = \exp X$ or $g = -\exp X$

Finite dimensional representation of sl(C,2) and SL(C,2)

(Kosmann)

1. Notational conventions

There are the conventions used in physics, and in fact, everywhere.

i) The representation of sl(C,2) on a n dimensional vector space E (which is unique) is labelled by j=n/2 which is either integer or half-integer. The representation on a 2j dimensional vector space is then denoted $\left(E^j, D^j\right)$.

ii) The 2j vectors of a basis of E^j are labelled by an integer m which runs from -j to +j by increments of 1 and are denoted $|j, m\rangle = e_m :$ | j,-j > , | j,-j+1 >,......| j,j-1 >, | j,j >

iii) The map f_{2j} is denoted D^j and given by the rules of the action of $D^j(X)$ on the vectors $|j, m\rangle$ for the vectors $X = J_3, J_+, J_-$:

$D^j(J_3)(|j, m\rangle) = m|j, m\rangle$

$D^j(J_\epsilon)(|j, m\rangle) = \sqrt{j(j+1) - m(m+\epsilon)}(|j, m+\epsilon\rangle)$

2. Casimir elements:

The elements of the representation of the universal envelopping algebra are expressed as products of matrices. The Casimir element $D^j(\Omega_2)$ is represented by the matrix $J^2 = J_1^2 + J_2^2 + J_3^2 = J_+J_- + J_3(J_3 - I) = J_-J_+ + J_3(J_3 + I)$. It acts by scalar :

$D^j(\Omega_2)|j, m\rangle = j(j+1)|j, m\rangle \Leftrightarrow \left[D^j(\Omega_2)\right] = j(j+1)[I]$

3. Sesquilinear form:

A sesquilinear form over E is defined by taking $|j, m\rangle$ as an orthonormal basis : $< mt, j||j, m\rangle = \delta_{mm'}$:

$\langle \sum_m x^m|j, m\rangle, \sum_m y^m|j, m\rangle \rangle = \sum_m \overline{x}^m y^m$

E^j becomes a Hilbert space, and the adjoint of an operator has for matrix in this basis the adjoint of the matrix of the operator : $\left[D^j(X)^*\right] = \left[D^j(X)\right]^*$ and we have : $\left[D^j(J_\epsilon)^*\right] = \left[D^j(J_\epsilon)\right]^* = \left[D^j(J_{-\epsilon})\right], \left[D^j(J_3)^*\right] = \left[D^j(J_3)\right], \left[D^j(\Omega_2)^*\right] = \left[D^j(\Omega_r)\right]$ so J_3, Ω_2 are hermitian operators.

4. Representations of SL(C,2):

Theorem 1879 *Any finite dimensional representation of the Lie algebra sl(C,2) lifts to a representation of the group SL(C,2) and can be computed by the exponential of matrices.*

Proof. Any $g \in SL(\mathbb{C}, 2)$ can be written as : $g = \epsilon \exp X$ for a unique $X \in sl(\mathbb{C}, 2), \epsilon = \pm 1$

Take : $\Phi(g(t)) = \epsilon \exp(tD^j(X))$

$\frac{\partial g}{\partial t}|_{t=0} = \Phi'(g)|_{g=1} = \epsilon D^j(X)$

so, as SL(C,2) is simply connected $\left(E, \epsilon \exp\left(D^j(X)\right)\right)$ is a representation of SL(C,2).

As the vector spaces are finite dimensional, the exponential of the morphisms can be computed as exponential of matrices. ∎

As the computation of these exponential is complicated, the finite dimensional representations of SL(C,2) are obtained more easily as functional representations on spaces of polynomials (see below).

Infinite dimensional representations of SL(C,2) (Knapp 1986 p. 31) The only irreducible representations (other than the trivial one) (H,f) of SL(C,2) are the following :

i) The principal unitary series :

$H = L^2(\mathbb{C}, \mu = dxdy)$: that is the space of maps $\varphi : \mathbb{C} \to \mathbb{C}$ such that $\int_{\mathbb{R}^2} |\varphi(x + iy)|^2 dxdy < \infty$

the morphisms f are parametrized by two scalars $(k, v) \in (\mathbb{Z}, \mathbb{R})$

$f_{k,v}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) \varphi(z) = -|-bz + d|^{-2-iv} \left(\frac{-bz+d}{|-bz+d|}\right)^{-k} \varphi\left(\frac{az-c}{-bz+d}\right)$

We have : $(H, f_{k,v}) \sim (H, f_{-k,-v})$

ii) The non unitary principal series :

$H = L^2\left(\mathbb{C}, \mu = \left(1 + |z|^2\right)^{\mathrm{Re}\, w} dxdy\right)$: that is the space of maps $\varphi : \mathbb{C} \to \mathbb{C}$

such that $\int_{\mathbb{R}^2} |\varphi(x + iy)|^2 \left(1 + |x + iy|^2\right)^{\mathrm{Re}\, w} dxdy < \infty$

the morphisms f are parametrized by two scalars $(k, v) \in (\mathbb{Z}, \mathbb{C})$

$f_{k,w}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) \varphi(z) = -|-bz + d|^{-2-w} \left(\frac{-bz+d}{|-bz+d|}\right)^{-k} \varphi\left(\frac{az-c}{-bz+d}\right)$

If w if purely imaginary we get back the previous series.

iii) the complementary unitary series :

$H = L^2(\mathbb{C}, \nu)$ maps $\varphi : \mathbb{C} \to \mathbb{C}$ with the scalar product : $\langle \varphi, \psi \rangle = \int_{\mathbb{C} \times \mathbb{C}} \frac{\overline{\varphi(z_1)\psi(z_2)}}{|z_1 - z_2|^{-2-w}} dz_1 dz_2 = \int_{\mathbb{C} \times \mathbb{C}} \overline{\varphi(z_1)} \psi(z_2) \nu$

the morphisms f are parametrized by two scalars $(k, w) \in (\mathbb{Z},]0, 2[\subset \mathbb{R})$

$f_{k,w}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) f(z) = -|-bz + d|^{-2-w} \left(\frac{-bz+d}{|-bz+d|}\right)^{-k} f\left(\frac{az-c}{-bz+d}\right)$

v) the non unitary principal series contain all the finite dimensional irreducible representations, by taking :

H= polynomial of degree m in z and n in $\overline{z}$

$f_{m,n}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) P(z) = (-bz + d)^m \overline{(-bz + d)}^n P\left(\frac{az-c}{-bz+d}\right)$

They are in bijective correspondance with the representations $\left(E^j, D^j\right)$ of sl(C,2)

23.4.4   Representations of SU(n)

SU(n) is a real Lie group, compact, not semi simple, connected and simply connected with Lie algebra su(n)={ X∈L($\mathbb{C}$,n):X+X*=0, TrX=0}. As SU(n) is compact, any unitary, any irreducible representation is finite dimensional.

The complexified of the Lie algebra $su\,(n)_{\mathbb{C}} = sl\,(\mathbb{C}, n)$ and the complexified $SU\,(n)_{\mathbb{C}} = SL\,(\mathbb{C}, n)$. So the representations (E,f) of SU(n) are in bijective correspondance with the representations of SL(C,n), by restriction of f to the subgroup SU(n) of SL(C,n). And the irreducible representations of SU(n) are the restrictions of the irreducible representations of SL(C,n). The same applies to the representations of the algebra su(n). So one finds the representations of SU(2) in the non unitary principal series of SL(C,2).

Representations of SU(2)
(Kosmann)
1. Basis of su(2):
We must restrict the actions of the elements of sl(C,2) to elements of su(2). The previous basis (J) of sl(C,2) is not a basis of su(2), so it is more convenient to take the matrices :

$K_1 = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = i\sigma_1; K_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = -i\sigma_2; K_3 = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} = i\sigma_2$

where σ_i are the Pauli's matrices (see Matrices). We add for convenience : $K_0 = I_{2\times2}$

$K_i^2 = -I, [K_1, K_2] = 2K_3, [K_2, K_3] = 2K_1; [K_3, K_1] = 2K_2$

The exponential is surjective on SU(2) so : $\forall g \in SU(2), \exists X \in su(2) : g = \exp X$

Using the specificities of the matrices it is easy to check that any matrix in SU(2) can be written as :

$\forall g \in SU(2) : [g] = \sum_{k=0}^{3} a_k [K_k]$ with $a_k \in \mathbb{R}, \sum_{k=0}^{3} |a_k|^2 = 1$

2. Finite dimensional representations :

All the finite dimensional representations of SL(C,2) stand in representations over polynomials. After adjusting to SU(2) (basically that $[g]^{-1} = [g]^*$) we have the following :

Theorem 1880 *The finite dimensional representations $\left(P^j, D^j\right)$ of SU(2) are the left regular representations over the degree 2j homogeneous polynomials with two complex variables z_1, z_2.*

P^j is a 2j+1 complex dimensional vector space with basis : $|j, m >= z_1^{j+m} z_2^{j-m}, -j \le m \le j$

$D^j\left(g\right) P\left(\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right) = P\left([g]^{-1} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right) = P\left([g]^* \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right)$

The representation is unitary with the scalar product over P^j defined by taking $|j, m>$ as orthonormal basis :

$\langle \sum_m x^m |j, m>, \sum_m y^m |j, m>\rangle = \sum_m \overline{x}^m y^m$

Any irreducible representation of SU(2) is equivalent to one of the $\left(P^j, D^j\right)$ for some $j \in \frac{1}{2}\mathbb{N}$.

3. Characters:

The characters of the representations can be obtained by taking the characters for a maximal torus which are of the kind:

$T(t) = \begin{bmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{bmatrix} \in SU(2)$

and we have :

$t \in]0, \pi[: \chi_{D^j}(T(t)) = \frac{\sin(2j+1)t}{\sin t}$

$\chi_{D^j}(T(0)) = 2j + 1 = \dim P^j$

$\chi_{D^j}(T(\pi)) = (-1)^{2j}(2j + 1)$

So if we take the tensorial product of two representations : $\left(P^{j_1} \otimes P^{j_2}, D^{j_1} \otimes D^{j_2}\right)$ we have :

$\chi_{j_1 \otimes j_2}(t) = \chi_{j_1}(t)\chi_{j_2}(t) = \chi_{|j_2-j_1|}(t) + \chi_{|j_2-j_1|+1}(t) + ...\chi_{j_2+j_1}(t)$

And the tensorial product is reducible in the sum of representations according to the **Clebsch-Jordan formula** :

$D^{j_1 \otimes j_2} = D^{|j_2-j_1|} \oplus D^{|j_2-j_1|+1} \oplus ... \oplus D^{j_2+j_1}$

The basis of $\left(P^{j_1} \otimes P^{j_2}, D^{j_1} \otimes D^{j_2}\right)$ is comprised of the vectors :

$|j_1, m_1> \otimes |j_2, m_2>$ with $, -j_1 \leq m_1 \leq j_1, , -j_2 \leq m_2 \leq j_2$

it can be decomposed in a sum of bases on each of these vector spaces :

$|J, M>$ with $|j_1 - j_2| \leq J \leq j_1 + j_2, -J \leq M \leq J$

So we have a matrix to go from one basis to the other :

$|J, M> = \sum_{m_1, m_2} C(J, M, j_1, j_2, m_1, m_2) |j_1, m_1> \otimes |j_2, m_2>$

whose coefficients are the **Clebsch-Jordan coefficients.** They are tabulated**.**

23.4.5   Representations of the Spin groups Spin(K,n)

The Pin and Spin groups are subsets of their respective Clifford algebras. Every Clifford algebra is isomorphic to an algebra of matrices over a field K', so a Clifford algebra has a faithful irreducible representation on a vector space over the field K', with the adequate dimension and orthonormal basis such that the representative of a Clifford member is one of these matrices. Then we get a representation of the Pin or Spin group by restriction of the representation of the Clifford algebra. In the Clifford algebra section the method to build the matrices algebra is given.

Notice that the dimension of the representation is given. Other representations can be deduced from there by tensorial products, but they are not necessarily irreducible.

$\text{Spin}(\mathbb{C}, n), Spin(\mathbb{R}, n)$ are double covers of $\text{SO}(\mathbb{C}, n), SO(\mathbb{R}, n)$, thus they have the same Lie algebra, which is compact. All their irreducible unitary representations are finite dimensional.

23.4.6 Representations of SO(K,n)

Theorem 1881 *Any irreducible representation of SO(K,n) is finite dimensional. The representations of SO(ℝ,n) can be deduced by restriction from the representations of SO(ℂ,n).*

Proof. SO(K,n) is semi-simple for n>2, compact, connected, not simply connected.

so($\mathbb{C}$,n) is the complexified of so($\mathbb{R}$,n), SO($\mathbb{C}$,n) is the complexified of SO($\mathbb{R}$,n)

As the groups are compact, all unitary irreducible representations are finite dimensional. ■

Theorem 1882 *(Knapp p.344) so(C,m) has two sets of fundamental weights and fundamental representations according to the parity of n:*

i) m odd : $m = 2n + 1 : so(C, 2n + 1), n \geq 2$ belongs to the B_n family

Fundamental weights : $w_l = \sum_{k=1}^{l} e_k, l \leq n - 1$ and $w_n = \frac{1}{2} \sum_{k=1}^{n} e_k$

The fundamental representation for l<n is the tensorial product $\left(\Lambda^l \mathbb{C}^{2n+1}, D_A^l f\right)$ of the standard representation ($\mathbb{C}^{2n+1}$,f) on orthonormal bases.. The $\mathbb{C}$−dimension of $\Lambda^l \mathbb{C}^{2n+1}$ is C_{2n+1}^l

The fundamental representation for w_n is the spin representation Γ_{2n+1}

ii) m even : $m = 2n : so(2n, C), n \geq 2$ belongs to the D_n family

Fundamental weights :

for $l \leq n - 2$: $w_l = \sum_{k=1}^{l} e_k, l < n$ and $w_{n-1} = \left(\frac{1}{2} \sum_{k=1}^{n-1} e_k\right) - e_n$ and $w_n = \left(\frac{1}{2} \sum_{k=1}^{n-1} e_k\right) + e_n$

The fundamental representation for $l < n$ is the tensorial product $\left(\Lambda^l \mathbb{C}^{2n}, D_A^l f\right)$ of the standard representation ($\mathbb{C}^{2n+1}$,f) on orthonormal bases. The $\mathbb{C}$−dimension of the representation is C_{2n}^l

The fundamental representation for w_n is the spin representation Γ_{2n}

The Spin representations deserve some comments :

i) Spin(K,n) and SO(K,n) have the same Lie algebra, isomorphic to so(K,n). So they share the same representations for their Lie algebras. These representations are computed by the usual method of roots, and the "spin representations" above correspond to specific value of the roots.

ii) At the group level the picture is different. The representations of the Spin groups are deduced from the representations of the Clifford algebra. For small size of the parameters isomorphims open the possibility to compute representations of the spin group by more direct methods.

iii) There is a Lie group morphism : $\pi : Spin(K, n) \rightarrow SO(K, n) :: \pi\left(\pm s\right) = g$

If (E,f) is a representation of SO(K,n), then (E,$f \circ \pi$) is a representation of Spin(K,n). Conversely a representation (E,$\widehat{f}$) of Spin(K,n) is a representation of SP(K,n) iff it meets the condition : $f \circ \pi\left(-s\right) = \widehat{f}\left(-s\right) = f \circ \pi\left(s\right) = \widehat{f}\left(s\right)$

$\widehat{f}$ must be such that $\widehat{f}\left(-s\right) = \widehat{f}\left(s\right)$, and we can find all the related representations of SO(K,n) among the representations of Spin(K,n) which have this symmetry.

The following example - one of the most important - explains the procedure.

Representations of SO($\mathbb{R}$,3)

1. Irreducible representations:

Theorem 1883 *(Kosmann) The only irreducible representations of SO($\mathbb{R}$, 3) are equivalent to the $\left(P^j, D^j\right)$ with $j \in \mathbb{N}$. The spin representation for j=1 is equivalent to the standard representation.*

Proof. To find all the representations of SO(R,3,1) we explore the representations of $Spin(\mathbb{R}, 3) \simeq SU(2)$. So we have to look among the representations $\left(P^j, D^j\right)$ and find which ones meet the condition above. It is easy to check that j must be an integer (so the half integer representations are excluded). ∎

2. Representations by harmonic functions:

Any matrix of SO(R,3) can be written as : $g = \exp\left(j(r)\right) = I_3 + j(r)\frac{\sin\sqrt{r^t r}}{\sqrt{r^t r}} + j(r)j(r)\frac{1-\cos\sqrt{r^t r}}{r^r r}$ where r is the vector of the rotation in $\mathbb{R}^3$ whose matrix is g in an orthonormal basis. So it is more convenient to relate the representations to functions in $\mathbb{R}^3$ than to polynomials of complex variables as it comes from SU(2).

The representation $\left(P^j, D^j\right)$ is equivalent to the representation :

i) Vector space : the homogeneous polynomials $\varphi\left(x_1, x_2, x_3\right)$ on $\mathbb{R}^3$ (meaning three real variables) with degree j with complex coefficients, which are harmonic, that is : $\Delta\varphi = 0$ where Δ is the laplacian $\Delta = \sum_{k=1}^3 \frac{\partial^2}{\partial x_k^2}$. This is a 2j+1 vector space, as P^j

ii) Morphism : the left regular representation Λ, meaning $f\left(g\right)\varphi\left(x\right) = \varphi\left([g]_{3\times3}^{-1}[x]_{3\times1}\right)$

So $f\left(\exp j(r)\right)\varphi\left(x\right) = \varphi\left(\left(I_3 - j(r)\frac{\sin\sqrt{r^t r}}{\sqrt{r^t r}} + j(r)j(r)\frac{1-\cos\sqrt{r^t r}}{r^r r}\right)[x]\right)$

3. Representation by spherical harmonics:

In order to have a unitary representation we need a Hilbert space. Harmonic functions are fully defined by their value on any sphere centered at 0. Let $L^2\left(S^2, \sigma, \mathbb{C}\right)$ the space of function defined over the shere S² of $\mathbb{R}^3$ and valued in $\mathbb{C}$, endowed with the sesquilinear form : $\langle\varphi, \psi\rangle = \int_{S^2}\overline{\varphi\left(x\right)}\psi\left(x\right)\sigma\left(x\right)$ where σ is the measure induced on S² by the Lebesgue measure of $\mathbb{R}^3$.

This is a Hilbert space, and the set of harmonic homogeneous polynomials of degree j on $\mathbb{R}^3$ is a closed vector subspace, so a Hilbert space denoted H^j. Thus $\left(H^j, \Lambda\right)$ is a unitary irreducible representation of SO(R,3) equivalent to $\left(P^j, D^j\right)$.

$L^2\left(S^2, \sigma, \mathbb{C}\right)$ is the Hilbert sum of the H^j : $L^2\left(S^2, \sigma, \mathbb{C}\right) = \oplus_{j\in\mathbb{N}}H^j$

By derivation we get a representations of the Lie algebra so(R,3) on the same Hilbert spaces H^j and the elements of so(R,3) are represented by antihermitian differential operators on H^j :

$D\Lambda\left(j\left(r\right)\right) = j\left(r\right)\left[\frac{\partial}{\partial x_i}\right]$

The Casimir operator $D\Lambda(\Omega_2) = -\Delta_{S^2}$ where Δ_{S^2} is the spherical laplacian defined on S^2. Its spectrum is discrete, with values j(j+1).

As we stay on S^2 it is convenient to use spherical coordinates :

$x_1 = \rho\sin\theta\cos\phi, x, x_2 = \rho\sin\theta\sin\phi, x_3 = \rho\cos\phi$

and on S^2 : $\rho = 1$ so $\varphi \in H^j : \varphi(\theta,\phi)$

A Hilbert basis of $L^2(S^2,\sigma,\mathbb{C})$ (thus orthonormal) comprises the vectors, called the **spherical harmonics** :

$j \in \mathbb{N}, -j \le m \le +j : |j,m> = Y_m^j(\theta,\phi)$

$m \ge 0 : Y_m^j(\theta,\phi) = C_m^j Z_m^j(\theta)\, e^{im\phi}, Z_m^j(\theta) = (\sin^m\theta)\, Q_m^j(\cos\theta)\,; Q_m^j(z) = \frac{d^{j+m}}{dz^{j+m}}(1-z^2)^j\,; C_m^j = \frac{(-1)^{j+m}}{2^l l!}\sqrt{\frac{2j+1}{4\pi}}\sqrt{\frac{(j-m)!}{(j+m)!}}$

$m < 0 : Y_m^j = (-1)^m\, \overline{Y_{-m}^j}$

which are eigen vectors of $D\Lambda(\Omega_2)$ with the eigen value j(j+1)

23.4.7 Representations of SO($\mathbb{R}$,p,q)

The group $SO(\mathbb{C},p,q)$ is isomorphic to $SO(\mathbb{C},p+q)$, thus semi-simple for n>2, compact, connected, not simply connected.

$O(\mathbb{R},p,q)$ is semi-simple for n>2, has two connected components, not simply connected. The universal covering group of $SO_0(\mathbb{R},p,q)$ is Spin($\mathbb{R}$,p,q).

$SO(\mathbb{C},p+q)$ is the complexified of $SO(\mathbb{R},p,q)$.

If we take the complexified we are back to representations of SO(C,p+q), with its tensorial product of the standard representations and the Spin representations. The Spin representations are deduced as restrictions of representations of the Clifford algebra Cl(p,q). The representations of the connected components SO_0(R,p,q) are then selected through the double cover in a way similar to $SO(\mathbb{C},n)$: the representations of Spin(R,p,q) must be such that $\forall s \in Spin(\mathbb{R},p,q) : \widehat{f}(-s) = \widehat{f}(s)$

Representations of SO(R,3,1)

1. To find all the representations of SO(R,3) we explore the representations of Spin($\mathbb{R}$,3,1) $\simeq$ SL($\mathbb{C}$,2)

The double cover is expressed by the map :

$\Psi : SL(\mathbb{C},2) \to SO_0(\mathbb{R},3,1) :: \Psi(\pm(\exp i\phi(\operatorname{Im} z))\exp\phi(\operatorname{Re} z)) = (\exp J(\operatorname{Im} z))(\exp K(\operatorname{Re} z))$

where : $\phi : \mathbb{C}^3 \to sl(\mathbb{C},2) :: \phi(z) = \frac{1}{2}\begin{bmatrix} z_3 & z_1+iz_2 \\ z_1-iz_2 & -z_3 \end{bmatrix}$, J and K are presented in the Section Linear groups.

2. The irreducible representations of SL(C,2) have been given previously.

i) The principal unitary series :

$H = L^2(\mathbb{C},\mu = dxdy)$

$f_{k,v}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)\varphi(z) = -\,|-bz+d|^{-2-iv}\left(\frac{-bz+d}{|-bz+d|}\right)^{-k}\varphi\left(\frac{az-c}{-bz+d}\right)$

and here we need k even $\in \mathbb{Z}, v \in \mathbb{R}$

ii) The non unitary principal series :

$H = L^2\left(\mathbb{C},\mu = \left(1+|z|^2\right)^{\operatorname{Re} w} dxdy\right)$

$$f_{k,w}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)\varphi(z) = -\left|-bz+d\right|^{-2-w}\left(\frac{-bz+d}{|-bz+d|}\right)^{-k}\varphi\left(\frac{az-c}{-bz+d}\right)$$

and here we need k even $\in \mathbb{Z}, v \in \mathbb{R}$

iii) the complementary unitary series :

$H = L^2\left(\mathbb{C}, dxdy\right)$ with the scalar product : $\langle\varphi,\psi\rangle = \int_{\mathbb{C}\times\mathbb{C}} \frac{\overline{\varphi(z_1)}\psi(z_2)}{|z_1-z_2|^{-2-w}}dz_1dz_2$

$$f_{k,w}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)f(z) = -\left|-bz+d\right|^{-2-w}\left(\frac{-bz+d}{|-bz+d|}\right)^{-k}f\left(\frac{az-c}{-bz+d}\right)$$

and here we need k even $\in \mathbb{Z}, w \in]0,2[\subset \mathbb{R}$

iv) the non unitary principal series contain all the finite dimensional irreducible representations, by taking :

H= polynomial of degree m in z and n in $\overline{z}$ with m,n even.

$$f_{m,n}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)P(z) = (-bz+d)^m\overline{(-bz+d)}^n P\left(\frac{az-c}{-bz+d}\right)$$

4. And we have for a representation (H,f) :

$f\left(\left(\exp J\left(\operatorname{Im} z\right)\right)\left(\exp K\left(\operatorname{Re} z\right)\right)\right) = f\left(\exp J\left(\operatorname{Im} z\right)\right)\circ f\left(\exp K\left(\operatorname{Re} z\right)\right)$

5. Another common way to find representations of SO(R,3,1) is to notice that so(R,3,1) is isomorphic to the direct product su(2)xsu(2) and from there the finite dimensional representations of SO(3,1) are the tensorial product of two irreducible representations of SU(2): $\left(P^{j_1}\otimes P^{j_2}, D^{j_1}\otimes D^{j_2}\right)$, which is then reducible to a sum of irreducible representations $\left(P^k, D^k\right)$ with the Clebsch-Jordan coefficients. This way we can use the well known tricks of the SU(2) representations, but the link with the generators of so(R,3,1) is less obvious.

23.4.8 Symplectic groups Sp(K,n)

$K=\mathbb{R}, \mathbb{C}$. They are semi-simple, connected, non compact group.

Their algebra belongs to the C_n family with the fundamental weights : $w_i = \sum_{k=1}^{i} e_k$

23.4.9 Affine transformations

1. An affine map d over an affine space E is the combination of a linear map g and a translation t. If g belongs to some linear group G then we have a group D of affine maps, which is the semi product of G and the abelian group $T\simeq\left(\overrightarrow{E},+\right)$ of translations : $:D = G \propto_\lambda T$, λ being the action of G on the vectors of $\overrightarrow{E} : \lambda : G\times\overrightarrow{E}\to\overrightarrow{E}$.

The composition law is : $(g,t)\times(g',t') = (gg', \lambda(g,t')+t), (g,t)^{-1} = \left(g^{-1}, -\lambda\left(g^{-1}, t\right)\right)$

Over finite dimensional space, in a basis, an affine group is characterized by a couple (A,B) of a square inversible matrix A belonging to a group of matrices G and a column matrix B, corresponding to the translation.

2. An affine group in a n dimensional affine space over the field K has a standard representation by (n+1)x(n+1) matrices over K as follows :

$$D = \begin{bmatrix} a_{11} & .. & a_{1n} & b_1 \\ .. & .. & .. & .. \\ a_{n1} & .. & a_{nn} & b_n \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & 1 \end{bmatrix}$$

and we can check that the composition of two affine maps is given by the product of the matrices. The inverse is :

$$D^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}B \\ 0 & 1 \end{bmatrix}$$

From there finding representations of groups come back to find representations of a group of matrices. Of special interest are the affine maps such that A belongs to a group of orthonormal matrices SO(K,n) or, in the Minkovski space, to SO(R,3,1) (this is the Poincaré group).

Part VI
PART 6 : FIBER BUNDLES AND JETS

The theory of fiber bundles is fairly recent - for mathematics - but in 40 or 50 years it has become an essential tool in many fields of research. ln mathematics it gives a unified and strong foundation for the study of whatever object one puts over a manifold, and thus it is the natural prolongation of differential geometry, notably whith the language of categories. In physics it is the starting point of all gauge theories. As usual we address first the general definitions, for a good understanding of the key principles, before the study of vector bundles, principal bundles and associated bundles. A fiber bundle E can be seen, at least locally and let us say in the practical calculii, as the product of two manifolds MxV, associated with a projection from E onto M. If V is a vector space we have a vector bundle, if V is a Lie group we have a principal bundle, and an associated fiber bundle is just a vector bundle where a preferred frame of reference has been identified. All that has been said previously about tensor bundle, vector and tensor fields can be generalized very simply to vector bundles using functors from the category of manifolds to the category of fibered manifolds.

However constructions above manifolds involve quite often differential operators, which link in some manner the base manifold and the objects upon it, for example connection on a tangent bundle. Objects involving partial derivatives on a manifold can be defined in the terms of jets : jets are just maps which have same derivatives at some order. Using these concepts we have a comprehensive body of tools to classify almost all non barbaric differential operators on manifolds, and using once more functors on manifolds one can show that they turn out to be combinations of well known operators.

One of these is connection, which is studied here in the more general context of fiber bundle, but with many similarities to the object defined over the tangent bundle in the previous part of this book.

In differential geometry we usually strive to get "intrinsic" results, meaning which are expressed by the same formulas whatever the charts. With jets, coordinates become the central focus, giving a precise and general study of all objects with derivatives content. So the combination of fiber bundle and jets enables to give a consistent definition of connections, matching the pure geometrical definition and the usual definition through Christoffel symbols.

519

24 FIBER BUNDLES

24.1 General fiber bundles

24.1.1 Fiber bundle

There are several generally accepted definitions of fiber bundles, with more or less strict requirements. We will stand with two definitions which encompasse all current usages, while offering the useful ingredients.

Fibered manifold

Definition 1884 *A **fibered manifold** E(M,π) is a triple of two Hausdorff manifolds E,M and a map $\pi : E \to M$ which is a surjective submersion. M is called the base space and π projects E on M.*

Theorem 1885 *(Giachetta p.7) E is a fibered manifold iff there is a manifold V, and E admits an atlas $(O_a, \varphi_a)_{a \in A}$ such that :*

$\varphi_a : O_a \to B_M \times B_V$ are two Banach vector spaces

$\varphi_a (x,.) : \pi (O_a) \to B_M$ is a chart ψ_a of M

and the transitions functions on $O_a \cap O_b$ are $\varphi_{ba} (\xi_a, \eta_a) = (\psi_{ba} (\xi_a), \phi (\xi_a, \eta_a)) = (\xi_b, \eta_b)$

Such a chart of E is said to be a chart adapted to the fibered structure. It is easier to see with the fiber bundle structure.

The coordinates on E are a pair (ξ_a, η_a), ξ corresponds to x.

Definition 1886 *A fibered manifold E'(N,π') is a **subbundle** of E(M,π) if N is a submanifold of M and the restriction $\pi_{E'} = \pi'$*

Theorem 1887 *A fibered manifold E(M,π) has the universal property : if f is a map f$\in C_r (M; P)$ in another manifold P then $f \circ \pi$ is class r iff f is class r (all the manifolds are assumed to be of class r).*

Fiber bundle

Definition 1888 *A **fiber bundle** denoted E(M,V,π) is a mathematical structure on a set E, comprises of :*

*i) three Hausdorff manifolds : E, the fibration (also called the **total space**), M which is the **base**, V which is the **standard fiber**, all over the same field K and the same class r (at least 1)*

*ii) a class r surjective submersion : $\pi : E \to M$ called the **projection***

iii) an open cover $(O_a)_{a \in A}$ of M and a set of diffeomorphisms

$\varphi_a : O_a \times V \subset M \times V \to \pi^{-1}(O_a) \subset E :: p = \varphi_a(x, u)$.

*The set of charts $(O_a, \varphi_a)_{a \in A}$ is called a **trivialization** (or fiber bundle atlas).*

*iv) a set of class r diffeomorphisms $(\varphi_{ab})_{a,b \in A}$ $\varphi_{ab} : (O_a \cap O_b) \times V \to V$ called the **transition maps**, such that :*

$\forall p \in \pi^{-1}(O_a \cap O_b), p = \varphi_a(x, u_a) = \varphi_b(x, u_b) \Rightarrow u_b = \varphi_{ba}(x, u_a)$

*and meeting the **cocycle conditions** :*

$\forall a, b, c \in A : \varphi_{aa}(x, u) = u; \varphi_{ab}(x, \varphi_{bc}(x, u)) = \varphi_{ac}(x, u)$

$$
\begin{array}{ccc}
\text{E} & & \\
\pi^{-1}(O_a) & & \\
\pi \downarrow & \nwarrow \varphi_a & \\
O_a & \to & O_a \times V \\
\text{M} & & \text{MxV}
\end{array}
$$

Notation 1889 $E(M, V, \pi)$ *is a fiber bundle with total space E, base M, standard fiber V, projection π*

The key points

1. A fiber bundle is a fibered manifold where V and a trivialization have been identified. As shown by the theorems below a fibered manifold is usually a fiber bundle, so there is not much difference between them, but the structure of fiber bundle is more explicit and therefore easier to understand. Fiber bundles are also called "locally trivial fibered manifolds with fiber V" (Hosemoller). We opt for simplicity.

2. For any p in E there is a *unique* x in M, but usually *several* u in V. And for the same x in E *all* u in V gives some point in E, which are identical. The set $\pi^{-1}(x) = E(x) \subset E$ is called the **fiber over x**.

3. The projection π is a surjective submersion so : $\pi'(p)$ is surjective, $\text{rank}\pi'(p) = \dim M$, π is an open map. For any x in M, the fiber $\pi^{-1}(x)$ is a submanifold of E.

4. Any product MxV=E is a fiber bundle, called a **trivial fiber bundle**. Then all the transition maps are the identity. But the converse is not true : usually E is built from pieces of MxV patched together through the trivializations. Anyway locally a fiber bundle is isomorphic to the product MxV, so it is useful to keep this model in mind.

5. If the manifolds are finite dimensional we have : $\dim E = \dim M + \dim V$ and $\dim \pi^{-1}(x) = \dim V$.

So If dimE=dimM then V is a manifold of dim0, meaning a discrete set. We will always assume that this is not the case.

6. The couples $(O_a, \varphi_a)_{a \in A}$ and the set $(\varphi_{ab})_{a,b \in A}$ play a role similar to the charts and transition maps for manifolds, but here they are defined between manifolds, without a Banach space.

The same fiber bundle can be defined by two different altlas $(O_a, \varphi_a)_{a \in A}, (Q_i, \psi_i)_{i \in I}$. Notice that the projection π is always the same. But they must meet some consistency conditions. Denote

$\tau_a : \pi^{-1}(O_a) \to V :: p = \varphi_a(\pi(p), \tau_a(p))$

$\theta_i : \pi^{-1}(Q_i) \to V :: p = \psi_i(\pi(p), \theta_i(p))$

$\forall p \in \pi^{-1}(O_a) \cap \pi^{-1}(Q_i) : \varphi_a(\pi(p), \tau_a(p)) = \psi_i(\pi(p), \theta_i(p))$

As φ_a, ψ_i are diffeomorphisms there are maps : $(\chi_{ia})_{(i,a) \in I \times A}$:

$\forall p \in \pi^{-1}(O_a) \cap \pi^{-1}(Q_i) : \theta_i(p) = \chi_{ia}(\pi(p), \tau_a(p))$

So these conditions read in a similar way than the transition conditions in an atlas.

When we want to check that some quantity defined on a manifold is "intrinsic", that it is does not depend on the choice of atlas, we check that it transforms according to the right formulas in a change of charts. Similarly to check that some quantity defined on a fiber bundle is "intrinsic" we check that it transforms according to the right formulas on the transition between two charts, and these conditions read the same as in a change of atlas.

7. Conversely, on the same manifold E different structures of fiber bundle can be defined.

8. The manifold V determines the kind of fiber bundle : if V is a vector space this is a vector bundle, if V is a Lie group this is a principal bundle.

9. Some practical advices :

i) it is common to define the trivializations as maps : $\varphi_a : \pi^{-1}(O_a) \to O_a \times V :: \varphi_a(p) = (x, u)$.

As the maps φ_a are diffeomorphism they give the same result. But from my personal experience the convention adopted here is more convenient.

ii) the indexes a,b in the transition maps : $p = \varphi_a(x, u_a) = \varphi_b(x, u_b) \Rightarrow u_b = \varphi_{ba}(x, u_a)$ are a constant hassle. Careful to stick as often as possible to some simple rules about the meaning of φ_{ba} : the order matters !

Example : the tangent bundle   Let M be a manifold with atlas $\left(B, (\Omega_i, \psi_i)_{i \in I}\right)$ then $TM(M, B, \pi) = \cup_{x \in M} \{u_x \in T_x M\}$ is a fiber bundle with base M, the canonial projection : $\pi : TM \to M :: \pi(u_x) = x$ and the trivializations : $(\Omega_i, \varphi_i)_{i \in I} : \varphi_i : O_i \times B \to TM :: (\psi_i'(x))^{-1} u = u_x$

The transition maps : $\varphi_{ji}(x) = \psi_j'(x) \circ (\psi_i'(x))^{-1}$ are linear and do not depend on u.

The tangent bundle is trivial iff it is parallelizable.

General theorems

Theorem 1890 *(Hosemoller p.15) A fibered manifold E(M,π) is a fiber bundle E(M,V,π) iff each fiber $\pi^{-1}(x)$ is diffeomorphic to V.*

Theorem 1891 *(Kolar p.77) If $\pi : E \to M$ is a proper surjective submersion from a manifold E to a connected manifold M, both real finite dimensional, then there is a structure of fiber bundle $E(M,V,\pi)$ for some V.*

Theorem 1892 *For any Hausdorff manifolds M,V, M connected, if there are an open cover $(O_a)_{a \in A}$ of M , a set of class r diffeomorphisms $(\varphi_{ab})_{a,b \in A}$ $\varphi_{ab} : (O_a \cap O_b) \times V \to V$, there is a fiber bundle structure $E(M,V,\pi)$.*

This shows that the transitions maps are an essential ingredient in the definition of fiber bundles.

Proof. Let X be the set : $X = \cup_{a \in A} O_a \times V$ and the equivalence relation :

$\Re : (x,u) \sim (x',u'), x \in O_a, x' \in O_b \Leftrightarrow x = x', u' = \varphi_{ba}(x,u)$

Then the fibration is $E = X/\Re$

The projection is the map : $\pi : E \to M :: \pi([x,u]) = x$

The trivializations are : $\varphi_a : O_a \times V \to E :: \varphi_a(x,u) = [x,u]$

and the cocycle conditions are met. ∎

Theorem 1893 *(Giachetta p.10) If M is reducible to a point, then any fiber bundle with basis M is trivial (but this is untrue for fibered manifold).*

Theorem 1894 *(Giachetta p.9) A fibered manifold whose fibers are diffemorphic either to a compact manifold or $\mathbb{R}^r$ is a fiber bundle.*

Theorem 1895 *(Giachetta p.10) Any finite dimensional fiber bundle admits a countable open cover whose each element has a compact closure.*

24.1.2 Sections

Fibered manifolds

Definition 1896 *A **section** on a fibered manifold $E(M,\pi)$ is a map $S : M \to E$ such that $\pi \circ S = Id$*

Fiber bundles

A *local section* on a fiber bundle $E(M,V,\pi)$ is a map : $S : O \to E$ with domain O$\subset M$

A *section* on a fiber bundle is defined by a map : $\sigma : M \to V$ such that $S = \varphi(x,\sigma(x))$. However we must solve the transition problem between two opens O_a, O_b. So our approach will be similar to the one used for tensor fields over a manifold : a section is defined by a family of maps $\sigma(x)$ with the condition that they define the same element S(x) at the transitions. Thus it will be fully consistent with what is done usually in differential geometry for tensor fields.

Definition 1897 *A section S on a fiber bundle $E(M,V,\pi)$ with trivialization $(O_a, \varphi_a)_{a \in A}$ is a family of maps $(\sigma_a)_{a \in A}, S_a : O_a \to V$ such that:*

$\forall a \in A, x \in O_a : S(x) = \varphi_a(x, \sigma_a(x))$

$\forall a, b \in A, O_a \cap O_b \neq \varnothing, \forall x \in O_a \cap O_b : \sigma_b(x) = \varphi_{ba}(x, \sigma_a(x))$

If the fiber bundle if of class r, a section is of class s$\leq r$ iff the maps σ_a are of class s.

Notation 1898 $\mathfrak{X}_r(E)$ *is the set of class r sections of the fiber bundle E*

A **global section on a fiber bundle** E(M,V,π) is a section defined by a single map : $\sigma : M \to V$ such that :

$\forall a \in A, x \in O_a : S_a(x) = \varphi_a(x, \sigma(x))$

On a fiber bundle we always have sections but not all fiber bundles have a global section. For a global section we must have : $\forall x \in O_a \cap O_b, \sigma(x) = \varphi_{ba}(x, \sigma(x))$

Example : On the tangent bundle $TM(M, B, \pi)$ over a manifold the sections are vector fields V(p) and the set of sections is $\mathfrak{X}(TM)$

24.1.3 Morphisms

Definition 1899 *A fibered manifold morphism or* **fibered map** *between two fibered manifolds* $E_1(M_1, \pi_1), E_2(M_2, \pi_2)$ *is a couple (F,f) of maps :* $F : E_1 \to E_2, f : M_1 \to M_2$ *such that :* $\pi_2 \circ F = f \circ \pi_1$.

The following diagram commutes :

$$
\begin{array}{ccccc}
E_1 & \to & \to & \to & E_2 \\
\downarrow & & F & & \downarrow \\
\downarrow & \pi_1 & & & \downarrow \quad \pi_2 \\
\downarrow & & & & \downarrow \\
M_1 & \to & \to & \to & M_2 \\
& & f & &
\end{array}
$$

Theorem 1900 *If (F,f) is fibered manifold morphism* $E_1(M_1, \pi_1) \to E_2(M_2, \pi_2)$ *and f is a local diffeomorphim, then if S_1 is a section on E_1 then $S_2 = F \circ S_1 \circ f^{-1}$ is a section on E_2 :*

Proof. $\pi_2 \circ S_2(x_2) = \pi_2 \circ F \circ S_1(x_1) = f \circ \pi_1 \circ S_1(x_1) = f(x_1) = x_2$ ∎

If (F,f) is an morphism on E, such that f is a diffeomorphism, then one can define the transport of a section on E :

$S \to \widetilde{S} : \widetilde{S}(y) = F\left(S\left(f^{-1}(y)\right)\right)$

Definition 1901 *A* **base preserving morphism** *is a morphism* $F : E_1 \to E_2$ *between two fibered manifolds* $E_1(M, \pi_1), E_2(M, \pi_2)$ *over the same base M such that $\pi_2 \circ F = \pi_1$ which means that f is the identity.*

Definition 1902 *A morphism (F,f) of fibered manifolds is injective if F,f are both injective, surjective if F,f are both surjective and is an isomorphism if F,f are both diffeomorphisms. Two fibered manifolds are* **isomorphic** *if there is a fibered map (F,f) such that F and f are diffeomorphisms.*

24.1.4 The category of fibered manifolds

The fiber bundle theory is the general framework to study all mathematical objects that are defined "over" a manifold (vector bundle, tensor bundles, connections, metric,..). And the language of categories is well suited for this : a mathematical object over a manifold is a functor from the category of manifolds to the category of fibered manifolds.

Definition 1903 *The **category of fibered manifolds** $\mathfrak{FM}$ over a field K has for objects fibered manifolds E, and for morphisms the morphisms of fibered manifolds denoted* $\hom(E_1, E_2)$

The category of class r fibered manifolds is the obvious subcategory.

In the language of categories : monomorphism = injective morphism, epimorphism = surjective morphism

Notation 1904 $\mathfrak{M}$ *is the category of manifolds over a field K with their morphisms (class r morphisms if necessary).*

Definition 1905 *The **base functor** $\mathfrak{B} : \mathfrak{FM} \mapsto \mathfrak{M}$ is the functor from the category category $\mathfrak{M}$ of manifolds to the category $\mathfrak{FM}$ of fibered manifolds which associates to each fibered manifold its base, and to each fibered morphism (F,f) the morphism f*

24.1.5 Product and sum of fiber bundles

Definition 1906 *The **product** of two fibered manifolds $E_1(M_1, \pi_1), E_2(M_2, \pi_2)$ is the fibered manifold $(E_1 \times E_2)(M_1 \times M_2, \pi_1 \times \pi_2)$*

The product of two fiber bundle $E_1(M_1, V_1, \pi_1), E_2(M_2, V_2, \pi_2)$ is the fibered manifold $(E_1 \times E_2)(M_1 \times M_2, V_1 \times V_2, \pi_1 \times \pi_2)$

Definition 1907 *The **Whitney sum** of two fibered manifolds $E_1(M, \pi_1), E_2(M, \pi_2)$ is the fibered manifold denoted $E_1 \oplus E_2$ where : M is the base, the total space is : $E_1 \oplus E_2 = \{(p_1, p_2) : \pi_1(p_1) = \pi_2(p_2)\}$, the projection $\pi(p_1, p_2) = \pi_1(p_1) = \pi_2(p_2)$*

The Whitney sum of two fiber bundles $E_1(M, V_1, \pi_1), E_2(M, V_2, \pi_2)$ is the fiber bundle denoted $E_1 \oplus E_2$ where : M is the base, the total space is : $E_1 \oplus E_2 = \{(p_1, p_2) : \pi_1(p_1) = \pi_2(p_2)\}$, the projection $\pi(p_1, p_2) = \pi_1(p_1) = \pi_2(p_2)$, the fiber is $V = V_1 \times V_2$ and the trivializations : $\varphi(x, (u_1, u_2)) = (p_1, p_2) = (\varphi_1(x, u_1), \varphi_2(x, u_2))$

If the fibers bundles are trivial then their Whitney sum is trivial.

24.1.6 Pull back

Also called induced bundle

Definition 1908 *The **pull back** f^*E of a fibered manifold $E(M, \pi)$ on a manifold N by a continuous map $f : N \to M$ is the fibered manifold $f^*E\,(N, \widetilde{\pi})$ with total space : $f^*E = \{(y, p) \in N \times E : f\,(y) = \pi\,(p)\}$,projection : $\widetilde{\pi} : f^*E \to N :: \widetilde{\pi}\,(y, p) = y$*

Definition 1909 *The pull back f^*E of a fiber bundle $E(M, V, \pi)$ with atlas $(O_a, \varphi_a)_{a \in A}$ on a manifold N by a continuous map $f : N \to M$ is the fiber bundle $f^*E\,(N, V, \widetilde{\pi})$ with total space : $f^*E = \{(y, p) \in N \times E : f\,(y) = \pi\,(p)\}$,projection : $\widetilde{\pi} : f^*E \to N :: \widetilde{\pi}\,(y, p) = y$, open cover : $f^{-1}\,(O_a)$, trivializations : $\widetilde{\varphi}_a : f^{-1}\,(O_a) \times V \to \widetilde{\pi}^{-1} \circ f^{-1}\,(O_a) :: \widetilde{\varphi}_a(y, u) = (y, \varphi_a(f(y), u))$*

$$
\begin{array}{ccc}
f^*E & \xrightarrow{\pi^* f} \to \to & E \\
\downarrow & & \downarrow \\
f^*\pi \downarrow & & \downarrow \pi \\
\downarrow & f & \downarrow \\
N & \to \to \to \to \to & M
\end{array}
$$

The projection $\widetilde{\pi}$ is an open map.

For any section $S : M \to E$ we have the pull back $f^*S : N \to E :: f^*S\,(y) = (f\,(y), S\,(f\,(y)))$

24.1.7 Tangent space over a fiber bundle

The key point is that the tangent space $T_p E$ of a fiber bundle is isomorphic to $T_x M \times T_u V$

Local charts

Theorem 1910 *The total space E of a fiber bundle $E(M, V, \pi)$ admits, as a manifold, an atlas $\left(B_M \times B_V, (O_a \times U_i, \Phi_{ai})_{(a,i) \in A \times I} \right)$ where*

$\left(B_M, (O_a, \psi_a)_{a \in A} \right)$ *is an atlas of the manifold M*
$\left(B_V, (U_i, \phi_i)_{i \in I} \right)$ *is an atlas of the manifold V*
$\Phi_{ai}\,(p) = (\psi_a \circ \pi\,(p), \phi_i \tau_a\,(p)) = (\xi_a, \eta_a)$

Proof. Let $(O_a, \varphi_a)_{a \in A}$ be an atlas of E as fiber bundle. $E(M, V, \pi)$

By taking a refinement of the open covers we can assume that $\left(B_M, (O_a, \psi_a)_{a \in A} \right)$ is an atlas of M

As the φ_a are diffeomorphisms the following maps are well defined :
$\tau_a : \pi^{-1}\,(O_a) \to V :: \varphi_a\,(x, \tau_a\,(p)) = p = \varphi_a\,(\pi\,(p), \tau_a\,(p))$
$\forall p \in \pi^{-1}\,(O_a) \cap \pi^{-1}\,(Q_i) : \tau_b\,(p) = \varphi_{ba}\,(\pi\,(p), \tau_a\,(p))$
and the φ_a are open maps so $\varphi_a\,(O_a, V_i)$ is an open cover of E.
The map :
$\Phi_{ai} : \varphi_a\,(O_a, V_i) \to B_M \times B_V :: \Phi_{ai}\,(p) = (\psi_a \circ \pi\,(p), \phi_i \tau_a\,(p)) = (\xi_a, \eta_a)$

is bijective and differentiable.

If $p \in \Omega_{ai} \cap \Omega_{bj} : \Phi_{bj}(p) = (\psi_b \circ \pi(p), \phi_j \circ \tau_b(p)) = (\xi_b, \eta_b)$

$\xi_b = \psi_b \circ \pi \circ \pi^{-1} \circ \psi_a^{-1}(\xi_a) = \psi_b \circ \psi_a^{-1}(\xi_a) = \psi_{ba}(\xi_a)$ ∎

Proof. $\eta_j = \phi_j \circ \tau_b \circ \tau_a^{-1} \circ \phi_i^{-1}(\eta_i) = \phi_j \circ \varphi_{ba}(\psi_a^{-1}(\xi_a), \phi_i^{-1}(\eta_i))$ ∎

Tangent space

Theorem 1911 *Any vector* $v_p \in T_p E$ *of the fiber bundle* $E(M, V, \pi)$ *has a unique decomposition :* $v_p = \varphi_a'(x, u_a)(v_x, v_{au})$ *where :* $v_x = \pi'(p)v_p \in T_{\pi(p)}M$ *does not depend on the trivialization and* $v_{au} = \tau_a'(p)v_p$

Proof. The differentiation of :

$p = \varphi_a(\pi(p), \tau_a(p)) \rightarrow v_p = \varphi_{ax}'(\pi(p), \tau_a(p))\pi'(p)v_p + \varphi_{au}'(\pi(p), \tau_a(p))\tau_a'(p)v_p$

$x = \pi(p) \rightarrow v_x = \pi'(p)v_p$

$p = \varphi_a(x, u_a) \rightarrow v_p = \varphi_{ax}'(x, u_a)v_{ax} + \varphi_{au}'(x, u_a)v_{au}$

$\Rightarrow v_{ax} = \pi'(p)v_p$ does not depend on the trivialization

$\varphi_{au}'(x, u_a)v_{au} = \varphi_{au}'(\pi(p), \tau_a(p))\tau_a'(p)v_p \Rightarrow v_{au} = \tau_a'(p)v_p$ ∎

Theorem 1912 *Any vector of* $T_p E$ *can be uniquely written:* $v_p = \sum_\alpha v_x^\alpha \partial x_\alpha + \sum_i v_{au}^i \partial u_i$ *with the basis, called a* ***holonomic*** *basis,*

$\partial x_\alpha = \varphi_{ax}'(x, u)\partial \xi_\alpha, \partial u_i = \varphi_{au}'(x, u)\partial \eta_i$ *where* $\partial \xi_\alpha, \partial \eta_i$ *are holonomic bases of* $T_x M, T_u V$

Proof. $v_x = \sum_\alpha v_x^\alpha \partial \xi_\alpha, v_{au} = \sum_i v_{au}^i \partial \eta_i$

$v_p = \varphi_{ax}'(x, u_a)v_x + \varphi_{au}'(x, u_a)v_{au} = \sum_\alpha v_x^\alpha \varphi_{ax}'(x, u)\partial \xi_\alpha + \sum_i v_{au}^i \varphi_{au}'(x, u)\partial \eta_i$ ∎

In the following :

Notation 1913 ∂x_α *(latine letter, greek indices) is the part of the basis on TE induced by M*

∂u_i *(latine letter, latine indices) is the part of the basis on TE induced by V.*

$\partial \xi_\alpha$ *(greek letter, greek indices) is a holonomic basis on TM.*

$\partial \eta_i$ *(greek letter, latine indices) is a holonomic basis on TV.*

$v_p = \sum_\alpha v_x^\alpha \partial x_\alpha + \sum_i v_u^i \partial u_i$ *is any vector* $v_p \in T_p E$

With this notation it is clear that a holonomic basis on TE split in a part related to M (the base) and V (the standard fiber)

Transition:

Usually the holonomic bases in M and V change also in a transition. So to keep it simple we can stay at the level of v_x, v_u

Theorem 1914 *At the transitions between charts :* $v_p = \varphi_a'(x, u_a)(v_x, v_{au}) = \varphi_b'(x, u_b)(v_x, v_{bu})$ *we have the identities :*

$\varphi_{ax}'(x, u_a) = \varphi_{bx}'(x, u_b) + \varphi_{bu}'(x, u_b)\varphi_{bax}'(x, u_a)$

$\varphi_{au}'(x, u_a) = \varphi_{bu}'(x, u_b)\varphi_{bau}'(x, u_a)$

$v_{bu} = \varphi_{ba}'(x, u_a)(v_x, v_{au})$

Proof. $p = \varphi_a\left(x, u_a\right) = \varphi_b\left(x, u_b\right) = \varphi_b\left(x, \varphi_{ba}\left(x, u_a\right)\right)$

The differentiation of :$\varphi_a\left(x, u_a\right) = \varphi_b\left(x, \varphi_{ba}\left(x, u_a\right)\right)$ with respect to u$_a$ gives
:

$\varphi'_{au}\left(x, u_a\right) v_{au} = \varphi'_{bu}\left(x, u_b\right) \varphi'_{bau}\left(x, u_a\right) v_{au}$

The differentiation of :$\varphi_a\left(x, u_a\right) = \varphi_b\left(x, \varphi_{ba}\left(x, u_a\right)\right)$ with respect to x gives
:

$\varphi'_{ax}\left(x, u_a\right) v_x = \left(\varphi'_{bx}\left(x, u_b\right) + \varphi'_{bu}\left(x, u_b\right) \varphi'_{bax}\left(x, u_a\right)\right) v_x$

$v_p = \varphi'_{ax}\left(x, u_a\right) v_x + \varphi'_{au}\left(x, u_a\right) v_{au} = \varphi'_{bx}\left(x, u_b\right) v_x + \varphi'_{bu}\left(x, u_b\right) v_{bu}$

$\left(\varphi'_{bx}\left(x, u_b\right) + \varphi'_{bu}\left(x, u_b\right) \varphi'_{bax}\left(x, u_a\right)\right) v_x + \varphi'_{bu}\left(x, u_b\right) \varphi'_{bau}\left(x, u_a\right) v_{au} = \varphi'_{bx}\left(x, u_b\right) v_x + \varphi'_{bu}\left(x, u_b\right) v_{bu}$

$\varphi'_{bu}\left(x, u_b\right) \varphi'_{bax}\left(x, u_a\right) v_x + \varphi'_{bu}\left(x, u_b\right) \varphi'_{bau}\left(x, u_a\right) v_{au} = \varphi'_{bu}\left(x, u_b\right) v_{bu}$

$v_{bu} = \varphi'_{bax}\left(x, u_a\right) v_x + \varphi'_{bau}\left(x, u_a\right) v_{au} = \varphi'_{ba}\left(x, u_a\right)\left(v_x, v_{au}\right)$ ∎

Theorem 1915 *If S is a section on E, then $S'\left(x\right) : T_x M \to T_{S(x)} E$ is such that : $\pi'\left(S\left(x\right)\right) S'\left(x\right) = Id$*

Proof. $\pi\left(S\left(x\right)\right) = x \Rightarrow \pi'\left(S\left(x\right)\right) S'\left(x\right) v_x = v_x$ ∎

Vertical space

Definition 1916 *The **vertical space** at p to the fiber bundle E(M,V,π) is : $V_p E = \ker \pi'\left(p\right)$. It is isomorphic to $T_u V$*

This is a vector subspace of $T_p E$. which does not depend on the trivialization

As : $\pi\left(p\right) = x \Rightarrow \pi\left(p\right)' \varphi'_a\left(x, u\right)\left(v_x, v_u\right) = v_x$ so $V_p E = \left\{\varphi'_a\left(x, u\right)\left(v_x, v_u\right), v_x = 0\right\} = \left\{\varphi'_{au}\left(x, u\right) v_u, v_u \in T_u V\right\}$

It is isomorphic to $T_u V$ thus to B_V

Fiber bundle structures

Theorem 1917 *The tangent bundle TE of a fiber bundle $E\left(M, V, \pi\right)$ is the vector bundle $TE\left(TM, TV, \pi \times \pi'\right)$*

Total space : $TE = \cup_{p \in E}\left\{v_p \in T_p E\right\}$

Base : $TM = \cup_{x \in M}\left\{x, v_x \in T_x M\right\}$

Projection : $\pi \times \pi'\left(p\right) v_p = \left(x, v_x\right)$

Open cover : $\left\{TM, x \in O_a\right\}$

Trivializations :

$\left(x, v_x\right) \times \left(u, v_u\right) \in TM \times TV \to \left(\varphi_a(x, u), \varphi'_a(x, u)(v_x, v_u)\right\} \in TE$

Transitions :

$v_{ax} = v_{bx}$

$v_{bu} = \varphi'_{ba}\left(x, u_a\right)\left(v_x, v_{au}\right)$

Basis : $v_p = \sum_\alpha v_x^\alpha \partial x_\alpha + \sum_i v_u^i \partial u_i, \partial x_\alpha = \varphi'_{ax}\left(x, u\right) \partial \xi_\alpha, \partial u_i = \varphi'_{au}\left(x, u\right) \partial \eta_i$

The coordinates of v_p in this atlas are : $\left(\xi^\alpha, \eta^i, v_x^\alpha, v_u^i\right)$

Theorem 1918 *The **vertical bundle** of a fiber bundle $E\left(M, V, \pi\right)$ is the vector bundle : $VE\left(M, TV, \pi\right)$*

total space : $VE = \cup_{p \in E} \{v_p \in V_p E\}$
base : M
projection : $\pi(v_p) = p$
trivializations : $M \times TV \to VE :: \varphi'_{au}(x, u)v_u \in VE$
open cover : $\{O_a\}$
transitions : $\varphi'_{au}(x, u_a) v_{au} = \varphi'_{bu}(x, u_b) v_{bu}$

Forms defined on a fiber bundle

As it has been done previously :

E is a manifold, so $\Lambda_r(E) = \Lambda_r(E; K)$ is the usual space of r forms on E valued in the field K.

Similarly $\Lambda_r(E; H)$ is the space of r forms on E valued in a fixed vector space H.

We will denote :

Notation 1919 $\Lambda_r(E; TE)$ *is the space of r forms on E valued in the tangent bundle TE of E,*

$\Lambda_r(E; VE)$ *is the space of r forms on E valued in the vertical bundle TE of E,*

$\Lambda_r(M; TE)$ *is the space of r forms on M valued in the tangent bundle TE of E,*

$\Lambda_r(M; VE)$ *is the space of r forms on M valued in the vertical bundle TE of E,*

The sections over E are considered as 0 forms on M valued in E

Definition 1920 *A r form on a fiber bundle* $E(M, V, \pi)$ *is said to be* **horizontal** *if it is null whenever one of the vector is vertical.*

A horizontal r form, valued in the vertical bundle, reads :
$\varpi = \sum_{\{\alpha_1 ... \alpha_r\}} \sum_i \varpi^i_{\alpha_1 ... \alpha_r}(p) dx^{\alpha_1} \wedge .. \wedge dx^r \otimes \partial u_i$
A horizontal 1-form, valued in the vertical bundle, sometimes called a soldering form, reads :
$\theta = \sum_{\alpha,i} \theta^i_\alpha(p) dx^\alpha \otimes \partial u_i$

24.1.8 Vector fields on a fiber bundle

Definition 1921 *A vector field on the tangent bundle TE of the fiber bundle* $E(M, V, \pi)$ *is a map :* $W : E \to TE$. *In an atlas* $(O_a, \varphi_a)_{a \in A}$ *of E it is defined by a family* $(W_{ax}, W_{au})_{a \in A}$ *with* $W_{ax} : \pi^{-1}(O_a) \to TM, W_{au} : \pi^{-1}(O_a) \to TV$ *such that for* $x \in O_a \cap O_b : p = \varphi_a(x, u_a)$ $W_{ax}(p) = W_{bx}(p), W_{bu}(p) = (\varphi_{ba}(x, u_a))'(W_x(p), W_{au}(p))$

So : $W_a(\varphi_a(x, u_a)) = \varphi'_a(x, u_a)(W_{ax}(p), W_{au}(p))$
$= \sum_{\alpha \in A} W^\alpha_{ax} \partial x_\alpha + \sum_{i \in I} W^i_{au} \partial u_{ai}$

Definition 1922 *The **commutator** of the vector fields $X, Y \in \mathfrak{X}(TE)$ on the tangent bundle TE of the fiber bundle $E(M, V, \pi)$ is the vector field :* $[X, Y]_{TE}(\varphi_a(x, u_a)) = \varphi'_a(x, u_a)([X_{ax}, Y_{ax}]_{TM}(p), [X_{au}, Y_{au}]_{TV}(p))$

$$[X, Y]_{TE} = \sum_{\alpha\beta} \left(X_x^\beta \partial_\beta Y_x^\alpha - Y_x^\beta \partial_\beta X_x^\alpha \right) \partial x_\alpha + \sum_{ij} \left(X_u^j \partial_j Y_u^i - Y_u^j \partial_j X_u^i \right) \partial u_i$$

24.1.9 Lie derivative

Projectable vector field

For any vector field $W \in \mathfrak{X}(TE) : \pi \times \pi'(W(p)) \in TM$. But there is no guarantee that this projection is itself a vector field over M.

Definition 1923 *A vector field $W \in \mathfrak{X}(TE)$ on a fiber bundle $E(M, V, \pi)$ is **projectable** if $\pi \times \pi'(W) \in \mathfrak{X}(TM)$*

$\forall p \in E : \pi \times \pi'(W(p)) = (\pi(p), Y(\pi(p))), Y \in \mathfrak{X}(M)$
That we can write : $\pi_* W(\pi(p)) = \pi'(p) W(p) \Leftrightarrow Y = \pi_* W$
A vertical vector field W is projectable : indeed if $W(p) \in V_p E$ we have $\pi \times \pi'(W(p)) = 0$.
A vector field : $\pi'(p) W(p) = \pi'(p)(\varphi'_{ax}(x, u_a)(v_x(p), v_{au}(p))) = v_x(p)$ so W is projectable iff $v_x(p)$ does not depend on u.

Theorem 1924 *The flow of a projectable vector field on a fiber bundle E is a fibered manifold morphism*

Proof. The flow of a vector field W on E is a local diffeomorphism on E with : $\frac{\partial}{\partial t} \Phi_W(p, t)|_{t=\theta} = W(\Phi_W(p, \theta))$
It is a fibered manifold morphism if there is : $f : M \to M$ such that : $\pi(\Phi_W(p, t)) = f(\pi(p), t)$
If W is a projectable vector field :
$\frac{\partial}{\partial t} \pi(\Phi_W(p, t))|_{t=\theta} = \pi'(\Phi_W(p, \theta)) \frac{\partial}{\partial t} \Phi_W(p, t)|_{t=\theta} = \pi'(\Phi_W(p, \theta)) W(\Phi_W(p, \theta)) = Y(\pi(\Phi_W(p, t)))$
So we have : $\pi(\Phi_W(p, t)) = \Phi_Y(\pi(p), t)$ and $(\Phi_W(p, t), \Phi_Y(\pi(p), t)) = \Phi_W^E(p, t)$ is a fibered manifold morphism. ∎
If W is vertical, then Y=0 : the map is contained inside the same fiber : $\Phi_W^E(p, t) = (\Phi_W(p, t), Id_M)$

Theorem 1925 *A projectable vector field defines a section on E:*

Proof. take p in E, in a neighborhood n(x) of $x = \pi(p)$ the flow $\Phi_Y(x, t)$ is defined for some interval J, and $\forall x' \in n(x), \exists t : x' = \Phi_Y(x, t)$
take $S(x') = \Phi_W(p, t) \Rightarrow \pi(S(x')) = \Phi_Y(\pi(p), t) = x'$ ∎

Lie derivative of a section

Lie derivatives can be extended to sections on fiber bundles, but with some adjustments (from Kolar p.377). As for the usual Lie derivative, they are a way to compare sections on a fiber bundle at different points, without the need for a covariant derivative.

Theorem 1926 *The Lie derivative of a section S of a fiber bundle $E(M,V,\pi)$ along a projectable vector field W is the section of the vertical bundle :* $\pounds_W S = \frac{\partial}{\partial t}\Phi_W\left(S\left(\Phi_Y\left(x,-t\right)\right),t\right)|_{t=0} \in \mathfrak{X}\left(VE\right)$ *with* $\pi'\left(p\right)W\left(p\right) = Y\left(\pi\left(p\right)\right)$

Proof. The flow $\Phi_W\left(p,t\right)$ is a fibered manifold morphism and a local diffeomorphism on E : $\Phi_W\left(p,t\right) : E \to E$

On a neighborhood of (x,0) in M×ℝ the map :

$F_W\left(x,t\right) = \Phi_W^E\left(S\left(\Phi_Y\left(x,-t\right)\right),t\right) : E \to E$

defines the transport of a section S on E : $S \to \widetilde{S} : \widetilde{S}\left(x\right) = F_W\left(x,t\right)$

We stay in the same fiber because W is projectable and S is a section :

$\pi\left(\Phi_W^E\left(S\left(\Phi_Y\left(x,-t\right)\right),t\right)\right) = \Phi_Y\left(\pi\left(S\left(\Phi_Y\left(x,-t\right)\right)\right),t\right) = \Phi_Y\left(\Phi_Y\left(x,-t\right),t\right) = x$

Therefore if we differentiate in t=0 we have a vector in x, which is vertical :

$\frac{\partial}{\partial t}\Phi_W\left(S\left(\Phi_Y\left(x,-t\right)\right),t\right)|_{t=0}$

$= \frac{\partial}{\partial t}\Phi_W\left(p\left(t\right),t\right)|_{t=0} = \frac{\partial}{\partial p}\Phi_W\left(p\left(t\right),t\right)|_{t=0}\frac{\partial p}{\partial t}|_{t=0} + W\left(p\left(t\right)\right)|_{t=0}$

$= \frac{\partial}{\partial p}\Phi_W\left(p\left(t\right),t\right)|_{t=0}\frac{\partial}{\partial t}S\left(\Phi_Y\left(x,-t\right)\right)|_{t=0} + W\left(S\left(\Phi_Y\left(x,-t\right)\right)\right)|_{t=0}$

$= \frac{\partial}{\partial p}\Phi_W\left(p\left(t\right),t\right)|_{t=0}\frac{\partial}{\partial x}S\left(y\left(t\right)\right)|_{t=0}\frac{\partial}{\partial t}\Phi_Y\left(x,-t\right)|_{t=0} + W\left(S\left(x\right)\right)$

$= -\frac{\partial}{\partial p}\Phi_W\left(S\left(x\right)\right)\frac{\partial}{\partial x}S\left(x\right)Y\left(x\right) + W\left(S\left(x\right)\right)$

$\pounds_W S = -\frac{\partial \Phi_W}{\partial p}\left(S\left(x\right)\right)S'\left(x\right)Y\left(x\right) + W\left(S\left(x\right)\right) \in V_p E$ ∎

So the Lie derivative is a map : $\pounds_W S : M \to VE$

$F_W\left(x,t\right) \circ F_W\left(x,s\right) = F_V\left(x,s+t\right)$ whenever the flows are defined

so : $\frac{\partial}{\partial s}F_W\left(x,s+t\right)|_{s=0} = F_W\left(x,t\right) \circ \frac{\partial}{\partial s}F_W\left(x,s\right)|_{s=0} = F_W\left(x,t\right) \circ \pounds_W S = \pounds_W\left(F_W\left(x,t\right)\right)$

In coordinates :

$S\left(x\right) = \varphi\left(x,\sigma\left(x\right)\right) \to S'\left(x\right) = \sum_\alpha \partial x_\alpha \otimes d\xi^\alpha + \left(\sum_{i\alpha}\left(\partial_\alpha \sigma^i\right)\partial u_i\right) \otimes d\xi^\alpha$

$S'\left(x\right)Y\left(x\right) = \sum_\alpha Y^\alpha\left(x\right)\partial x_\alpha + \sum_{i\alpha}Y^\alpha\left(x\right)\left(\partial_\alpha \sigma^i\right)\partial u_i$

$\pounds_W S = -\frac{\partial \Phi_W}{\partial p}\left(S\left(x\right)\right)\left(\sum_\alpha Y^\alpha\left(x\right)\partial x_\alpha + \sum_i Y^\alpha\left(x\right)\left(\partial_\alpha \sigma^i\right)\partial u_i\right) + W\left(S\left(x\right)\right)$

$\frac{\partial \Phi_W}{\partial p} = \frac{\partial \Phi_W^\beta}{\partial x^\alpha}dx^\alpha \otimes \partial x_\beta + \frac{\partial \Phi_W^i}{\partial x^\alpha}dx^\alpha \otimes \partial u_i + \frac{\partial \Phi_W^\alpha}{\partial u^i}du^i \otimes \partial x_\alpha + \frac{\partial \Phi_W^j}{\partial u^i}du^i \otimes \partial u_j$

$\pounds_W S = \left(W^\alpha - \frac{\partial \Phi_W^\alpha}{\partial x^\beta}Y^\beta - \frac{\partial \Phi_W^\alpha}{\partial u^i}Y^\beta\left(\partial_\beta \sigma^i\right)\right)\partial x_\alpha + \left(W^i - \frac{\partial \Phi_W^i}{\partial x^\alpha}Y^\alpha - \frac{\partial \Phi_W^i}{\partial u^j}Y^\alpha\left(\partial_\alpha \sigma^j\right)\right)\partial u_i$

$W^\alpha = \frac{\partial \Phi_W^\alpha}{\partial x^\beta}Y^\beta + \frac{\partial \Phi_W^\alpha}{\partial u^i}Y^\beta\left(\partial_\beta \sigma^i\right)$

$\pounds_W S = \sum_i \left(W^i - \frac{\partial \Phi_W^i}{\partial x^\alpha}Y^\alpha - \frac{\partial \Phi_W^i}{\partial u^j}Y^\alpha\left(\partial_\alpha \sigma^j\right)\right)\partial u_i$

Notice that this expression involves the derivative of the flow with respect to x and u.

Theorem 1927 *(Kolar p.57) The Lie derivative of a section S of a fiber bundle E along a projectable vector field W has the following properties :*

i) A section is invariant by the flow of a projectable vector field iff its lie derivative is null.

ii) If W,Z are two projectable vector fields on E then $[W,Z]$ is a projectable vector field and we have : $\pounds_{[W,Z]}S = \pounds_W \circ \pounds_Z S - \pounds_Z \circ \pounds_W S$

iii) If W is a vertical vector field $\pounds_W S = W(S(x))$

For ii) $\pi_* ([W,Z])(\pi(p)) = [\pi_* W, \pi_* Z]_M = [W_x, Z_x](\pi(p))$

The result holds because the functor which makes TE a vector bundle is natural (Kolar p.391).

For iii) : If W is a vertical vector field it is projectable and Y=0 so : $\pounds_W S = \frac{\partial}{\partial t}\Phi_W(S(x),t)|_{t=0} = \sum_i W^i \partial u_i = W(S(x))$

Lie derivative of a morphism

The definition can be extended as follows (Kolar p.378):

Definition 1928 *The Lie derivative of the base preserving morphism $F : E_1 \to E_2$ between two fibered manifolds over the same base M: $E_1(M,\pi_1), E_2(M,\pi_2)$, with respect to the vector fields $W_1 \in \mathfrak{X}(TE_1), W_2 \in \mathfrak{X}(TE_2)$ projectable on the same vector field $Y \in \mathfrak{X}(TM)$ is : $\pounds_{(W_1,W_2)}F(p) = \frac{\partial}{\partial t}\Phi_{W2}(F(\Phi_{W_1}(p,-t)),t)|_{t=0} \in \mathfrak{X}(VE_2)$*

Proof. 1st step : $p \in E_1 : \Phi_{W_1}(p,-t) :: \pi_1(\Phi_{W_1}(p,-t)) = \pi_1(p)$ because W_1 is projectable

2nd step : $F(\Phi_{W_1}(p,-t)) \in E_2 :: \pi_2(F(\Phi_{W_1}(p,-t))) = \pi_1(p)$ because F is base preserving

3d step : $\Phi_{W2}(F(\Phi_{W_1}(p,-t)),t) \in E_2 :: \pi_2(\Phi_{W2}(F(\Phi_{W_1}(p,-t)),t)) = \pi_1(p)$ because W_2 is projectable

$\frac{\partial}{\partial t}\pi_2(\Phi_{W2}(F(\Phi_{W_1}(p,-t)),t))|_{t=0} = 0$

So $\pounds_{(W_1,W_2)}F(p)$ is a vertical vector field on E_2 ∎

24.2   Vector bundles

This is the generalization of the vector bundle over a manifold. With a vector bundle we can use vector spaces located at each point over a manifold and their tensorial products, for any dimension, and get more freedom in the choice of a basis. The drawback is that each vector bundle must be defined through specific atlas, whereas the common vector bundle was borned from the manifold structure itself.

24.2.1   Definitions

Definition 1929 *A **vector bundle** $E(M,V,\pi)$ is a fiber bundle whose standard fiber V is a Banach vector space and transitions functions at each point $\varphi_{ab}(x,.) : V \to V$ are continuous linear invertible maps :$\varphi_{ab}(x) \in G\mathcal{L}(V;V)$*

So : with an atlas $(O_a, \varphi_a)_{a \in A}$ of E :

$\forall a, b : O_a \cap O_b \neq \varnothing : \exists \varphi_{ba} : O_a \cap O_b \to G\mathcal{L}(V; V) ::$

$\forall p \in \pi^{-1}(O_a \cap O_b), p = \varphi_a(x, u_a) = \varphi_b(x, u_b) \Rightarrow u_b = \varphi_{ba}(x) u_a$

The transitions maps must meet the cocycle conditions :

$\forall a, b, c \in A : \varphi_{aa}(x) = 1; \varphi_{ab}(x) \varphi_{bc}(x) = \varphi_{ac}(x)$

With the following maps:

$\tau_a : \pi^{-1}(O_a) \to V :: \varphi_a(x, \tau_a(p)) = p = \varphi_a(\pi(p), \tau_a(p))$

$\forall p \in \pi^{-1}(O_a) \cap \pi^{-1}(Q_i) : \tau_b(p) = \varphi_{ba}(\pi(p), \tau_a(p)) = \varphi_{ba}(\pi(p)) \tau_a(p) \Leftrightarrow$
$\varphi_{ba}(\pi(p)) = \tau_b(p) \circ \tau_a(p)^{-1}$

Examples :

the tangent bundle $TM(M, B, \pi)$ of a manifold M modelled on the Banach B.

the tangent bundle $TE(TM, B_V, \pi \times \pi')$ of a fiber bundle $E(M, V, \pi)$ where V is modelled on B_V

Theorem 1930 *(Kolar p.69) Any finite dimensional vector bundle admits a finite vector bundle atlas*

Remark : in the definition of fiber bundles we have required that all the manifolds are on the same field K. However for vector bundles we can be a bit more flexible.

Definition 1931 *A **complex vector bundle** $E(M, V, \pi)$ over a real manifold M is a fiber bundle whose standard fiber V is a Banach complex vector space and transitions functions at each point $\varphi_{ab}(x, .) : V \to V$ are complex continuous linear maps :$\varphi_{ab}(x) \in \mathcal{L}(V; V)$.*

24.2.2 Vector space structure

The main property of a vector bundle is that each fiber has a vector space structure, isomorphic to V : the fiber $\pi^{-1}(x)$ over x is just a copy of V located at x.

Definition of the operations

Theorem 1932 *The fiber over each point of a vector bundle $E(M, V, \pi)$ has a canonical structure of vector space, isomorphic to V*

Proof. Define the operations, pointwise with an atlas $(O_a, \varphi_a)_{a \in A}$ of E

$p = \varphi_a(x, u_a), q = \varphi_a(x, v_a), k, k' \in K : kp + k'q = \varphi_a(x, ku_a + k'v_a)$

then :

$p = \varphi_b(x, u_b), q = \varphi_b(x, v_b), u_b = \varphi_{ba}(x, u_a), v_b = \varphi_{ba}(x, v_a)$

$kp + k'q = \varphi_b(x, ku_b + k'v_b) = \varphi_b(x, k\varphi_{ba}(x, u_a) + k'\varphi_{ba}(x, v_a))$

$= \varphi_b(x, \varphi_{ba}(x, ku_a + k'v_a)) = \varphi_a(x, ku_a + k'v_a)$ ∎

With this structure of vector space on E(x) the trivializations are linear in u : $\varphi(x, .) \in \mathcal{L}(V; E(x))$

Holonomic basis

Definition 1933 *The **holonomic basis** of the vector bundle $E(M, V, \pi)$ associated to the atlas $(O_a, \varphi_a)_{a \in A}$ of E and a basis $(e_i)_{i \in I}$ of V is the basis of each fiber defined by : $e_{ia} : O_a \to E :: e_{ia}(x) = \varphi_a(x, e_i)$. At the transitions :*
$e_{ib}(x) = \varphi_{ab}(x) e_{ia}(x)$

Warning ! we take the image of the *same* vector e_i in each open. So at the transitions $e_{ia}(x) \neq e_{ib}(x)$. *They are not sections.* This is similar to the holonomic bases of manifolds : $\partial x_\alpha = \varphi'_a(x)^{-1}(\varepsilon_\alpha)$

Proof. At the transitions :

$e_{ia}(x) = \varphi_a(x, e_i), e_{ib}(x) = \varphi_b(x, e_i) = \varphi_a(x, u_a) \Rightarrow e_i = \varphi_{ba}(x) u_a \Leftrightarrow u_a = \varphi_{ab}(x) e_i$

$e_{ib}(x) = \varphi_a(x, \varphi_{ab}(x) e_i) = \varphi_{ab}(x) \varphi_a(x, e_i) = \varphi_{ab}(x) e_{ia}(x)$ ■

In the holonomic basis a vector of E(x) has the components : $v = \sum_{i \in I} v_a^i e_{ai}(x) = \varphi_a\left(x, \sum_{i \in I} v_a^i e_{ai}\right) = \varphi_a(x, v_a)$ and at the transitions the components $v_b^i = \sum_j \varphi_{ba}(x)_j^i v_a^j$

$v = \sum_{i \in I} v_a^i e_{ai}(x) = \sum_{i \in I} v_b^i e_{bi}(x) = \varphi_a(x, v_a) = \varphi_b(x, v_b)$

$v_b = \varphi_{ba}(x) v_a \Leftrightarrow [v_b] = [\varphi_{ba}(x)] [v_a]$

The map : $\tau_a : \pi^{-1}(O_a) \to V :: p = \varphi_a(\pi(p), \tau_a(p))$ is linear :

$\tau_a\left(\sum_i u_a^i e_{ai}(x)\right) = \sum_i u_a^i e_{ai}$

Conversely, if we have for each open O_a a set of local maps $e_{ia} : O_a \to E$ such that $(e_{ia}(x))_{i \in I}$ is a basis of $\pi^{-1}(x)$, pick up a basis $(e_i)_{i \in I}$ of V, then each vector $u \in V$ reads : $u = \sum_{i \in I} u^i e_i$, and we can define the trivializations by :

$\varphi_a : O_a \times V \to u = \sum_{i \in I} u^i e_{ia}(x) = \varphi_a\left(x, \sum_{i \in I} u^i e_i\right)$

The bases of a vector bundle are not limited to the holonomic basis defined by a trivialization. Any other basis can be defined at any point, by an endomorphism in the fiber. Notice that a vector of E does not depend on a basis : it is the same as any other vector space.

Sections

Definition 1934 *With an atlas $(O_a, \varphi_a)_{a \in A}$ of E a section $U \in \mathfrak{X}(E)$ of the vector bundle $E(M, V, \pi)$ is defined by a family of maps $(u_a)_{a \in A}, u_a : O_a \to V$ such that :*

$x \in O_a : U(x) = \varphi_a(x, u_a(x)) = \sum_{i \in I} u_a^i(x) e_{ai}(x)$

$\forall x \in O_a \cap O_b : U(x) = \sum_{i \in I} u_a^i(x) e_{ai}(x) = \sum_{i \in I} u_b^i(x) e_{bi}(x) \Leftrightarrow u_b^i(x) = \sum_{j \in I} \left([\varphi_{ba}(x)]\right)_j^i u_a^j(x)$

Theorem 1935 *The set of sections $\mathfrak{X}(E)$ over a vector bundle has the structure of a vector space.*

Notice that it is infinite dimensional, and thus usually is not isomorphic to V.

Theorem 1936 *(Giachetta p.13) If a finite dimensional vector bundle E admits a family of global sections which spans each fiber the fiber bundle is trivial.*

Warning ! there is no commutator of sections $\mathfrak{X}(E)$ on a vector bundle. But there is a commutator for sections $\mathfrak{X}(TE)$ of TE.

Tangent bundle

Theorem 1937 *The tangent space to a vector bundle $E(M,V,\pi)$ has the vector bundle structure : $TE\,(TM, V \times V, \pi \times \pi')$ with $\pi'(p)(v_p) = v_x$. A vector v_p of T_pE is a couple $(v_x, v_u) \in T_xM \times V$ which reads : $v_p = \sum_{\alpha \in A} v_x^\alpha \partial x_\alpha + \sum_{i \in I} v_u^i e_i(x)$*

Proof. With an atlas $(O_a, \varphi_a)_{a \in A}$ of E the tangent space T_pE to E at p=$\varphi_a(x,u)$ has the basis $\partial x_\alpha = \varphi'_{ax}(x,u)\,\partial \xi_\alpha, \partial u_i = \varphi'_{au}(x,u)\,\partial \eta_i$ with holonomic bases $\partial \xi_\alpha \in T_xM, \partial \eta_i \in T_uV$

$v_p = \sum_\alpha v_x^\alpha \partial x_\alpha + \sum_i v_u^i \partial u_i$

But : $\partial \eta_i = e_i$ and φ is linear with respect to u, so : $\partial u_i = \varphi'_{au}(x,u)\,e_i = \varphi_a(x, e_i) = e_{ai}(x)$ ∎

So the basis of T_pE is $\partial x_\alpha = \varphi'_{ax}(x,u)\,\partial \xi_\alpha, \partial u_i = e_{ai}(x)$

The coordinates of v_p in this atlas are : $\left(\xi^\alpha, \eta^i, v_x^\alpha, v_u^i\right)$. The coordinates of u appears by $\varphi'_{ax}(x,u)$ and p

At the transitions we have the identities : $v_p = \varphi'_a(x, u_a)(v_x, v_{au}) = \varphi'_b(x, u_b)(v_x, v_{bu})$ with $v_{bu} = (\varphi_{ba}(x)(u_a))'(v_x, v_{au}) = \left(\varphi_{ba}(x)'\,v_x\right)(u_a) + \varphi_{ba}(x)(v_{au})$

Theorem 1938 *The vertical bundle $VE\,(M,V,\pi)$ is a trivial bundle isomorphic to $E \times_M E$*

we need both p (for $e_i(x)$) and v_u for a point in VE : $v_p \in V_pE$: $v_p = \sum_{i \in I} v_u^i e_i(x)$

The vertical cotangent bundle is the dual of the vertical tangent bundle, and is not a subbundle of the cotangent bundle.

Vector fields on the tangent bundle TE are defined by a family $(W_{ax}, W_{au})_{a \in A}$ with $W_{ax} \in \mathfrak{X}(TO_a), W_{au} \in C(O_a; V)$ such that for $x \in O_a \cap O_b$: $W_{ax}(p) = W_{bx}(p), W_{bu}(p) = (\varphi_{ba}(x)(u_a))'(W_x(p), W_{au}(p))$

So : $W_a(\varphi_a(x, u_a)) = \sum_{\alpha \in A} W_{ax}^\alpha(p)\,\partial x_\alpha + \sum_{i \in I} W_{au}^i(p)\,e_{ai}(x)$

Norm on a vector bundle

Theorem 1939 *Each fiber of a vector bundle is a normed vector space*

We have assumed that V is a Banach vector space, thus a normed vector space. With an atlas $(O_a, \varphi_a)_{a \in A}$ of E we can define pointwise a norm on $E(M,V,\pi)$ by :

$\| \|_E : E \times E \to \mathbb{R}_+ :: \|\varphi_a(x, u_a)\|_E = \|u_a\|_V$

The definition is consistent iff :

$\forall a, b \in A, O_a \cap O_b \neq \varnothing, \forall x \in O_a \cap O_b :$

$\|\varphi_a (x, u_a)\|_E = \|u_a\|_V = \|\varphi_b (x, u_b)\|_E = \|u_b\|_V = \|\varphi_{ba} (x) u_a (x)\|$

The transitions maps : $\varphi_{ba} (x) \in \mathcal{L}(V; V)$ are continuous so :

$\|u_b\|_V \leq \|\varphi_{ba} (x)\| \|u_a (x)\|_V, \|u_a\|_V \leq \|\varphi_{ab} (x)\| \|u_b (x)\|_V$

and the norms defined on O_a, O_b are equivalent : they define the same topology (see Normed vector spaces). So for any matter involving only the topology (such as convergence or continuity) we have a unique definition, however the norm are not the same.

Norms and topology on the space of sections $\mathfrak{X}(E)$ can be defined, but we need a way to agregate the results. This depends upon the maps $\sigma_a : O_a \to V$ (see Functional analysis) and without further restrictions upon the maps $M \to TE$ we have *not* a normed vector space.

Lie derivative

The general results for fiber bundle hold. As any vector bundle can be made a tensorial bundle, whose sections are tensor fields, the Lie derivative of a tensor field is a tensor field of the same type.

1. Projectable vector fields W on TE are such that :

$\pi'(p) W(p) = \pi'(p) \left(\sum_\alpha W_x^\alpha \partial x_\alpha + \sum_i W_u^i \partial u_i \right) = Y(\pi(p)) = \sum_\alpha W_x^\alpha \partial \xi_\alpha$

the components W_x^α must depend on x only (and not u).

2. The Lie derivative of a section X on E, that is a vector field $\sum X^i(x) e_i(x) = X(x)$, along a projectable vector field W on TE is :

$\pounds_W X = \frac{\partial}{\partial t} \Phi_W (X(\Phi_Y(x, -t)), t) |_{t=0}$

This is a map : $\pounds_W : \mathfrak{X}(E) \to \mathfrak{X}(VE)$

As the vertical bundle is isomorphic to $E \times_M E$ we have : $\widehat{\pounds}_W X = (X, \pounds_W X)$ and the second part $\pounds_W S$ only is significant and can be seen as a section of E. So there is a map : $divW \in \mathcal{L}(E; E) : \pounds_W X = divW(X)$

Definition 1940 *The divergence of a projectable vector field W on a vector bundle E is the linear map : $divW \in \mathcal{L}(E; E)$ such that for any section X of E and projectable vector field W : $\pounds_W X = divW(X)$*

24.2.3 Maps between vectors bundles

Morphism of vector bundles

Definition 1941 *A **morphism between vector bundles** $E_1(M_1, V_1, \pi_1), E_2(M_2, V_2, \pi_2)$ is a map : $F : E_1 \to E_2$ such that :*

$\forall x \in M_1 : F\left(\pi_1^{-1}(x)\right) \in \mathcal{L}\left(\pi_1^{-1}(x); \pi_2^{-1}(f(x))\right)$ *where : $f : M_1 \to M_2 ::$ $f \circ \pi_1 = \pi_2 \circ F$*

so it preserves both the fiber and the vector structure of the fiber.

$\pi_1^{-1}(x) \simeq V_1, \pi_2^{-1}(f(x)) \simeq V_2$

So $\forall x \in M_1 : F(x) \in \mathcal{L}(V_1; V_2)$ and we can define a vector bundle morphism as a couple :

$f : M_1 \to M_2$

$F : M_1 \to \mathcal{L}(V_1; V_2)$

then with the trivializations : $p = \varphi_a(x, u) \in E_1, q = \psi_A(y, v) \in E_2$

$p = \varphi_a(x, u) \in E_1 \to q = \psi_A(f(x), F(x)u)$

A vector bundle morphism is usually base preserving : f is the identity. Then it can be defined by a single map : $\forall x \in M : F(x) \in \mathcal{L}(E_1(x); E_2(x))$

Definition 1942 *A vector bundle $E_1(M_1, V_1, \pi_1)$ is a **vector subbundle** of $E_2(M_2, V_2, \pi_2)$ if :*

i) $E_1(M_1, \pi_1)$ is a fibered submanifold of $E_2(M_2, \pi_2)$: M_1 is a submanifold of $M_2, \pi_2|_{M_1} = \pi_1$

ii) there is a vector bundle morphism (F,f) : $E_1 \to E_2$

Pull back of vector bundles

Theorem 1943 *The pull back of a vector bundle is a vector bundle*

Proof. They have the same standard fiber.

Let $E(M, V, \pi)$ a vector bundle with atlas $(O_a, \varphi_a)_{a \in A}$, N a manifold, f a continuous map $f : N \to M$ then the fiber in f*E over (y,p) is

$\widetilde{\pi}^{-1}(y, p) = (y, \varphi_a(f(y), u))$ with $p = \varphi_a(f(y), u)$

This is a vector space with the operations (the first factor is neutralized) :

$k(y, \varphi_a(f(y), u)) + k'(y, \varphi_a(f(y), v)) = (y, \varphi_a(f(y), ku + k'v))$

For a section $e_{ia} : O_a \to E :: e_{ia}(x) = \varphi_a(x, e_i)$ we have the pull back $f^* e_{ia} : N \to E :: f^* e_i(y) = (f(y), e_{ia}(f(y)))$ so it is a basis of $\widetilde{\pi}^{-1}(y, p)$ with the previous vector space structure. ■

Whitney sum

The Whitney sum $E_1 \oplus E_2$ of two vector bundles $E_1(M, V_1, \pi_1), E_2(M, V_2, \pi_2)$ can be identified with $E(M, V_1 \oplus V_2, \pi)$ with :

$E = \{p_1 + p_2 : \pi_1(p_1) = \pi_2(p_2)\}, \pi(p_1, p_2) = \pi_1(p_1) = \pi_2(p_2)$

Theorem 1944 *(Kolar p.69) For any finite dimensional vector bundle $E_1(M, V_1, \pi_1)$ there is a second vector bundle $E_2(M, V_2, \pi_2)$ such that the Whitney sum $E_1 \oplus E_2$ is trivial*

Exact sequence

(Giachetta p.14) Let $E_1(M, V_1, \pi_1), E_2(M, V_2, \pi_2), E_3(M, V_3, \pi_3)$ be 3 vector bundles on *the same base M*

They form a **short exact sequence**, denoted :

$0 \to E_1 \overset{f_1}{\to} E_2 \overset{f_2}{\to} E_3 \to 0$

if : f_1 is a vector bundle monomorphism (injective), f_2 is a vector bundle epimorphism (surjective)

then : E_3 is the factor bundle, and $V_2/V_1 \simeq V_3$

Then there is the dual exact sequence :

$$0 \to E_3^* \overset{f_1^*}{\to} E_2^* \overset{f_2^*}{\to} E_1^* \to 0$$

with the dual maps and the dual vector spaces $V_i^*, i = 1, 2, 3$

Every exact sequence of vector bundle splits : there is a vector bundle monomorphism (surjective) : $g : E_3 \to E_2$such that : $f_3 \circ g = Id_{E_3}$ then : $E_2 = E_1 \oplus E_3, V_2 = V_1 \oplus V_3 = V_1 \oplus (V_2/V_1)$

24.2.4 Tensorial bundles

Functor on vector bundles

Theorem 1945 *The vector bundles on a field K and their morphisms constitute a category* $\mathfrak{VM}$

The functors over the category of vector spaces on a field K : $\mathcal{F} = \mathfrak{D}, \mathfrak{T}^r, \mathfrak{T}_s :$ $\mathfrak{V} \to \mathfrak{V}$ (see Algebra-tensors) transform a vector space V in its dual V^*, its tensor powers $\otimes^r V, \otimes_s V^*$ and linear maps between vector spaces in similarly morphisms.

We can restrict their action to the subcategory of Banach vector spaces and continuous morphisms.

We define the functors $\mathcal{F}\mathfrak{M} = \mathfrak{D}\mathfrak{M}, \mathfrak{T}^r\mathfrak{M}, \mathfrak{T}_s\mathfrak{M} : \mathfrak{VM} \to \mathfrak{VM}$ as follows :

i) To any vector bundle $E(M, V, \pi)$ we associate the vector bundle :

\- with the same base manifold

\- with standard fiber $\mathcal{F}(V)$

\- to each transition map $\varphi_{ba}(x) \in GL(V; V)$ the map : $\mathcal{F}(\varphi_{ba}(x)) \in$ $GL(\mathcal{F}(V); \mathcal{F}(V))$

then we have a vector bundle $\mathcal{F}\mathfrak{M}(E(M, V, \pi)) = \mathcal{F}E(M, \mathcal{F}V, \pi)$

ii) To any morphism of vector bundle $F \in \hom_{\mathfrak{VM}}(E_1(M_1, V_1, \pi_1), E_2(M_2, V_2, \pi_2))$ such that : $\forall x \in M_1 : F(\pi_1^{-1}(x)) \in \mathcal{L}(\pi_1^{-1}(x); \pi_2^{-1}(f(x)))$ where : $f : M_1 \to M_2 :: f \circ \pi_1(p) = \pi_2 \circ F$ we associate the morphism $\mathcal{F}\mathfrak{M}F \in \hom_{\mathfrak{VM}}(\mathcal{F}\mathfrak{M}E_1(M_1, \mathcal{F}V_1, \pi_1), \mathcal{F}E_2(M_2, \mathcal{F}V_2, \pi_2))$ with : $\mathcal{F}\mathfrak{M}f = f; \mathcal{F}\mathfrak{M}F(\pi_1^{-1}(x)) = \mathcal{F}F(\pi_1^{-1}(x))$

The tensors over a vector bundle do not depend on their definition through a basis, either holonomic or not, and neither on a definition on V. This is similar to the tensors on the tangent bundle of a manifold.

Dual bundle
The dual bundle of $E(M, V, \pi)$ is the result of the application of the functor $\mathfrak{D}\mathfrak{M}$

If $E(M, V, \pi)$ is a vector bundle with trivializations $(O_a, \varphi_a)_{a \in A}$ and transitions maps : $\varphi_{ab}(x) \in \mathcal{L}(V; V)$ its dual bundle denoted $E'(M, V', \pi)$ has :

same base M

the fiber V' : the topological dual of V, this is a Banach vector space if V is a Banach.

for each transition map the transpose : $\varphi_{ab}^t(x) \in \mathcal{L}(V'; V') : \varphi_{ab}^t(x)(\lambda)(u) = \lambda(\varphi_{ab}(x)(u))$

for trivialization $(O_a, \varphi_a^t)_{a \in A}$.

If $(e^i)_{i \in I}$ is a basis of V' such that : $e^i (e_j) = \delta_{ij}$ on each fiber E'(x) we define : $e_a^i (x) = \varphi_a^t (x, e^i)$

At the transitions we have :

$e_b^i (x) = \varphi_b^t (x, e^i) = \varphi_a^t (x, \varphi_{ba}^t (x) e^i) = \sum_{j \in I} [\varphi_{ab}^t (x)]_j^i e_a^j (x)$

$e_b^i (x) (e_{jb} (x)) = \left(\sum_{k \in I} [\varphi_{ab}^t (x)]_k^i e_a^k (x) \right) \left(\sum_{l \in I} [\varphi_{ab} (x)]_j^l e_{la} (x) \right) = \sum_{k \in I} [\varphi_{ab}^t (x)]_k^i [\varphi_{ab} (x)]_j^k = \delta_j^i$

Thus : $[\varphi_{ab}^t (x)] = [\varphi_{ab} (x)]^{-1} \Leftrightarrow \varphi_{ab}^t (x) = \varphi_{ba} (x)$

$e_b^i (x) = \sum_{j \in I} [\varphi_{ba} (x)]_j^i e_a^j (x)$

A section Λ of the dual bundle is defined by a family of maps $(\lambda_a)_{a \in A}, \lambda_a : O_a \to V'$ such that :

$x \in O_a : \Lambda (x) = \varphi_a^t (x, \lambda_a (x)) = \sum_{i \in I} \lambda_{ai} (x) e_a^i (x)$

$\forall x \in O_a \cap O_b : \lambda_{bi} (x) = \sum_{j \in I} [\varphi_{ab} (x)]_i^j \lambda_{aj} (x)$

So we can define pointwise the action of $\mathfrak{X} (E')$ on $\mathfrak{X} (E) : \mathfrak{X} (E') \times \mathfrak{X} (E) \to C (M; K) :: \Lambda (x) (U (x)) = \lambda_a (x) u_a (x)$

Tensorial product of vector bundles

The action of each of the others functors on a vector bundle : $E(M, V, \pi)$ with trivializations $(O_a, \varphi_a)_{a \in A}$ and transitions maps : $\varphi_{ab} (x) \in \mathcal{L} (V; V)$ gives :

Notation 1946 $\otimes^r E$ is the vector bundle $\otimes^r E (M, \otimes^r V, \pi)) = \mathfrak{T}^r \mathfrak{M} (E(M, V, \pi))$

$\otimes_s E$ is the vector bundle $\otimes_s E (M, \otimes_s V, \pi)) = \mathfrak{T}_s \mathfrak{M} (E(M, V, \pi))$

Moreover as the transition maps are invertible, we can implement the functor $\mathfrak{T}_s^r$ and get $\otimes_s^r E(M, (\otimes^r V) \otimes (\otimes_s V'), \pi)$, as well.

Notation 1947 $\otimes_s^r E$ is the vector bundle $\otimes_s^r E (M, \otimes_s V, \pi)) = \mathfrak{T}_s^r \mathfrak{M} (E(M, V, \pi))$

Similarly we have the algebras of symmetric tensors and antisymmetric tensors.

Notation 1948 $\odot^r E$ is the vector bundle $\odot^r E (M, \odot^r V, \pi))$

$\wedge_s E'$ is the vector bundle $\wedge_s E (M, \wedge_s V', \pi))$

The sections of these vector bundles are denoted accordingly : $\mathfrak{X} (\otimes_s^r E)$

Notice :

i) The tensorial bundles are not related to the tangent bundle of the vector bundle or of the base manifold

ii) $\Lambda_s E' (M, \Lambda_s V', \pi) \neq \Lambda_s (M; E)$

These vector bundles have all the properties of the tensor bundles over a manifold, as seen in the Differential geometry part: linear combination of tensors of the same type, tensorial product, contraction, as long as we consider only pointwise operations which do not involve the own tangent bundle of the base M.

The trivializations are defined on the same open cover and any map $\varphi_a(x,u)$ can be uniquely extended to a map

$\Phi_{ar,s} : O_a \times \otimes^r V \otimes \otimes_s V' \to \otimes_s^r E$

such that : $\forall i_k, j_l \in I : U = \Phi_{ar,s}\left(x, e_{i_1} \otimes ... \otimes e_{i_r} \otimes e^{j_1} \otimes ... \otimes e^{j_s}\right) = e_{ai_1}(x) \otimes ... \otimes e_{ai_r}(x) \otimes e^{aj_1}(x) \otimes ... \otimes e^{aj_s}(x)$

Sections of the tensorial bundles $\mathfrak{X}\left(\otimes_s^r E\right)$ are families of maps : $T_a : O_a \to \otimes_s^r E$ such that on the transitions we have :

$T(x) = \sum_{i_1...i_r j_1...j_s} T_{aj_1...j_s}^{i_1...i_r} e_{ai_1}(x) \otimes ... \otimes e_{ai_r}(x) \otimes e^{aj_1}(x) \otimes ... \otimes e^{aj_s}(x)$

$T(x) = \sum_{i_1...i_r j_1...j_s} T_{bj_1...j_s}^{i_1...i_r} e_{bi_1}(x) \otimes ... \otimes e_{bi_r}(x) \otimes e^{bj_1}(x) \otimes ... \otimes e^{bj_s}(x)$

$e_{ib}(x) = \sum_{j \in I} [\varphi_{ab}(x)]_i^j e_{ja}(x)$

$e_b^j(x) = \sum_{j \in I} [\varphi_{ba}(x)]_i^j e_a^j(x)$

So :

$T_{bj_1...j_s}^{i_1...i_r} = \sum_{k_1...k_r} \sum_{l_1....l_s} T_{al_1...l_s}^{k_1...k_r} [J]_{k_1}^{i_1} .. [J]_{k_r}^{i_r} \left[J^{-1}\right]_{j_1}^{l_1} .. \left[J^{-1}\right]_{j_s}^{l_s}$

with $[J] = [\varphi_{ba}(x)]$, $[J]^{-1} = [\varphi_{ab}(x)]$

For a r-form in $\Lambda_r E'$ these expressions give the formula :

$T_{bi_1...i_r}(x) = \sum_{\{j_1,...,j_r\}} T_{aj_1...j_r} \det\left[J^{-1}\right]_{i_1...i_r}^{j_1...r}$

where $\det\left[J^{-1}\right]_{i_1...i_r}^{j_1...r}$ is the determinant of the matrix $\left[J^{-1}\right]$ with r column $(i_1,..i_r)$ comprised each of the components $\{j_1...j_r\}$

We still call contravariant a section of $\otimes^r E$ and covariant a section of $\otimes_s E'$.

More generally we can define the tensorial product of two different vector bundle *over the same manifold*:

Definition 1949 *The tensorial product $E_1 \otimes E_2$ of two vector bundles $E_1(M, V_1, \pi_1)$, $E_2(M, V_2, \pi_2)$ over the same manifold is the set : $\cup_{x \in M} E_1(x) \otimes E_2(x)$, which has the structure of the vector bundle $E_1 \otimes E_2(M, V_1 \otimes V_2, \pi)$*

24.2.5 Scalar product on a vector bundle

The key point is that a scalar product on V can induce a scalar product on $E(M, V, \pi)$ if it is preserved by the transition maps. M is not involved, except if E=TM.

General definition

Definition 1950 *A scalar product on a vector bundle $E(M, V, \pi)$ is a map defined on M such that $\forall x \in M$ $g(x)$ is a non degenerate, bilinear symmetric form if E is real, or a sesquilinear hermitian form if E is complex.*

For any atlas $(O_a, \varphi_a)_{a \in A}$ of E with transition maps φ_{ba}, g is defined by a family $(g_a)_{a \in A}$ of forms defined on each domain O_a :

$g(x)\left(\sum_i u^i e_{ai}(x), \sum_i v^i e_{ai}(x)\right) = \sum_{ij} \overline{u}^i v^j g_{aij}(x)$ with : $g_{aij}(x) = g(x)\left(e_a^i(x), e_a^j(x)\right)$

or equivalently by a family of maps γ_a such that $\gamma_a(x)$ is a scalar product on V.

At the transitions :

$x \in O_a \cap O_b : g_{bij}(x) = g(x) \left(e_b^i(x), e_b^j(x) \right) = g(x) \left(\varphi_{ab}(x) e_a^i(x), \varphi_{ab}(x) e_a^j(x) \right)$

So : $[g_b(x)] = [\varphi_{ab}(x)]^* [g_a(x)] [\varphi_{ab}(x)]$

There are always orthonormal basis. If the basis $e_{ai}(x)$ is orthonormal then it will be orthonormal all over E if the transition maps are orthonormal : $\sum_k [\varphi_{ab}(x)]_i^k [\varphi_{ab}(x)]_j^k = \eta_{ij}$

A metric g on a manifold defines a scalar product on the tangent and cotangent bundle in the usual way. It is possible to extend these metric to any tensor bundle (see Algebra - tensorial product of maps). The most usual is the metric defined on the exterior algebra $\mathfrak{X}(\Lambda TM^*)$ through the Hodge formalism.

Tensorial definition

1. Real case :

Theorem 1951 *A symmetric covariant tensor $g \in \mathfrak{X}(\odot_2 E)$ defines a bilinear symmetric form on each fiber of a real vector bundle $E(M, V, \pi)$ and a scalar product on E if this form is non degenerate.*

Such a tensor reads in a holonomic basis of E :

$g(x) = \sum_{ij} g_{aij}(x) e_a^i(x) \otimes e_a^j(x)$ with $g_{aij}(x) = g_{aji}(x)$

at the transitions : $g_{bij}(x) = \sum_{kl} [\varphi_{ab}(x)]_i^k [\varphi_{ab}(x)]_j^l g_{akl}(x) \Leftrightarrow [g_b] = [\varphi_{ab}(x)]^t [g_a] [\varphi_{ab}(x)]$

For : X,Y$\in \mathfrak{X}(E) : g(x)(X(x), Y(x)) = \sum_{ij} g_{aij}(x) X_a^i(x) Y_a^j(x)$

2. Complex case :

Definition 1952 *A **real structure on a complex vector bundle** $E(M, V, \pi)$ is a map $\sigma \in C_0(M; E)$ such that $\sigma(x)$ is antilinear on E(x) and $\sigma^2(x) = Id_{E(x)}$*

Theorem 1953 *A real structure on a complex vector bundle $E(M, V, \pi)$ with atlas $(O_a, \varphi_a)_{a \in A}$ and transition maps φ_{ab} is defined by a family of maps $(\sigma_a)_{a \in A}$ defined on each domain O_a such that at the transitions :$[\sigma_b(x)] = [\varphi_{ba}(x)] [\sigma_a(x)] \overline{[\varphi_{ab}(x)]}$*

Proof. σ is defined by a family of maps : $\sigma \left(\sum u_a^i e_{ai}(x) \right) = \sum \overline{u}_a^i \sigma(e_{ai}(x))$ in matrix notation :$\sigma(e_{ai}(x)) = \sum_j [\sigma_a(x)]_i^j e_{aj}(x)$ so : $\sigma(u) = \sum_j [\sigma_a(x)]_j^i \overline{u}_a^j e_{ai}(x)$

At the transitions :

$\sigma(e_{bi}(x)) = \sum_j \sigma \left([\varphi_{ab}(x)]_i^j e_{ja}(x) \right) = \sum_j \overline{[\varphi_{ab}(x)]}_i^j \sigma(e_{ja}(x)) = \sum_j \overline{[\varphi_{ab}(x)]}_i^j [\sigma(e_a(x))]_j^k e_{ka}(x)$

$= \sum_{jkl} \overline{[\varphi_{ab}(x)]}_i^j [\sigma(e_a(x))]_j^k [\varphi_{ba}(x)]_k^l e_{lb}(x) = \sum_k [\sigma(e_b(x))]_i^l e_{lb}(x) \Leftrightarrow [\sigma_b(x)]_i^l = [\varphi_{ba}(x)]_k^l [\sigma_a(x)]_j^k \overline{[\varphi_{ab}(x)]}_i^j$

In matrix notations : $[\sigma_b(x)] = [\varphi_{ba}(x)] [\sigma_a(x)] \overline{[\varphi_{ab}(x)]}$ ∎

Warning ! the matrix notation is convenient but don't forget that σ is antilinear

Theorem 1954 *A real structure σ on a complex vector space V induces a real structure on a complex vector bundle $E(M, V, \pi)$ iff the transition maps φ_{ab} are real maps with respect to σ*

Proof. A real structure σ on V it induces a real structure on E by : $\sigma(x)(u) = \varphi_a(x, \sigma(u_a))$

The definition is consistent iff : $\sigma(x)(u) = \varphi_a(x, \sigma(u_a)) = \varphi_b(x, \sigma(u_b))$ with $u_b = \varphi_{ba}(x)(u_a)$

$\sigma(\varphi_{ba}(x)(u_a)) = \varphi_{ba}(x)\sigma(u_a) \Leftrightarrow \sigma \circ \varphi_{ba}(x) = \varphi_{ba}(x) \circ \sigma$

So iff $\varphi_{ab}(x)$ is a real map with respect to σ ■

Theorem 1955 *A complex structure σ and a covariant tensor $g \in \mathfrak{X}(\otimes_2 E)$ such that $g(x)(\sigma(x)(u), \sigma(v)) = \overline{g(x)(v,u)}$ define a hermitian sequilinear form, on a complex vector bundle $E(M, V, \pi)$ by $\gamma(x)(u,v) = g(x)(\sigma(x)u, v)$*

Proof. g defines a C-bilinear form g on E(x). This is the application of a general theorem (see Algebra).

g is a tensor and so is defined by a family of maps in a holonomic basis of E :

$g(x) = \sum_{ij} g_{aij}(x) e_a^i(x) \otimes e_a^j(x)$ at the transitions : $[g_b] = [\varphi_{ab}(x)]^t [g_a] [\varphi_{ab}(x)]$

g is not a symmetric tensor. The condition $g(x)(\sigma(x)(u), \sigma(v)) = \overline{g(x)(v,u)}$ reads in matrix notation :

$\overline{[u_a]}^t [\sigma_a(x)]^t [g_a(x)] [\sigma_a(x)] \overline{[v_a]} = \overline{[v_a]}^t [g_a(x)] [u_a] = \overline{[v_a]}^t \overline{[g_a(x)][u_a]}$

$= \overline{[u_a]}^t [g_a(x)]^* [v_a]$

$[\sigma_a(x)]^t [g_a(x)] [\sigma_a(x)] = [g_a(x)]^*$

At the transitions this property is preserved :

$[\sigma_b(x)]^t [g_b(x)] [\sigma_b(x)]$

$= [\varphi_{ab}(x)]^* [\sigma(e_a(x))]^t [\varphi_{ba}(x)]^t [\varphi_{ab}(x)]^t [g_a] [\varphi_{ab}(x)] [\varphi_{ba}(x)] [\sigma(e_a(x))] \overline{[\varphi_{ab}(x)]}$

$= [\varphi_{ab}(x)]^* [\sigma(e_a(x))]^t [g_a] [\sigma(e_a(x))] \overline{[\varphi_{ab}(x)]} = [\varphi_{ab}(x)]^* [g_a(x)]^* \overline{[\varphi_{ab}(x)]} = [g_b]^*$ ■

Induced scalar product

Theorem 1956 *A scalar product g on a vector space V induces a scalar product on a vector bundle $E(M, V, \pi)$ iff the transitions maps preserve g*

Proof. Let $E(M, V, \pi)$ a vector bundle on a field K=$\mathbb{R}$ or $\mathbb{C}$, with trivializations $(O_a, \varphi_a)_{a \in A}$ and transitions maps : $\varphi_{ab}(x) \in \mathcal{L}(V; V)$ and V endowed with a map : $g : V \times V \to K$ which is either symmetric bilinear (if K=$\mathbb{R}$) or sesquilinear (if K=$\mathbb{C}$) and non degenerate.

Define : $x \in O_a : g_a(x)(\varphi_a(x, u_a), \varphi_a(x, v_a)) = g(u_a, v_a)$

The definition is consistent iff :

$\forall x \in O_a \cap O_b : g_a(x)(\varphi_a(x, u_a), \varphi_a(x, v_a)) = g_b(x)(\varphi_b(x, u_a), \varphi_b(x, v_a))$

$\Leftrightarrow g(u_b, v_b) = g(u_a, v_a) = g(\varphi_{ba}(x)u_a, \varphi_{ba}(x)v_a)$

that is : $\forall x, \varphi_{ba}(x)$ preserves the scalar product, they are orthogonal (unitary) with respect to g. ■

It is assumed that V is a Banach space, thus endowed with a norm, which induces a norm on each fiber E(x). If this norm is induced on V by a definite

positive scalar product, then V is a Hilbert space, and each fiber is itself a Hilbert space if the transitions maps preserve the scalar product.

Notice that a scalar product on V induces a scalar product on E, but this scalar product is not necessarily defined by a tensor in the complex case. Indeed for this we need also an induced complex structure which is defined iff the transitions maps are real.

24.2.6 Affine bundles

Definition 1957 *An **affine bundle** is a pair of :*

a vector bundle $\overrightarrow{E}\left(M, \overrightarrow{V}, \overrightarrow{\pi}\right)$ with atlas (O_a, φ_a)

a fiber bundle $E\left(M, V, \pi\right)$ where V is an affine space modelled on $\overrightarrow{V}$ with atlas (O_a, ψ_a) such that each fiber $E(x)$ has the structure of an affine space modelled on the vector space $\overrightarrow{E(x)}$

So there is a fiber preserving map :
$$E\left(x\right) \times E\left(x\right) \to \overrightarrow{E}\left(x\right) :: \left(\psi_a\left(x, A\right), \psi_a\left(x, B\right)\right) = \varphi_a\left(x; \overrightarrow{AB}\right)$$
At the transitions :
$$\left(\psi_a\left(x, A_a\right), \psi_a\left(x, B_a\right)\right) = \varphi_a\left(x; \overrightarrow{A_a B_a}\right) = \left(\psi_b\left(x, A_b\right), \psi_b\left(x, B_b\right)\right) = \varphi_b\left(x; \overrightarrow{A_b B_b}\right)$$
$$A_b = \psi_{ba}\left(x, A_a\right), B_b = \psi_{ba}\left(x, B_a\right) \Rightarrow \overrightarrow{A_b B_b} = \overrightarrow{\psi_{ba}\left(x, A_a\right) \psi_{ba}\left(x, B_a\right)}$$
$$\overrightarrow{A_b B_b} = \varphi_{ba}\left(x\right) \overrightarrow{A_a B_a}$$
$$\overrightarrow{\psi_{ba}\left(x, A_a\right) \psi_{ba}\left(x, B_a\right)} = \varphi_{ba}\left(x\right) \overrightarrow{A_a B_a}$$
So the transition maps : ψ_{ba} must be affine maps, with underlying linear map $\varphi_{ba}\left(x\right)$

Taking one element in the fiber E(x) and a basis in $\overrightarrow{E}\left(x\right)$ we have a **frame of reference** :
$$\psi_a\left(x, A\right) = \psi_a\left(x, O\right) + \sum_{i \in I} u^i e_i\left(x\right)$$

24.2.7 Higher order tangent bundle

Definition

TM is a manifold, therefore it has a tangent bundle, denoted $T^2 M$ and called the bitangent bundle which is a manifold with dimension 2x2xdimM. More generally we can define the r order tangent bundle : $T^r M = T\left(T^{r-1} M\right)$ which is a manifold of dimension $2^r \times \dim M$. The set $T^2 M$ can be endowed with different structures.

Theorem 1958 *If M is a manifold with atlas $\left(E, (O_i, \varphi_i)_{i \in I}\right)$ then the bitangent bundle is a vector bundle $T^2 M\left(TM, E \times E, \pi \times \pi'\right)$*

This is the application of the general theorem about vector bundles.

A point in TM is some $u_x \in T_x M$ and a vector at u_x in the tangent space $T_{u_x}\left(TM\right)$ has for coordinates : $(x, u, v, w) \in O_i \times E^3$

Let us write the maps :

$\psi_i : U_i \to M :: x = \psi_i(\xi)$ with $U_i = \varphi_i(O_i) \subset E$

$\psi'_i(\xi) : U_i \times E \to T_x M :: \psi'_i(\xi) u = u_x$

$\psi"_i(\xi)$ is a symmetric bilinear map:

$\psi"_i(\xi, u) : U_i \times E \times E \to T_{u_x} TM$

By derivation of : $\psi'_i(\xi) u = u_x$ with respect to $\xi : U_{u_x} = \psi"_i(\xi)(u, v) + \psi'_i(\xi) w$

The trivialization is : $\Psi_i : O_i \times E^3 \to T^2 M :: \Psi_i(x, u, v, w) = \psi"_i(\xi)(u, v) + \psi'_i(\xi) w$

With $\partial_{\alpha\beta}\psi_i(\xi) = \partial x_{\alpha\beta}, \partial_\alpha \psi_i(\xi) = \partial x_\alpha : U_{u_x} = \sum \left(u^\alpha v^\beta \partial x_{\alpha\beta}(u, v) + w^\gamma \partial x_\gamma \right)$

The **canonical flip** is the map : $\kappa : T^2 M \to T^2 M :: \kappa(\Psi_i(x, u, v, w)) = (\Psi_i(x, v, u, w))$

Lifts on $T^2 M$

1. Acceleration:

The acceleration of of a path $c : \mathbb{R} \to M$ on a manifold M is the path in $T^2 M$:

$C_2 : t \to T^2 M :: C_2(t) = c"(t) c'(t) = \psi"_i(x(t))(x'(t), x'(t)) + \psi'_i(x(t)) x"(t)$

2. Lift of a path:

The lift of a path $c : \mathbb{R} \to M$ on a manifold M to a path in the fiber bundle TM is a path: $C : \mathbb{R} \to TM$ such that $\pi'(C(t)) = c'(t)$

C is such that :

$C : t \to TM :: C(t) = (c(t), c'(t)) \in T_{c(t)} M$

3. Lift of a vector field:

The lift of a vector field $V \in \mathfrak{X}(TM)$ is a vector field $W \in \mathfrak{X}(T^2 M)$ such that : $\pi'(V(x)) W(V(x)) = V(x)$

W is a projectable vector field on $\mathfrak{X}(T^2 M)$ with projection V.

W is defined through the flow of V: it is the derivative of the lift to T^2M of an integral curve of V

$W(V(x)) = \frac{\partial}{\partial y}(V(y))|_{y=x}(V(x))$

Proof. The lift of an integral curve c(t)=$\Phi_V(x, t)$ of V :

C(t)=$(c(t), c'(t)) = \left(\Phi_V(x, t), \frac{d}{d\theta}\Phi_V(x, \theta)|_{\theta=t} \right) = (\Phi_V(x, t), V(\Phi_V(x, t)))$

Its derivative is in T^2M with components :

$\frac{d}{dt} C(t) = \left(\Phi_V(x, t), V(\Phi_V(x, t)), \frac{d}{dy} V(\Phi_V(x, t))|_{y=x} \frac{d}{d\theta}\Phi_V(x, \theta)|_{\theta=t} \right)$

$= \left(\Phi_V(x, t), V(\Phi_V(x, t)), \frac{d}{dy} V(\Phi_V(x, t))|_{y=x} V(\Phi_V(x, \theta)) \right)$

with t=0 :

$W(V(x)) = \left(x, V(x), \frac{d}{dy} V(\Phi_V(x, 0))|_{y=x} V(x) \right)$ ∎

A vector field $W \in \mathfrak{X}(T^2 M)$ such that : $\pi'(u_x) W(u_x) = u_x$ is called a second order vector field (Lang p.96).

Family of curves

A classic problem in differential geometry is, given a curve c, build a family of curves depending on a real parameter s which are the deformation of c (they are homotopic).

So let be : $c : [0, 1] \to M$ with c(0)=A,c(1)=B the given curve.

We want a map $f : \mathbb{R} \times [0,1] \to M$ with f(s,0)=A,f(s,1)=B

Take a compact subset P of M such that A,B are in P, and a vector field V with compact support in P, such that V(A)=V(B)=0

The flow of V is complete so define : $f(s,t) = \Phi_V(c(t), s)$

$f(s,0) = \Phi_V(A, s) = A, f(s,1) = \Phi_V(B, s) = B$ because V(A)=V(B)=0, so all the paths : $f(s, .) : [0,1] \to M$ go through A and B. The family $(f(s, .))_{s \in \mathbb{R}}$ is what we wanted.

The lift of f(s,t) on TM gives :

$F(s, .) : [0,1] \to TM :: F(s,t) = \frac{\partial}{\partial t} F(s,t) = V(f(s,t))$

The lift of V on T^{2}M gives :

$W(u_x) = \frac{\partial}{\partial y}(V(y))|_{y=x}(u_x)$

so for $u_x = V(f(s,t)) : W(V(f(s,t))) = \frac{\partial}{\partial y}(V(y))|_{y=f(s,t)}(V(f(s,t)))$

But : $\frac{\partial}{\partial s} V(f(s,t)) = \frac{\partial}{\partial y}(V(y))|_{y=f(s,t)} \frac{\partial}{\partial s} f(s,t) = \frac{\partial}{\partial y}(V(y))|_{y=f(s,t)}(V(f(s,t)))$

So we can write : $W(V(f(s,t))) = \frac{\partial}{\partial s} V(f(s,t))$: the vector field W gives the transversal deviation of the family of curves.

24.3   Principal fiber bundles

A principal bundle is a fiber bundle the standard fiber of which is a group.

24.3.1   Definitions

G-bundle

Definition 1959 *A fiber bundle $E(M, V, \pi)$ with an atlas $(O_a, \varphi_a)_{a \in A}$ has a* **G-bundle structure** *if there is :*
- a Lie group G (on the same field as E)
- a left action of G on the standard fiber $V : \lambda : G \times V \to V$
- a family $(g_{ab})_{a,b \in A}$ of maps : $g_{ab} : (O_a \cap O_b) \to G$
such that : $\forall x \in O_a \cap O_b, p = \varphi_a(x, u_a) = \varphi_b(x, u_b) \Rightarrow u_b = \lambda(g_{ba}(x), u_a)$

which means :
$\varphi_{ba} : (O_a \cap O_b) \times V \to V :: \varphi_{ba}(x, u) = \lambda(g_{ba}(x), u)$
All maps are assumed to be of the same class r.

Notation 1960 *$E = M \times_G V$ is a G-bundle with base M, standard fiber V and left action of the group G*

Example : a vector bundle $E(M, V, \pi)$ has a G-bundle structure with G=GL(V)

Principal bundle

Definition 1961 *A* **principal (fiber) bundle** *$P(M, G, \pi)$ is a G-bundle with standard fiber G itself and the left action the left translation.*

So we have :

i) 3 manifolds : the total bundle P, the base M, a Lie group G, all manifolds on the same field

ii) a class r surjective submersion : $\pi : P \to M$

iii) an atlas $(O_a, \varphi_a)_{a \in A}$ with an open cover $(O_a)_{a \in A}$ of M and a set of diffeomorphisms

$\varphi_a : O_a \times G \subset M \times G \to \pi^{-1}(O_a) \subset P :: p = \varphi_a(x, g)$.

iv) a set of class r maps $(g_{ab})_{a,b \in A}$ $g_{ab} : (O_a \cap O_b) \to G$ such that :

$\forall p \in \pi^{-1}(O_a \cap O_b), p = \varphi_a(x, g_a) = \varphi_b(x, g_b) \Rightarrow g_b = g_{ba}(x) g_a$

meeting the cocycle conditions : $\forall a, b, c \in A : g_{aa}(x) = 1; g_{ab}(x) g_{bc}(x) = g_{ac}(x)$

Right action

One of the main features of principal bundles is the existence of a right action of G on P

Definition 1962 *The right action of G on a principal bundle $P(M, G, \pi)$ with atlas $(O_a, \varphi_a)_{a \in A}$ is the map :*

$\rho : P \times G \to P :: \rho(\varphi_a(x, g), h) = \varphi_a(x, gh)$

So $\pi(\rho(p, g)) = \pi(p)$

ρ does not depend on the trivialization :

Proof. $\rho_a(\varphi_a(x, g_a), h) = \varphi_a(x, g_a h)$

$\rho_b(\varphi_b(x, g_b), h) = \varphi_b(x, g_b h) = \varphi_b(x, \varphi_{ba}(x) g_a h) = \varphi_a(x, \varphi_{ab}(x) \varphi_{ba}(x) g_a h) = \rho_a(\varphi_a(x, g_a), h)$ ∎

This action is free and for $p \in \pi^{-1}(x)$ the map $\rho(p, .) : G \to \pi^{-1}(x)$ is a diffeomorphism. The orbits are the sets $\pi^{-1}(x)$.

A principal bundle can be defined equivalently by a right action on a fibered manifold :

Theorem 1963 *(Kolar p.87) Let $P(M, \pi)$ a fibered manifold, and G a Lie group which acts freely on P on the right, such that the orbits of the action $\{\rho(p, g), g \in G\} = \pi^{-1}(x)$. Then $P(M, G; \pi)$ is a principal bundle.*

Remark : this is still true with P infinite dimensional, if all the manifolds are modeled on Banach spaces as usual.

The trivializations are :

O_a = the orbits of the action. Then pick up any $p_a = \tau_a(x)$ in each orbit and define the trivializations by :

$\varphi_a(x, g) = \rho(p_a, g) \in \pi^{-1}(x)$

Theorem 1964 *if $f : N \to M$ is a smooth map, then the pull back f^*P of the principal bundle $P(M, G; \pi)$ is still a principal bundle.*

The right action is an action of G on the manifold P, so we have the usual identities with the derivatives :

$\rho'_p(p, 1) = \Im_P$

$\rho'_g(p, g) = \rho'_g(\rho(p, g), 1)L'_{g^{-1}}(g) = \rho'_p(p, g)\rho'_g(p, 1)R'_{g^{-1}}(g)$

$\left(\rho'_p(p, g)\right)^{-1} = \rho'_p(\rho(p, g), g^{-1})$

and more with an atlas $(O_a, \varphi_a)_{a \in A}$ of E

$\pi\left(\rho\left(p, g\right)\right) = \pi\left(p\right) \Rightarrow$

$\pi'\left(\rho\left(p, g\right)\right)\rho'_p(p, g) = \pi'\left(p\right)$

$\pi'\left(\rho\left(p, g\right)\right)\rho'_g(p, g) = 0$

$\rho'_g(p, 1) = \varphi'_{ag}\left(x, g\right)\left(L'_g 1\right)$

Proof. Take : $p_a\left(x\right) = \varphi_a\left(x, 1\right) \Rightarrow \varphi_a\left(x, g\right) = \rho\left(p_a\left(x\right), g\right)$

$\Rightarrow \varphi'_{ag}\left(x, g\right) = \rho'_g\left(p_a\left(x\right), g\right) = \rho'_g(\rho(p_a\left(x\right), g), 1)L'_{g^{-1}}(g) = \rho'_g(p, 1)L'_{g^{-1}}(g)$

■

$\rho'_g(\varphi\left(x, h\right), g) = \rho'_g(\varphi\left(x, hg\right), 1)L'_{g^{-1}}(g) = \varphi'_{ag}\left(x, hg\right)\left(L'_{hg}1\right)L'_{g^{-1}}(g)$

Sections

Definition 1965 *A section $S(x)$ on a principal bundle $P(M,G,\pi)$ with an atlas $(O_a, \varphi_a)_{a \in A}$ is a family of maps : $\sigma_a : O_a \to G$ such that :*

$\forall a \in A, x \in O_a : S\left(x\right) = \varphi_a\left(x, \sigma_a\left(x\right)\right)$

$\forall a, b \in A, O_a \cap O_b \neq \varnothing, \forall x \in O_a \cap O_b : \sigma_b\left(x\right) = g_b = g_{ba}\left(x\right)\sigma_a\left(x\right)$

Theorem 1966 *(Giachetta p.172) A principal bundle admits a global section iff it is a trivial bundle.*

Definition 1967 *A **holonomic map** of a principal bundle $P(M,G,\pi)$ with an atlas $(O_a, \varphi_a)_{a \in A}$ is a family of maps : $p_a : O_a \to P :: p_a\left(x\right) = \varphi_a\left(x, 1\right)$. At the transitions : $p_b\left(x\right) = \rho\left(p_a, g_{ab}\left(x\right)\right)$.*

so it has not the same value at the intersections : *this is not a section.*

$x \in O_a \cap O_b : p_b\left(x\right) = \varphi_b\left(x, 1\right) = \varphi_a\left(x, g_{ab}\left(x\right)\right) = \rho\left(p_a, g_{ab}\left(x\right)\right)$

A section on the principal bundle is then defined by : $S(x) = \rho\left(p_a\left(x\right), \sigma_a\left(x\right)\right)$

24.3.2 Tangent space of a principal bundle

Tangent bundle

The tangent bundle $TG = \cup_{g \in G} T_g G$ of a Lie group is still a Lie group with the actions :

$M : TG \times TG \to TG :: M(X_g, Y_h) = R'_h(g)X_g + L'_g(g)Y_h \in T_{gh}G$

$\Im : TG \to TG :: \Im(X_g) = -R'_{g^{-1}}(1) \circ L'_{g^{-1}}(g)X_g = -L'_{g^{-1}}(g) \circ R'_{g^{-1}}(g)X_g \in T_{g^{-1}}G$

Identity : $X_1 = 0_1 \in T_1 G$

We have the general results :

Theorem 1968 *(Kolar p.99) The tangent bundle TP of a principal fiber bundle $P(M,G,\pi)$ is still a principal bundle $TP(TM, TG, \pi \times \pi')$.*

The vertical bundle VP is :

a trivial vector bundle over P : $VP(P, T_1G, \pi) \simeq P \times T_1G$

a principal bundle over M with group TG : $VP(M, TG, \pi)$

With an atlas $(O_a, \varphi_a)_{a \in A}$ of P the right action of TG on TP is : (ρ, ρ') :
$$TP \times TG \to TP :: (\rho, \rho')((p, v_p), (g, v_g)) = (\rho(p,g), \rho'(p,g)(v_p, v_g))$$
$$(\rho, \rho')((\varphi(x,h), \varphi'(x,h)(v_x, v_h)), (g, v_g))$$
$$= (\varphi(x, hg), \varphi'(p, gh)(v_x, (R'_g h) v_h + (L'_h g) v_g))$$

Fundamental vector fields

Fundamental vector fields are defined as for any action of a group on a manifold, with same properties.

Theorem 1969 *On a principal bundle $P(M,G,\pi)$ with atlas $(O_a, \varphi_a)_{a \in A}$ and right action ρ :*

i) The map : $\zeta : T_1 G \to \mathfrak{X}(VP) :: \zeta(X)(p) = \rho'_g(p,1) X = \varphi'_{ag}(x,g)(L'_g 1) X$ is linear and does not depend on the trivialization,
$$\zeta(X) = \rho_*(L'_g(1) X, 0)$$
$$\rho'_p(p,g)(\zeta(X)(p)) = \zeta(Ad_{g^{-1}} X)(\rho(p,g))$$

*ii) The **fundamental vector fields** on P are defined, for any fixed X in $T_1 G$, by :$\zeta(X) : M \to VP :: \zeta(X)(p) = \rho'_g(p,1) X$; They belong to the vertical bundle, and span an integrable distribution over P, whose leaves are the connected components of the orbits,*
$$\forall X, Y \in T_1 G : [\zeta(X), \zeta(Y)]_{VP} = \zeta([X,Y]_{T_1 G})$$

iii) The Lie derivative of a section S on P along a fundamental vector field $\zeta(X)$ is : $\pounds_{\zeta(X)} S = \zeta(X)(S)$

Component expressions

Theorem 1970 *A vector $v_p \in T_p P$ can be written $v_p = \sum_\alpha v_x^\alpha \partial x^\alpha + \zeta(u_g)(p)$ where $u_g \in T_1 G$*

At the transitions : $u_{bg} = u_{ag} + Ad_{g_a^{-1}} L'_{\varphi_{ba}^{-1}}(\varphi_{ba}) \varphi'_{ba}(x) v_x$

Proof. With the basis $(\varepsilon_i)_{i \in I}$ of the Lie algebra of G $(L'_g 1) \varepsilon_i$ is a basis of $T_g G$.

With an atlas $(O_a, \varphi_a)_{a \in A}$ of P a basis of $T_p P$ at $p = \varphi_a(x,g)$ is :
$$\partial x_\alpha = \varphi'_{ax}(x,g) \partial \xi_\alpha, \partial g_i = \varphi'_{ag}(x,g)(L'_g 1) \varepsilon_{ii} = \zeta(\varepsilon_i)(p)$$
$$(p, v_p) \in T_p P : v_p = \sum_\alpha v_x^\alpha \partial x^\alpha + \sum_i v_g^i \zeta(\varepsilon_i)(p) = \sum_\alpha v_x^\alpha \partial x_\alpha + \zeta(v_g)(p)$$
where $v_g = \sum_i v_g^i \varepsilon_i \in T_1 G$

At the transitions :
$$v_{bg} = (\varphi_{ba}(x) g_a)'(v_x, v_{ag}) = R'_{g_a}(\varphi_{ba}(x)) \varphi'_{ba}(x) v_x + L'_{\varphi_{ba}(x)}(g_a) v_{ag}$$
$$u_{bg} = L'_{g_b^{-1}} g_b \left(R'_{g_a}(\varphi_{ba}) \varphi'_{ba}(x) v_x + L'_{\varphi_{ba}}(g_a) L'_{g_a}(1) u_{ag} \right)$$
$$u_{bg} = L'_{g_b^{-1}}(g_b) \left(R'_{g_b}(1) R'_{\varphi_{ba}^{-1}}(\varphi_{ba}) \varphi'_{ba} v_x + L'_{g_b}(1) L'_{g_a^{-1}}(g_a) L'_{g_a}(1) u_{ag} \right)$$

$$u_{bg} = Ad_{g_b^{-1}} R'_{\varphi_{ba}^{-1}}(\varphi_{ba}) L'_{\varphi_{ba}}(1) L'_{\varphi_{ba}^{-1}}(\varphi_{ba}) \varphi'_{ba}(x) v_x + u_{ag}$$
$$u_{bg} = Ad_{g_b^{-1}} Ad_{\varphi_{ba}} L'_{\varphi_{ba}^{-1}}(\varphi_{ba}) \varphi'_{ba}(x) v_x + u_{ag}$$
$$u_{bg} = Ad_{g_a^{-1}} L'_{\varphi_{ba}^{-1}}(\varphi_{ba}) \varphi'_{ba}(x) v_x + u_{ag} \quad \blacksquare$$

Component expression of the right action of TG on TP:

Theorem 1971 *The right action of TG on the tangent bundle reads:*
$$(\rho, \rho') \left((\varphi(x,h), \varphi'(x,h) v_x + \zeta(u_h) p) \left(g, (L'_g 1) u_g \right) \right)$$
$$= \left(\rho(p,g), \varphi'_x(x,hg) v_x + \zeta\left(Ad_{g^{-1}} u_h + u_g \right) (\rho(p,g)) \right)$$

Proof. $(\rho, \rho') \left((\varphi(x,h), \varphi'(x,h) v_x + \zeta(u_h) p) \left(g, (L'_g 1) u_g \right) \right)$
$$= \left(\rho(p,g), \varphi'_x(x,hg) v_x + \varphi'_g(x,hg) \left((R'_g h)(L'_h 1) u_h + (L'_h g)(L'_g 1) u_g \right) \right)$$
$$\varphi'_g(x,gh) \left((R'_g h)(L'_h 1) u_h + (L'_h g)(L'_g 1) u_g \right)$$
$$= \varphi'_g(x,hg) \left(L'_{hg} 1 \right) L'_{(hg)^{-1}}(hg) \left((R'_g h)(L'_h 1) u_h + (L'_h g)(L'_g 1) u_g \right)$$
$$= \zeta \left(L'_{(hg)^{-1}}(hg) \left((R'_g h)(L'_h 1) u_h + (L'_h g)(L'_g 1) u_g \right) \right) (\rho(p,g))$$
$$L'_{(hg)^{-1}}(hg) \left((R'_g h)(L'_h 1) u_h + (L'_h g)(L'_g 1) u_g \right)$$
$$= L'_{(hg)^{-1}}(hg) \left(\left(R'_{hg}(1) R'_{h^{-1}}(h)(L'_h 1) u_h + L'_{hg}(1) L'_{g^{-1}}(g)(L'_g 1) u_g \right) \right)$$
$$= L'_{(hg)^{-1}}(hg) R'_{hg}(1) R'_{h^{-1}}(h)(L'_h 1) u_h + L'_{(hg)^{-1}}(hg) L'_{hg}(1) u_g$$
$$= Ad_{(hg)^{-1}} Ad_h u_h + u_g = Ad_{g^{-1}} u_h + u_g \quad \blacksquare$$

24.3.3 Morphisms

Morphisms

Definition 1972 *A **principal bundle morphism** between the principal bundles $P_1(M_1, G_1, \pi_1), P_2(M_2, G_2, \pi_2)$ is a couple (F,χ) of maps : $F : P_1 \to P_2$, and a Lie group morphism $\chi : G_1 \to G_2$ such that :*
$$\forall p \in P_1, g \in G_1 : F(\rho_1(p,g)) = \rho_2(F(p), \chi(g))$$

Then we have a fibered manifold morphism by taking : $f : M_1 \to M_2 ::$
$f(\pi_1(p)) = \pi_2(F(p))$
f is well defined because : $f(\pi_1(\rho_1(p,g))) = \pi_2(F(\rho_1(p,g))) = \pi_2(\rho_2(F(p), \chi(g))) = \pi_2(F(p)) = f \circ \pi_1(p)$

Reduction of a principal bundle by a subgroup

Definition 1973 *A morphism between the principal bundles $P(M, G, \pi_P), Q(M, H, \pi_Q)$ is a **reduction** of P is H is a Lie subgroup of G*

Theorem 1974 *A principal bundle $P(M, G, \pi)$ admits a reduction to $Q(M, H, \pi_Q)$ if it has an atlas with transition maps valued in H*

Proof. If P has the atlas (O_a, φ_a, g_{ab}) the restriction of the trivializations to H reads:

$\psi_a : O_a \times H \to P : \psi_a(x, h) = \varphi_a(x, h)$

For the transitions : $\psi_a(x, h_a) = \psi_b(x, h_b) \Rightarrow h_b = g_{ba}(x) h_a$ and because H is a subgroup : $g_{ba}(x) = h_b h_a^{-1} \in H$ ∎

Gauge transformation

Definition 1975 *An **automorphism** on a principal bundle P with right action ρ is a diffeomorphism $F : P \to P$ such that $:\forall g \in G : F(\rho(p, g)) = \rho(F(p), g)$*

Then $f : M \to M :: f = \pi \circ F$ it is a diffeomorphism on M

Definition 1976 *The **gauge group** of a principal bundle P is the set of fiber preserving automorphisms. Its elements F, called **gauge transformations**, are characterized by a collection of maps : $j_a : O_a \to G$ such that :*
$F(\varphi_a(x, g)) = \varphi_a(x, j_a(x) g), j_b(x) = g_{ba}(x) j_a(x) g_{ba}(x)^{-1}$

$F \in Gau(P) : F \in Aut(P), \pi \circ F = \pi$

A gauge transformation is also called vertical automorphism.

Notice that this is a *left* action by $j_a(x)$

If j depends on x we have a local gauge transformation, if j does not depend on x we have a global gauge transformation.

Proof. i) With an atlas $(O_a, \varphi_a)_{a \in A}$ of P define j from F

With τ_a previously defined as : $\tau_a : \pi^{-1}(O_a) \to G :: p = \varphi_a(\pi(p), \tau_a(p))$

$F(p) = \varphi_a(\pi(F(p)), \tau_a(F(p))) = \varphi_a(\pi(p), \tau_a(F(p)))$

Let us define :

$J_a(p) = \tau_a(F(p))$

$J_a(\rho(p, g)) = \tau_a(F(\rho(p, g))) = \tau_a(\rho(F(p), g)) = \tau_a(F(p)) g = J_a(p) g$

and define : $j_a(x) = J(\varphi_a(x, 1))$

$J_a(p) = J_a(\rho(\varphi_a(x, 1), \tau_a(p))) = J_a(\varphi_a(x, 1)) \tau_a(p) = j_a(x) \tau_a(p)$

$F(p) = \varphi_a(\pi(p), J_a(p)) = \varphi_a(\pi(p), j_a(x) \tau_a(p))$

$F(\varphi_a(x, g)) = \varphi_a(x, j_a(x) g)$

ii) conversely define F from j :

$F(\varphi_a(x, g)) = \varphi_a(x, j_a(x) g)$

then : $\pi \circ F = \pi$ and

$F(\rho(p, g)) = F(\rho(\varphi_a(x, g_a), g)) = F(\varphi_a(x, g_a g)) = F(\varphi_a(x, j_a(x) g_a g)) = \rho(F(p), g)$

iii) At the transitions the transition maps φ_{ab} are still the same, and j must be such that :

$F(\varphi_a(x, g_a)) = \varphi_a(x, j_a(x) g_a) = F(\varphi_b(x, g_b)) = \varphi_b(x, j_b(x) g_b)$

$j_b(x) g_b = g_{ba}(x) j_a(x) g_a$

$g_b = g_{ba}(x) g_a$

$j_b(x) g_{ba}(x) g_a = g_{ba}(x) j_a(x) g_a \Rightarrow j_b(x) = g_{ba}(x) j_a(x) g_{ba}(x)^{-1}$ ∎

Morphisms and change of trivializations

As for charts on manifolds, trivializations are not unique. However, as the theorem in the fiber bundle subsection tells, the key ingredients are the transition maps. Whenever the manifolds, the open cover and the transition maps are given, the fiber bundle structure is fixed. So a change of trivializations which keeps the same principal bundle structure must be compatible with the transition maps. Thus it must preserve the projection, and be fiber wise. However this is the identity map Id:$P \to P$: the same point in P has different coordinates.

A change of trivialization of the principal bundle, with the same open cover, can be expressed as :

$p = \varphi_a(x, g_a) = \psi_a(x, h_a)$

The condition is that the transition maps g_{ab} are the same :

$p = \psi_a(x, h_a) = \psi_b(x, h_b) \Rightarrow h_b = g_{ba}(x) h_a$

Let us define the map : $F : P \to P :: F(\varphi_a(x, g_a)) = \psi_a(x, h_a)$

The right action is the same (it does not depend on the trivialization) :

$\rho(\varphi_a(x, g_a), g) = \rho(\psi_a(x, h_a), g)$

F is a gauge transformation :

$\forall g \in G : F(\rho(p, g)) = F(\rho(\varphi_a(x, g_a), g)) = F(\varphi_a(x, g_a g)) = \psi_a(x, h_a g) = \rho(F(p), g)$

so there is some $j_a(x) : F(\varphi_a(x, g_a)) = \varphi_a(x, j_a(x) g_a) \Leftrightarrow h_a = j_a(x) g_a$

and $p = \varphi_a(x, g_a) = \psi_a(x, h_a)$ reads : $p = \varphi_a(x, g_a) = \psi_a(x, j_a(x) g_a)$

The condition about the transition maps is met:

$j_b(x) = g_{ba}(x) j_a(x) g_{ba}(x)^{-1}, g_b = g_{ba}(x) g_a$

$\Rightarrow h_b = j_b(x) g_b = g_{ba}(x) j_a(x) g_{ba}(x)^{-1} g_b = g_{ba}(x) j_a(x) g_{ba}(x)^{-1} g_{ba}(x) g_a = g_{ba}(x) j_a(x) g_a = g_{ba}(x) h_a$

So any change of trivialization can be written as : $p = \varphi_a(x, g_a) = \psi_a(x, j_a(x) g_a)$. This is a local change of gauge, and it is defined by a section of P : $J_a(x) = \varphi_a(x, j_a(x))$.

Theorem 1977 *Any change of trivialization on a principal bundle $P(M, V, \pi)$ can be expressed as a gauge transformation.*

24.3.4 Principal bundles of frames

This is the classical, and most important example of principal bundle. It is less obvious than it seems.

Principal bundle of linear frames

Theorem 1978 *The **bundle of linear frames** of a vector bundle $E(M, V, \pi)$ with atlas $(O_a, \varphi_a)_{a \in A}$ is the set of its bases. It has the structure of a principal bundle $P(M, G, \pi_P)$ with trivializations $(O_a, L_a)_{a \in A}$ where G is a subgroup of $G\mathcal{L}(V;V)$ and $L_a \in C_0(O_a; G\mathcal{L}(V;V))$*

Proof. A holonomic basis of E associated to an atlas is defined by : e_{ia} : $O_a \rightarrow E :: e_{ai}(x) = \varphi_a(x, e_i)$ where $(e_i)_{i \in I}$ is a basis of V. At the transitions : $e_{ib}(x) = \varphi_b(x, e_i) = \varphi_a(x, \varphi_{ab}(x) e_i)$ where $\varphi_{ab}(x) \in GL(V, V)$.

The crux of a fiber bundle is the transitions maps $(\varphi_{ab}(x))_{a \in A}$ . . Because of the cocycle conditions : $\varphi_{aa}(x) = Id_V, \varphi_{ba}(x) = \varphi_{ab}(x), \varphi_{ab}(x) \circ \varphi_{bc}(x) = \varphi_{ac}(x)$. Let G be the subset of $G\mathcal{L}$(V;V) generated by $\{\varphi_{ab}(x), a, b \in A, x \in M\}$ whenever they are defined.

Any basis of E is defined by a family of section $(S_i)_{i \in I} \in \mathfrak{X}(E)^I$ such that $(S_i(x))_{i \in I}$ are independant vectors in E(x). Equivalently it is defined by a family $(L_a)_{a \in A}$ of maps $L_a \in C_0(O_a; G\mathcal{L}(V; V))$ such that $S_i(x) = L_a(x) e_{ai}(x)$ and at the transitions :

for $x \in O_a \cap O_b : L_b(x) = g_{ba}(x) \circ L_a(x) \Leftrightarrow L_b(x) \circ L_a(x)^{-1} \in G$

P is the set of such families. ■

Notice that G is not necessarily identical to $G\mathcal{L}$(V;V). As a vector bundle E is characterized by a localization of a copy of V at each point of M, and by the group V. Usually the maps L_a do not belong to G, only the products : $L_b(x) \circ L_a(x)^{-1}$

As a particular case the bundle of linear frames of a n dimensional manifold on a field K is the bundle of frames of its tangent bundle. It has the structure of a principal bundle GL$(M, GL(K, n), \pi_{GL})$. A basis $(e_i)_{i=1}^n$ of T_xM is deduced from the holonomic basis $(\partial x_\alpha)_{\alpha=1}^n$ of TM associated to an atlas of M by an endomorphism of K^n represented by a matrix of GL(K,n). If its bundle of linear frame admits a global section, it is trivial and M is parallelizable.

Principal bundle of orthonormal frames

Theorem 1979 *The **bundle of orthonormal frames** on a m dimensional real manifold (M,g) endowed with a scalar product is the set of its orthogonal bases. It has the structure of a principal bundle O(M,O(ℝ,r,s),π) if g has the signature (r,s) with r+s=m.*

Proof. There is always an orthonormal basis at each point, by diagonalization of the matrix of g. These operations are continuous and differentiable if g is smooth. So at least locally we can define charts such that the holonomic bases are orthonormal. The transitions maps belong to O(ℝ,r,s) because the bases are orthonormal. We have a reduction of the bundle of linear frames. ■

If M is orientable, it is possible to define a continuous orientation, and then a principal bundle $SO(M, SO(\mathbb{R}, r, s), \pi)$

$O(\mathbb{R}, r, s)$ has four connected components, and two for $SO(\mathbb{R}, r, s)$. One can always restrict a trivialization to the connected component $SO_0(\mathbb{R}, r, s)$ and then the frames for the other connected component are given by a global gauge transformation.

24.4 Associated bundles

In a fiber bundle, the fibers $\pi^{-1}(x)$, which are diffeomorphic to V, have usually some additional structure. The main example is the vector bundle where the bases $(e_i(x))_{i \in I}$ are chosen to be some specific frame, such that orthonormal. So we have to combine in the same structure a fiber bundle with its fiber content given by V, and a principal bundle with its organizer G. This is the basic idea about associated bundles. They will be defined in the more general setting, even if the most common are the associated vector bundles.

24.4.1 General associated bundles

Associated bundle

Definition 1980 *An associated bundle is a structure which consists of :*
i) a principal bundle P(M,G,π_P) with right action $\rho : P \times G \to G$
ii) a manifold V and a left action $\lambda : G \times V \to V$ which is effective
iii) an action of G on PxV defined as : $\Lambda : G \times (P \times V) \to P \times V ::$
$\Lambda(g,(p,u)) = \left(\rho(p,g), \lambda\left(g^{-1}, u\right)\right)$
iv) the equivalence relation $\sim$ on PxV : $\forall g \in G : (p,u) \sim \left(\rho(p,g), \lambda\left(g^{-1}, u\right)\right)$
*The **associated bundle** is the set $E = (P \times V)/\sim$*
The projection $\pi_E : E \to M :: \pi_E\left([p,u]_\sim\right) = \pi_P(p)$

Notation 1981 $P[V, \lambda]$ *is the associated bundle, with principal bundle P, fiber V and action $\lambda : G \times V \to V$*

1. The power -1 comes from linear frames : coordinates change according to the inverse of the matrix to pass from one frame to the other.
2. The action Λ of G on E does not change the projection x in M.

The key point to understand associated bundles is that G is not really involved in the set E=P[V,λ]. One can perfectly work with a fiber bundle defined through one representative of the class of equivalence and, obviously, the simplest is to use a holonomic map on P for this purpose. So, practically, we have the following (from Kolar p.90) :

Theorem 1982 *For any principal bundle $P(M, G, \pi_P)$ with atlas $(O_a, \varphi_a)_{a \in A}$,holonomic maps $p_a(x) = \varphi_a(x, 1)$ and transition maps $g_{ab}(x)$, manifold V and effective left action : $\lambda : G \times V \to V$ there is an associated bundle E=P[V,λ].*
E is the quotient set : PxV/$\sim$ with the equivalence relation $(p,u) \sim \left(\rho(p,g), \lambda\left(g^{-1}, u\right)\right)$
If P,G, V are class r manifolds, then there is a unique structure of class r manifold on E such that the map :
$\Pr : P \times V \to E :: \Pr(p,u) = [p,u]_\sim$ *is a submersion*
E is a fiber bundle $E(M, V, \pi_E)$ with atlas $(O_a, \psi_a)_{a \in A}$:

Theorem 1983 *projection* $: \pi_E : E \to M :: \pi_E([p, u]) = \pi_P(p)$
trivializations $\psi_a : O_a \times V \to E :: \psi_a(x, u) = \text{Pr}((\varphi_a(x, 1), u))$
transitions maps $: \psi_a(x, u_a) = \psi_b(x, u_b) \Rightarrow u_b = \lambda(g_{ba}(x), u_a)$

$\psi_a(x, u_a) = \psi_b(x, u_b) \Leftrightarrow \exists g \in G : (\varphi_b(x, 1), u_b) = (\rho(\varphi_a(x, 1), g), \lambda(g^{-1}, u_b))$
$\Rightarrow \varphi_b(x, 1) = \varphi_a(x, g) \Leftrightarrow 1 = g_{ba}(x) g$
$\Rightarrow u_b = \lambda(g_{ba}(x), u_b)$
If P and V are finite dimensional then : $\dim E = \dim M + \dim V$
The canonical projection is invariant by the action on PxV : $\text{Pr}(\rho(p, g), \lambda(g^{-1}, u)) = [p, u]$
Conversely :

Theorem 1984 *(Kolar p.91) For a fiber bundle $E(M, V, \pi)$ endowed with a G-bundle structure with a left action of G on V, there is a unique principal bundle $P(M,G,\pi)$ such that E is an associated bundle $P[V, \lambda]$*

Sections

Definition 1985 *A section of an associated bundle $P[V, \lambda]$ is a pair of a section s on P and a map : $u : M \to V$*

The pair is not unique : any other pair s',u' such that :
$\forall x \in M : \exists g \in G : (\rho(s'(x), g), \lambda(g^{-1}, u'(x))) = (s(x), u(x))$ defines the same section.

With the standard trivialization $(O_a, \psi_a)_{a \in A}$ associated to an atlas of P $(O_a, \varphi_a)_{a \in A} : p_a(x) = \varphi_a(x, 1)$ a section S on $P[V, \lambda]$ is defined by a collection of maps : $u_a : O_a \to V :: S(x) = \psi_a(x, u_a(x)) = [p_a(x), u_a(x)]$ such that at the transitions :
$x \in O_a \cap O_b : u_b(x) = \lambda(g_{ba}(x), u_a(x))$

Morphisms

Morphisms and gauge transformations

As gauge transformations are one of the key ingredient of gauge theories in physics they deserve some comments.

1. Whenever a physical quantity is "located" (usually in the universe) over some space modeled as a manifold M, the natural model is that of a fiber bundle with base M. But we need to relate geometrical quantities to physical measures, meaning eventually figures. The measures are done with respect to some frame (equivalently in some gauge), and expressed in some space V. If the frames can be characterized by a group, then they are modeled by a principal bundle, and we have an associated bundle.

If we have two observers 1,2 which measure the same phenomenon, using frames $p_1, p_2 \in P$, they will get measures $u_1, u_2 \in V$. As they are related

to the same physical phenomenon they must be done at the same location, so $\pi(p_1) = \pi(p_2) = x$. Anyway $u_1, u_2 \in V$ are not expected to be equal, but the difference should come only from the fact that the frames are not the same: $p_2 = \rho(p_1, g)$. The basic assumption in an associated bundle is that u_1, u_2 are deemed equal $(p_1, u_1) \sim (p_2, u_2)$ if $u_2 = \lambda\left(g^{-1}, u_1\right)$: it assumes that a change : $p \to \rho(p, g)$ implies a change $\lambda\left(g^{-1}, u\right)$ for some λ.

In all these procedures the observers use the same trivializations φ_a and charts in V.

With an atlas $(O_a, \varphi_a)_{a \in A}$ of P, if both undergo the same change of gauge :
$p_1 = \varphi_a(x, g_1) \to F(p_1) = \varphi_a(x, j_a(x) g_1)$,
$p_2 = \varphi_a(x, g_2) \to F(p_2) = \varphi_a(x, j_a(x) g_2)$
they will get new measures v_1, v_2

If their measures were equivalent before the change, they should stay equivalent after :

If $(p_1, u_1) \sim (p_2, u_2)$ then $\exists g \in G : p_2 = \rho(p_1, g), u_2 = \lambda\left(g^{-1}, u_1\right)$
and : $F(p_2) = \varphi_a(x, j_a(x) g_1 g) = \rho(F(p_1), g)$
Thus : $(F(p_1), v_1) \sim (F(p_2), v_2) = (\rho(F(p_1), g), v_2) \Leftrightarrow v_2 = \lambda\left(g^{-1}, v_1\right)$

This leads to define automorphisms on an associated bundle, induced by a change of gauge, as maps :

$E \to E :: (F(p), f(u))$ such that f is equivariant by G: $\forall g \in G, u \in V$: $f(\lambda(g, u)) = \lambda(g, f(u))$

Definition of a morphism between associated bundles

Definition 1986 *(Kolar p.92) : A morphism between the associated bundles $P_1[V_1, \lambda_1], P_2[V_2, \lambda_2]$ is a map : $\Phi : P_1 \to P_2 :: \Phi(p, u) = (F(p), f(u))$ such that $F : P_1 \to P_2$ is a morphism of principal bundles and $f : V_1 \to V_2$ is equivariant : $\forall g \in G, u \in V_1 : f(\lambda_1(g, u)) = \lambda_2(\chi(g), f(u))$*

A morphism F between the principal bundles $P_1(M_1, G_1, \pi_1), P_2(M_2, G_2, \pi_2)$ is a couple (F, χ) with $\chi : G_1 \to G_2$ a Lie group morphism such that : $\forall p \in P_1, g \in G_1 : F(\rho_1(p, g)) = \rho_2(F(p), \chi(g))$

With a morphism of principal bundle $F : P_1 \to P_2$ and a an associated bundle $P_1[V, \lambda_1]$ it is always possible to define an associated bundle $P_2[V, \lambda_2]$ by f=Id and $\lambda_2(\chi(g), u) = \lambda_1(g, u)$

Morphisms and change of trivializations

Another issue is the change of trivializations. Any change of trivialization on a principal bundle can be expressed as a change of gauge. However they do not have the same meaning.

If a physical law is expressed as $L(\xi, \gamma, \eta) = 0$ with ξ, γ, η coordinates of x,g,u, to assert that it is equivariant in a change of charts and trivializations is equivalent to say that it can be expressed in a geometrical fashion as $\mathcal{L}(p, u) = 0$. Physically it means that the model of associated bundle is accurate to model the phenomenon.

If a physical law $\mathcal{L}(p,u) = 0$ is invariant with a change of gauge : $\mathcal{L}(p,u) = 0 \Rightarrow \mathcal{L}\left(\rho(p,g), \lambda\left(g^{-1}, u\right)\right) = 0$ it means that the relation between the frame p and the measure u is accurate.

So in both cases the invariance is the assertion that the specific model used is accurate.

To assert that, whatever the gauge transformation F, there is a G equivariant map $f : V \to V$ such that :

$\mathcal{L}(p,u) = 0 \Rightarrow \mathcal{L}(F(p), f(u)) = 0$

is the application of a general principle in physics : the law of physics should not depend on the observer.

One parameter group of gauge transformation

Theorem 1987 *The infinitesimal generators of one parameter groups of local gauge transformation on a principal bundle are given by sections of the adjoint bundle $P[T_1 G, Ad]$*

Proof. Let keep x fixed in M. A one parameter group of gauge transformation in x is given by :

$\phi : \mathbb{R} \to G : \phi(t+s) = \phi(t)\phi(s)$

G is a Lie group, so the one parameter groups are of the form with $X \in T_1 G$

$\phi_{X_L}(g,t) = (\exp tX)g, \phi_{X_R}(g,t) = g(\exp tX)$

A one parameter group of local gauge transformation is given by a map with an atlas $(O_a, \varphi_a)_{a \in A}$ of P

$X : O_a \to T_1 G$ and then : $j_a(x,t) = \exp tX_a(x), F_t(\varphi_a(x,g)) = \varphi_a(x, (\exp tX_a(x))g)$

and we have :

$\exp tX_b(x) = g_{ba}(x)(\exp tX_a(x)) g_{ba}(x)^{-1}$

$X_b(x) = Ad_{g_{ba}(x)} X_a(x)$

Such a map X is a section $X \in \mathfrak{X}(P[T_1 G, Ad])$ of the adjoint bundle of P.

∎

If a quantity U is represented in the associated bundle $P[V, \lambda]$ so we have for any vector field X :

$\forall t : U = (p, u) \sim (J(-t, p), \lambda(\exp tX, u))$

If we have some law : $\mathcal{L}(p,u) = 0$ it shall be such that :

$\mathcal{L}(p,u) = 0 \Rightarrow \mathcal{L}(J(-t,p), \lambda(\exp tX, u)) = 0$

$\left(\mathcal{L}'_p(p,u) \varphi'_{ag}(x,g) + \mathcal{L}'_u(p,u) \lambda'_g(x,g)\right) X(x) = 0$

$\mathcal{L}'_p(p,u) \varphi'_{ag}(x,g) + \mathcal{L}'_u(p,u) \lambda'_g(x,g) = 0$

Tangent bundles

Theorem 1988 *(Kolar p.99) The tangent bundle TE of the associated bundle E=$P[V, \lambda]$ is the associated bundle*

$TP[TV, \lambda \times \lambda'] \sim TP \times_{TG} TV$

This is straightforward by using the general theorems and is fully compliant with .

Proof. P is a principal bundle $P(M, G, \pi_P)$ with atlas $(O_a, \varphi_a)_{a \in A}$,holonomic maps $p_a(x) = \varphi_a(x, 1)$ and transition maps $g_{ab}(x)$

TP is a principal bundle $TP(TM, TG, \pi \times \pi')$ with right action of TG :
$(\rho, \rho') : TP \times TG \to TP :: (\rho, \rho')(p, v_p)(g, \kappa_g) = (\rho(p, g), \rho'(p, g)(v_p, \kappa_g))$
TV is a manifold with the left action of TG :
$(\lambda, \lambda') : TG \times TV \to TV :: (\lambda, \lambda')(g, \kappa_g)(u, v_u)$
$= \left(\lambda\left(g^{-1}, u\right), \lambda'\left(g^{-1}, u\right)\left(-R'_{g^{-1}}(1)\kappa_g, v_u\right) \right)$
with : $\lambda'\left(g^{-1}, u\right)(\kappa_g, v_u) = \lambda'_g\left(g^{-1}, u\right)\left(g^{-1}\right)' \kappa_g + \lambda'_u(\kappa_g, v_u) v_u$
$= -\lambda'_g\left(g^{-1}, u\right) R'_{g^{-1}}(1) L'_{g^{-1}}(g)\kappa_g + \lambda'_u(\kappa_g, v_u) v_u$
The product PxV has a manifold structure, whose tangent bundle is T(PxV)=TPxTV (see Differential geometry)

So by the general theorem TE is an associated bundle. ∎

The structures are the following (see general theorem above):
TE is the quotient set $TE = T(P \times V) / \sim$ with the equivalence relation :
$\forall (g, \kappa_g) \in TG :$
$(p, u) \sim \Lambda(g, (p, u))$
$[(p, v_p), (u, v_u)] \sim \left(\Lambda(g, (p, u)), \Lambda'_g(g, (p, u))(\kappa_g)(v_p, v_u) \right)$
$[(p, v_p), (u, v_u)]$
$\sim \left[(\rho(p, g), \rho'(p, g)(v_p, \kappa_g)), \left(\lambda\left(g^{-1}, u\right), \lambda'\left(g^{-1}, u\right)\left(-R'_{g^{-1}}(1)L'_{g^{-1}}(g)\kappa_g, v_u\right) \right) \right]$
The projection $Pr : TP \times TV \to TE :: Pr((p, v_p), (u, v_u)) = [(p, v_p), (u, v_u)]$
is a submersion

TE has the structure of a manifold, of a an associated bundle $TP[TV, \lambda \times \lambda']$ and of a fiber bundle $TE(TM, TV, \pi_{TE})$ with atlas $(\{T_x M, x \in O_a\}, \Psi_a)_{a \in A}$
The projection $\pi_{TE} : TE \to TM :: \pi_{TE}([(p, v_p), (u, v_u)]) = \pi'_P(p) v_p = v_x$
is a submersion

The trivializations $\Psi_a : \pi'(\{T_x M, x \in O_a\})^{-1} \times TV \to TE :: \Psi_a((x, v_x), (u_a, v_{au})) = Pr((p_a(x), p'_a(x) v_x), (u_a, v_{au}))$
are the derivatives of the trivializations of E : $\psi_a : O_a \times V \to E :: \psi_a(x, u) = Pr((\varphi_a(x, 1), u))$
$\Psi_a((x, v_x), (u_a, v_{au})) = \psi'_{ax}(x, u_a) v_x + \psi'_{au}(x, u_a) v_{au}$
The transitions maps : $\forall x \in O_a \cap O_b : \exists (g_{ba}(x), \kappa_{ba}(x)) \in TG :$
$\Psi_a((x, v_{ax}), (u_a, v_{au})) = \Psi_b((x, v_{bx}), (u_b, v_{bu}))$
$u_b = \lambda\left(g_{ba}^{-1}(x), u_a\right)$
$v_{bu} = \lambda'\left(g_{ba}^{-1}(x), u_a\right)\left(-R'_{g_{ba}^{-1}}(1) L'_{g_{ba}^{-1}}(g_{ba})\kappa_{ba}(x), v_{au}\right)$

Theorem 1989 *(Kolar p.99) The vertical bundle of the associated bundle $P[V, \lambda]$ is he associated bundle $P[TV, \lambda'_u]$ isomorphic to the G bundle $P \times_G TV$*

The vertical vectors are : $\Psi_a((x, 0), (u_a, v_{au})) = Pr((p_a(x), 0), (u_a, v_{au}))$
With the transitions maps for TE they transform as:
$\Psi_b((x, 0), (u_b, v_{bu})) = \Psi_b\left((x, 0), \left(\lambda\left(g_{ba}^{-1}(x), u_a\right), \lambda'_u\left(g_{ba}^{-1}(x), u_a\right) v_{au}\right)\right) \simeq (p_a(x),)$

557

Fundamental vector fields

The fundamental vector fields corresponding to the action $\Lambda : G \times (P \times V) \to (P \times V)$ are generated by vectors $X \in T_1 G$:

$Z(X)(p, u) = \Lambda'_g(1, (p, u))(X) = \left(\rho'_g(p, 1)X, \lambda'_g(1, u)X\right) =$

They have the properties (cf.Lie groups):

$[Z(X), Z(Y)]_{\mathfrak{X}(TM \times TV)} = Z\left([X, Y]_{T_1 G}\right)$

$\Lambda'_{(p,u)}(g, (p, u))|_{(p,u)=(p_0,u_0)} Z(X)(p_0, u_0) = Z(Ad_g X)(\Lambda(g, (p_0, u_0)))$

The fundamental vector fields span an integrable distribution over PxV, whose leaves are the connected components of the orbits. The orbits are the elements of E=P$[V, \lambda]$, thus :

Theorem 1990 *The fundamental vector fields of E span the vertical bundle VE.*

More precisely :

Any fundamental vector is equivalent to : $\forall (g, \kappa_g) \in TG$:

$Z(X)(p_a, u) = \left((p_a(x), \zeta(X)(p_a)), (u, \lambda'_g(1, u)X)\right)$

$\sim \left((\rho(p_a, g), \rho'(p_a, g)(\zeta(X)(p_a), \kappa_g)),\right.$

$\left.\left(\lambda(g^{-1}, u), \lambda'(g^{-1}, u)\left(-R'_{g^{-1}}(1)L'_{g^{-1}}(g)\kappa_g, \lambda'_g(1, u)X\right)\right)\right)$

$= \left(\left(p, \zeta\left(Ad_{g^{-1}}X + L'_{g^{-1}}(g)\kappa_g\right)(p)\right),\right.$

$\left.\left(\lambda(g^{-1}, u), \lambda'_g(1, \lambda(g^{-1}, u))\left(-L'_{g^{-1}}(g)\kappa_g + Ad_{g^{-1}}X\right)\right)\right)$

Take : $\kappa_g = -L'_g(1) Ad_{g^{-1}}X$

$Z(X)(p_a, u) \sim \left((p, 0), \left(\lambda(g^{-1}, u), 2\lambda'_g(1, \lambda(g^{-1}, u)) Ad_{g^{-1}}X\right)\right)$

Conversely any vertical vector is equivalent to a fundamental vector :

$((\varphi_a(x, g_a), 0), (u_a, v_{au})) \sim Z(X)(p_a, u) = Z\left(\frac{1}{2}Ad_{g_a}\lambda'_g(1, u_a)^{-1} v_{au}\right)(p_a, \lambda(g, u_a))$

24.4.2 Associated vector bundles

They are associated bundle with V a vector space. They are the most usual associated bundles.

Definition

Definition 1991 *An **associated vector bundle** is an associated bundle $P[V, r]$ where (V,r) is a continuous representation of G on a Banach vector space V on the same field as P*

So we have the equivalence relation on PxV :

$(p, u) \sim \left(\rho(p, g), r(g^{-1})u\right)$

Vector bundle structure

Theorem 1992 *An associated vector bundle $P[V,r]$ on a principal bundle P with atlas $(O_a, \varphi_a)_{a \in A}$ is a vector bundle $E(M, V, \pi_E)$ with atlas $(O_a, \psi_a)_{a \in A}$ with $\psi_a(x, u_a) = [(\varphi_a(x, 1), u_a)]_\sim$ and transition maps : $\psi_{ba}(x) = r(g_{ba}(x))$ where $g_{ba}(x)$ are the transition maps on $P : U = \psi_a(x, u_a) = \psi_b(x, u_b) \Leftrightarrow u_b = r(g_{ba}(x)) u_a$*

Proof. This is the application of the theorem for general associated fiber bundles.

The transition maps are linear in u : $u_b = \lambda(g_{ba}(x), u_a) = r(g_{ba}(x)) u_a$ ∎

The "p" in the couple (p,u) represents the frame, and u acts as the components. So the vector space operations are completed in E fiberwise and in the same frame :

Let $U_1 = (p_1, u_1), U_2 = (p_2, u_2), \pi(p_1) = \pi(p_2)$

There is some g such that: $p_2 = \rho(p_1, g)$ thus : $U_2 \sim (\rho(p_1, g^{-1}), r(g) u_2) = (p_1, r(g) u_2)$

and : $\forall k, k' \in K : kU_1 + k'U_2 = (p_1, ku_1 + k'r(g)u_2)$

Associated vector bundles of linear frames

This is the converse of the previous theorem.

Theorem 1993 *Any vector bundle $E(M, V, \pi)$ has the structure of an associated bundle $P[V,r]$ where $P(M, G, \pi_P)$ is the bundle of its linear frames and r the natural action of the group G on V.*

P is the bundle of linear frames P=$\{f_i(x), i \in I, x \in M\}$ with standard fiber G and trivializations $(O_a, L_a)_{a \in A}$.

In a trivialization of E :

$p = L_a(x), U = \varphi(x, L_a(x)(u)) \sim (\varphi(x, \rho(p, g)(L_a(x)^{-1} u))) = r(L_a(x)^{-1}, u)$

If M is a m dimensional manifold on a field K, its principal bundle $GL(M, GL(K, m), \pi)$ of linear frames gives, with the standard representation $(K^m, \imath)$ of $GL(K, m)$ the associated vector bundle $GL[K^m, \imath]$. This is the usual definition of the tangent bundle, where one can use any basis of the tangent space $T_x M$.

Similarly if M is endowed with a metric g of signature (r,s) its principal bundle $O(M, O(K, r, s), \pi)$ of orthonormal frames gives, with the standard unitary representation $(K^m, \jmath)$ of $O(K, r, s)$ the associated vector bundle $O[K^m, \jmath]$.

Conversely if we have a principal bundle $P(M, O(K, r, s), \pi)$, then with the standard unitary representation $(K^m, \jmath)$ of $O(K, r, s)$ the associated vector bundle E=$P[K^n, \jmath]$ is endowed with a scalar product corresponding to O(K,r,s).

A section U$\in \mathfrak{X}(P[K^m, \jmath])$ is a vector field whose components are defined in orthonomal frames.

Holonomic basis

Definition 1994 *A **holonomic basis** $e_{ai}(x) = (p_a(x), e_i)$ on an associated vector bundle $P[V, r]$ is defined by a holonomic map $p_a(x) = \varphi_a(x, 1)$ of P and a basis $(e_i)_{i \in I}$ of V. At the transitions : $e_{bi}(x) = r(g_{ba}(x)) e_{ai}(x)$*

This is not a section.
$e_{bi}(x) = (p_b(x), e_i) = (\rho(p_a(x), g_{ab}(x)), e_i)) = (p_a(x), r(g_{ba}(x)) e_i)$

Theorem 1995 *A section U on an associated vector bundle $P[V, r]$ is defined with respect to a holonomic basis by a family of maps : $u_a : O_a \to V$ such that : $u_b(x) = r(g_{ba}(x)) u_a(x)$ and $U(x) = \sum_{i \in I} u_a^i(x) e_{ai}(x)$*

Warning ! $u_b^i(x) = \sum_{ij} [r(g_{ab}(x))]_j^i u_a^j(x)$
Proof. $U(x) = \psi_a(x, u_a(x)) = \psi_b(x, u_b(x)) \Leftrightarrow u_b(x) = r(g_{ba}(x)) u_a(x)$
$\psi_a(x, u_a(x)) = \sum_i u_a^i(x) \psi_a(x, e_i) = \sum_i u_a^i(x) e_{ai}(x) = \sum_i u_b^i(x) e_{bi}(x) = \sum_i u_b^i(x) r(g_{ba}(x)) e_{ai}(x)$
$\sum_i u_a^i(x) e_{ai}(x) = \sum_{ij} [r(g_{ba}(x))]_j^i u_b^j(x) e_{ai}(x)$
$u_b^i(x) = \sum_{ij} [r(g_{ab}(x))]_j^i u_a^j(x)$ ∎

Tensorial bundle

Because E has a vector bundle structure we can import all the tensorial bundles defined with V : $\otimes_s^r V, V^*, \wedge_r V, ...$The change of bases can be implemented fiberwise, the rules are the usual with using the matrices of r(g) in V. They are not related to any holonomic map on M : both structures are fully independant. Everything happens as if we had a casual vector space, copy of V, located at some point x in M.

The advantage of the structure of associated vector bundle over that of simple vector bundle is that we have at our disposal the mean to define any frame p at any point x through the principal bundle structure. So the picture is fully consistent.

Complexification of an associated vector bundle

Theorem 1996 *The complexified of a real associated vector bundle $P[V, r]$ is $P[V_{\mathbb{C}}, r_{\mathbb{C}}]$*

with
$V_{\mathbb{C}} = V \oplus iV$
$r_{\mathbb{C}} : G \to V_{\mathbb{C}} :: r_{\mathbb{C}}(g)(u + iv) = r(g) u + ir(g) v$ so $r_{\mathbb{C}}$ is complex linear
A holonomic basis of $P[V, r]$ is a holonomic basis of $P[V_{\mathbb{C}}, r_{\mathbb{C}}]$ with complex components.
Notice that the group stays the same and the principal bundle is still P.

Scalar product

1. General case :

Theorem 1997 *On a complex vector bundle $E = P[V,r]$ associated to $P(M,G,\pi)$ with an atlas $(O_a, \varphi_a)_{a \in A}$ and transition maps g_{ab} of P a scalar product is defined by a family γ_a of hermitian sesquilinear maps on V, such that, for a holonomic basis at the transitions : $\gamma_a(x)(e_i, e_j) = \sum_{kl} \overline{[r(g_{ba}(x))]}_i^k [r(g_{ba}(x))]_j^l \gamma_b(x)(e_k, e_l)$*

Proof. $\widehat{\gamma}_a(x)(e_{ia}(x), e_{ja}(x)) = \gamma_a(x)(e_i, e_j) = \gamma_{aij}(x)$

For two sections U,V on E the scalar product reads : $\widehat{\gamma}(x)(U(x), V(x)) = \sum_{ij} \overline{u_a^i(x)} v_a^j(x) \gamma_{aij}(x)$

At the transitions : $\forall x \in O_a \cap O_b$:

$\gamma(x)(U(x), V(x)) = \sum_{ij} \overline{u_a^i(x)} v_a^j(x) \gamma_{aij}(x) = \sum_{ij} \overline{u_b^i(x)} v_b^j(x) \gamma_{bij}(x)$

with $u_b^i(x) = \sum_j [r(g_{ba}(x))]_j^i u_a^j(x)$

$\gamma_{aij}(x) = \sum_{kl} \overline{[r(g_{ba}(x))]}_i^k [r(g_{ba}(x))]_j^l \gamma_{bkl}(x)$ ∎

2. Induced scalar product :

Conversely if γ is a hermitian sesquilinear maps on V it induces a scalar product on E iff : $[\gamma] = [r(g_{ba}(x))]^* [\gamma] [r(g_{ba}(x))]$ that is iff (V,r) is a unitary representation of G.

Theorem 1998 *For any vector space V endowed with a scalar product γ, and any unitary representation (V,r) of a group G, there is a vector bundle $E = P[V,r]$ associated to any principal bundle $P(M,G,\pi)$, and E is endowed with a scalar product induced by γ.*

Tangent bundle

By applications of the general theorems

Theorem 1999 *The tangent bundle of the associated bundle $P[V,r]$ is the G bundle $TP \times_G TV$*

The vertical bundle is isomorphic to $P \times_G TV$

Ajoint bundle of a principal bundle

Definition 2000 *The **adjoint bundle** of a principal bundle $P(M,G,\pi)$ is the associated vector bundle $P[T_1G, Ad]$.*

With the holonomic maps $p_a(x) = \varphi_a(x,1)$ on P and a basis ε_i of T_1G.:the holonomic basis of $P[T_1G, Ad]$ is $\varepsilon_{ai}(x) = (p_a(x), \varepsilon_i)$ and a section U of $P[T_1G, Ad]$ is defined by a collection of maps : : $u_a : O_a \to T_1G$ such that : $u_b(x) = Ad_{g_{ba}(x)}(u_a(x))$ and $U(x) = \sum_{i \in I} u_a^i(x) \varepsilon_{ai}(x)$

Theorem 2001 *The tangent bundle of $P[T_1G, Ad]$ is the associated bundle $TP[T_1G \times T_1G, Ad \times (Ad)']$ with group TG*

A vector U_q at p=(p,u) reads $w_q = \sum_i v_g^i \varepsilon_{ai}(x) + \sum_\alpha v_x^\alpha \partial x_\alpha$

Theorem 2002 *The vertical bundle of the adjoint bundle is $P \times_G (T_1G \times T_1G)$*

24.4.3 Homogeneous spaces

Homogeneous spaces are the quotient sets of a Lie group G by a one of its subgroup H. They have the structure of a manifold (but not of a Lie group except if H is normal).

Principal fiber structure
   (see Lie groups)
   If H is a subgroup of the group G :
   The quotient set G/H is the set $G/\sim$ of classes of equivalence : $x \sim y \Leftrightarrow \exists h \in H : x = y \cdot h$
   The quotient set $H\backslash G$ is the set $G/\sim$ of classes of equivalence : $x \sim y \Leftrightarrow \exists h \in H : x = h \cdot y$
   They are groups iff H is normal : gH = Hg

Theorem 2003 *For any closed subgroup H of the Lie group G, G has the structure of a principal fiber bundle $G\left(G/H, H, \pi_L\right)$ (resp.$G\left(H\backslash G, H, \pi_R\right)$*

Proof. The homogeneous space G/H (resp. H\G) is a manifold
   The projection $\pi_L : G \to G/H$ (resp $\pi_R : G \to H\backslash G$) is a submersion.
   On any open cover $(O_a)_{a \in A}$ of G/H (resp. H\G), by choosing one element of G in each class, we can define the smooth maps :
   $\lambda_a : O_a \to \pi_L^{-1}(O_a) : \lambda_a(x) \in G$
   $\rho_a : O_a \to \pi_R^{-1}(O_a) : \rho_a(x) \in G$
   the trivializations are :
   $\varphi_a : O_a \times H \to \pi_L^{-1}(O_a) :: g = \varphi_a(x, h) = \lambda_a(x)h$
   $\varphi_a : H \times O_a \to \pi_R^{-1}(O_a) :: g = \varphi_a(x, h) = h\rho_a(x)$
   For $x \in O_a \cap O_b : \lambda_a(x)h_a = \lambda_b(x)h_b \Leftrightarrow h_b = \lambda_b(x)^{-1}\lambda_a(x)h_a$
   $\varphi_{ba}(x) = \lambda_b(x)^{-1}\lambda_a(x) = h_b h_a^{-1} \in H$ ∎
   The right actions of H on G are:
   $\rho(g, h') = \varphi_a(x, hh') = \lambda_a(x)hh'$
   $\rho(g, h') = \varphi_a(x, hh') = hh'\rho_a(x)$
   The translation induces a smooth transitive right (left) action of G on H\G (G/H) :
   $\Lambda : G \times G/H \to G/H :: \Lambda(g, x) = \pi_L(g\lambda_a(x))$
   $P : H\backslash G \times G \to H\backslash G : P(x, g) = \pi_R(\rho_a(x)g)$

Theorem 2004 *(Giachetta p.174) If G is a real finite dimensional Lie group, H a maximal compact subgroup, then the principal fiber bundle $G\left(G/H, H, \pi_L\right)$ is trivial.*

   Examples :
   $O(\mathbb{R}, n)\left(S^{n-1}, O(\mathbb{R}, n-1), \pi\right)$
   $SU(n)\left(S^{2n-1}, SU(n-1), \pi\right)$
   $U(n)\left(S^{2n-1}, U(n-1), \pi\right)$
   $Sp(\mathbb{R}, n)\left(S^{4n-1}, Sp(\mathbb{R}, n-1), \pi\right)$
   where $S^{n-1} \subset \mathbb{R}^n$ is the sphere

Tangent bundle

Theorem 2005 *(Kolar p.96) If H is a closed subgroup H of the Lie group G, the tangent bundle $T(G/H)$ has the structure of a G-bundle $(G \times_H (T_1G/T_1H)) (G/H, T_1G/T_1H, \pi)$*

Which implies the following :
$$\forall v_x \in T_x G/H, \exists Y_H \in T_1 H, Y_G \in T_1 G, h \in H : v_x = L'_{\lambda_a(x)} (1) Ad_h (Y_G - Y_H)$$
Proof. By differentiation of : $g = \lambda_a (x) h$
$$v_g = R'_h (\lambda_a (x)) \circ \lambda'_a (x) v_x + L'_{\lambda_a(x)} (h) v_h = L'_g 1 Y_G$$
$$Y_G = L'_{g^{-1}} (g) R'_h (\lambda_a (x)) \circ \lambda'_a (x) v_x + L'_{g^{-1}} (g) L'_{\lambda_a(x)} (h) L'_h 1 Y_H$$
$$L'_{g^{-1}} (g) R'_h (\lambda_a (x)) = L'_{g^{-1}} (g) R'_{\lambda_a(x)h} (1) R'_{\lambda_a(x)^{-1}} (\lambda_a (x))$$
$$= L'_{g^{-1}} (g) R'_g (1) R'_{\lambda_a(x)^{-1}} (\lambda_a (x)) = Ad_{g^{-1}} \circ R'_{\lambda_a(x)^{-1}} (\lambda_a (x))$$
$$L'_{g^{-1}} (g) \left(L'_{\lambda_a(x)} (h) \right) L'_h 1 = L'_{g^{-1}} (g) \left(L'_{\lambda_a(x)h}(1) L'_{h^{-1}}(h) \right) L'_h 1$$
$$= L'_{g^{-1}} (g) L'_g (1) = Id_{T_1 H}$$
$$Y_G = Ad_{g^{-1}} \circ R'_{g_a(x)^{-1}} (g_a (x)) v_x + Y_H$$
$$v_x = R'_{g_a(x)} (1) Ad_{g_a(x)h} (Y_G - Y_H) = L'_{g_a(x)} (1) Ad_h (Y_G - Y_H) \quad \blacksquare$$

24.4.4 Spin bundles

Spin bundles are common in physics. This is a somewhat complicated topic. See Clifford algebras in the Algebra part and Svetlichny.

Clifford bundles

Theorem 2006 *On a real finite dimensional manifold M endowed with a bilinear symmetric form g with signature (r,s) which has a bundle of orthornormal frames P, there is a vector bundle Cl(TM), called a **Clifford bundle**, such that each fiber is Clifford isomorphic to $Cl(\mathbb{R}, r, s)$. And $O(\mathbb{R}, r, s)$ has a left action on each fiber for which Cl(TM) is a G-bundle : $P \times_{O(\mathbb{R},r,s)} Cl(\mathbb{R}, r, s)$.*

Proof. i) On each tangent space $(T_x M, g(x))$ there is a structure of Clifford algebra $Cl(T_x M, g(x))$. All Clifford algebras on vector spaces of same dimension, endowed with a bilinear symmetric form with the same signature are Clifford isomorphic. So there are Clifford isomorphisms : $T(x) : Cl(\mathbb{R}, r, s) \to Cl(T_x M, g(x))$. These isomorphisms are geometric and do not depend on a basis. However it is useful to see how it works.

ii) Let be :
$(\mathbb{R}^m, \gamma)$ with basis $(\varepsilon_i)_{i=1}^m$ and bilinear symmetric form of signature (r,s)
$(\mathbb{R}^m, \jmath)$ the standard representation of $O(\mathbb{R}, r, s)$
$O(M, O(\mathbb{R}, r, s), \pi)$ the bundle of orthogonal frames of (M,g) with atlas (O_a, φ_a) with transition maps $g_{ba}(x)$

$E = P[\mathbb{R}^m, \jmath]$ the associated vector bundle with holonomic basis : $\varepsilon_{ai}(x) = (\varphi_a (x, 1), \varepsilon_i)$ endowed with the induced scalar product g(x). E is just TM with orthogonal frames. So there is a structure of Clifford algebra $Cl(E(x), g(x))$.

On each domain O_a the maps : $t_a(x) : \mathbb{R}^m \to E(x) : t_a(x)(\varepsilon_i) = \varepsilon_{ai}(x)$ preserve the scalar product, using the product of vectors on both $\mathrm{Cl}(\mathbb{R}^m, \gamma)$ and $\mathrm{Cl}(E(x), g(x)) : t_a(x)(\varepsilon_i \cdot \varepsilon_j) = \varepsilon_{ai}(x) \cdot \varepsilon_{aj}(x)$ the map $t_a(x)$ can be extended to a map : $T_a(x) : \mathrm{Cl}(\mathbb{R}^m, \gamma) \to \mathrm{Cl}(E(x), g(x))$ which an isomorphism of Clifford algebra. The trivializations are :

$\Phi_a(x, w) = T_a(x)(w)$

$T_a(x)$ is a linear map between the vector spaces $\mathrm{Cl}(\mathbb{R}^m, \gamma)$, $\mathrm{Cl}(E(x), g(x))$ which can be expressed in their bases.

The transitions are :

$x \in O_a \cap O_b : \Phi_a(x, w_a) = \Phi_b(x, w_b) \Leftrightarrow w_b = T_b(x)^{-1} \circ T_a(x)(w_a)$

so the transition maps are linear, and $\mathrm{Cl}(TM)$ is a vector bundle.

iii) The action of $O(\mathbb{R}, r, s)$ on the Clifford algebra $Cl(\mathbb{R}, r, s)$ is :

$\lambda : O(\mathbb{R}, r, s) \times Cl(\mathbb{R}, r, s) \to Cl(\mathbb{R}, r, s) :: \lambda(h, u) = \alpha(s) \cdot u \cdot s^{-1}$ where $s \in Pin(\mathbb{R}, r, s) : \mathbf{Ad}_w = h$

This action is extended on $\mathrm{Cl}(TM)$ fiberwise :

$\Lambda : O(\mathbb{R}, r, s) \times Cl(TM)(x) \to Cl(TM)(x) :: \Lambda(h, W_x) = \Phi_a(x, \alpha(s)) \cdot W_x \cdot \Phi_a(x, s^{-1})$

For each element of $O(\mathbb{R}, r, s)$ there are two elements $\pm w$ of $Pin(\mathbb{R}, r, s)$ but they give the same result.

With this action $\mathrm{Cl}(M, Cl(\mathbb{R}, r, s), \pi_c)$ is a $O(\mathbb{R}, r, s)$ bundle : $P \times_{O(\mathbb{R}, r, s)} Cl(\mathbb{R}, r, s)$. $\blacksquare$

Comments :

i) $\mathrm{Cl}(TM)$ can be seen as a vector bundle $\mathrm{Cl}(M, Cl(\mathbb{R}, r, s), \pi_c)$ with standard fiber $Cl(\mathbb{R}, p, q)$, or an associated vector bundle $\mathrm{P}[\mathrm{Cl}(\mathbb{R}, r, s), \lambda]$, or as a G-bundle : $P \times_{O(\mathbb{R}, p, q)} Cl(\mathbb{R}, r, s)$. But it has additional properties, as we have all the operations of Clifford algebras, notably the product of vectors, available fiberwise.

ii) This structure can always be built whenever we have a principal bundle modelled over an orthogonal group. This is always possible for a riemannian metric but there are topological obstruction for the existence of pseudo-riemannian manifolds.

iii) With the same notations as above.if we take the restriction $\widetilde{T}_a(x)$ of the isomorphisms $T_a(x)$ to the Pin group we have a group isomorphism : $\widetilde{T}_a(x) : Pin(\mathbb{R}, r, s) \to Pin(T_xM, g(x))$ and the maps $\widetilde{T}_b(x)^{-1} \circ \widetilde{T}_a(x)$ are group automorphims on $Pin(\mathbb{R}, r, s)$, so we have the structure of a fiber bundle $Pin(M, Pin(\mathbb{R}, r, s), \pi_p)$. However there is no guarantee that this is a principal fiber bundle, which requires $w_b = \left(T_b(x)^{-1} \circ T_a(x) \right)(w_a) = T_{ba}(x) \cdot w_a$ with $T_{ba}(x) \in Pin(\mathbb{R}, r, s)$. Indeed an automorphism on a group is not necessarily a translation.

Spin structure

For the reasons above, it is useful to define a principal spin bundle with respect to a principal bundle with an orthonormal group : this is a spin structure.

Definition 2007 *On a pseudo-riemannian manifold (M,g), with its principal*

bundle of orthogonal frames $O(M, O(\mathbb{R}, r, s), \pi)$, *an atlas* (O_a, φ_a) *of O, a* **spin structure** *(or spinor structure) is a family of maps* $(\chi_a)_{a \in A}$ *such that :* $\chi_a(x) : Pin(\mathbb{R}, r, s) \to O(\mathbb{R}, r, s)$ *is a group morphism.*

So there is a continuous map which selects, for each g$\in O(\mathbb{R}, r, s)$, one of the two elements of the Pin group $Pin(\mathbb{R}, r, s)$.

Theorem 2008 *A spin structure defines a principal pin bundle* $Pin(M, O(\mathbb{R}, r, s), \pi)$

Proof. The trivializations are :
$\psi_a : O_a \times Pin(\mathbb{R}, r, s) \to Pin(M, O(\mathbb{R}, r, s), \pi) :: \psi_a(x, s) = \varphi_a(x, \chi(s))$
At the transitions :
$\psi_b(x, s_b) = \psi_a(x, s_a) = \varphi_a(x, \chi(s_a)) = \varphi_b(x, \chi(s_b)) \Leftrightarrow \chi(s_b) = g_{ba}(x) \chi(s_a) = \chi(s_{ba}(x)) \chi(s_a)$
$s_b = s_{ba}(x) s_a$ ∎
There are topological obstructions to the existence of spin structures on a manifold (see Giachetta p.245).
If M is oriented we have a similar definition and result for a principal Spin bundle.
With a spin structure any associated bundle $P[V, \rho]$ can be extended to an associated bundle $Sp[V, \widetilde{\rho}]$ with the left action of Spin on V : $\widetilde{\rho}(s, u) = \rho(\chi(g), u)$
One can similarly build a principal manifold $SP_c(M, Spin_c(\mathbb{R}, p, q), \pi_s)$ with the group $Spin_c(\mathbb{R}, p, q)$ and define complex spin structure, and complex associated bundle.

Spin bundle

Definition 2009 *A* **spin bundle,** *on a manifold M endowed with a Clifford bundle structure Cl(TM) is a vector bundle* $E(M, V, \pi)$ *with a map R on M such that* $(E(x), R(x))$ *are geometric equivalent representations of* $Cl(TM)(x)$

A spin bundle is not a common vector bundle, or associated bundle. Such a structure does not always exist. For a given representation there are topological obstructions to their existence, depending on M. Spin manifolds are manifolds such that there is a spin bundle structure for any representation. The following theorem shows that any manifold with a structure of a principal bundle of Spin group is a spin bundle. So a manifold with a spin structure is a spin manifold.

Theorem 2010 *If there is a principal bundle* $Sp(M, Spin(\mathbb{R}, r, s), \pi_S)$ *on the r+s=m dimensional real manifold M, then for any representation (V,r) of the Clifford algebra* $Cl(\mathbb{R}, r, s)$ *there is a spin bundle on M.*

Proof. The ingredients are the following :
a principal bundle Sp(M,$Spin(\mathbb{R}, r, s)$,π_S) with atlas $(O_a, \varphi_a)_{a \in A}$.transition maps $\varphi_{ba}(x) \in Spin(\mathbb{R}, r, s)$ and right action ρ.

$(\mathbb{R}^m, g)$ endowed with the symmetric bilinear form of signature (r,s) on $\mathbb{R}^m$ and its basis $(\varepsilon_i)_{i=1}^m$

$(\mathbb{R}^m, \mathbf{Ad})$ the representation of $Spin(\mathbb{R}, r, s)$

(V,r) a representation of $Cl(\mathbb{R}, r, s)$, with a basis $(e_i)_{i=1}^n$ of V

From which we have :

an associated vector bundle $F = Sp[\mathbb{R}^m, \mathbf{Ad}]$ with atlas $(O_a, \phi_a)_{a \in A}$ and holonomic basis : $\varepsilon_{ai}(x) = \phi_a(x, \varepsilon_i)$, $\varepsilon_{bi}(x) = Ad_{\varphi_{ba}(x)} \varepsilon_{ai}(x)$. Because $\mathbf{Ad}$ preserves γ the vector bundle F can be endowed with a scalar product g.

an associated vector bundle E=Sp[V,r] with atlas $(O_a, \psi_a)_{a \in A}$ and holonomic basis : $e_{ai}(x) = \psi_a(x, e_i)$, $e_{bi}(x) = r(\varphi_{ba}(x)) e_{bi}(x)$

$\psi_a : O_a \times V \to E :: \psi_a(x, u) = (\varphi_a(x, 1), u)$

For $x \in O_a \cap O_b$:

$U_x = (\varphi_a(x, s_a), u_a) \sim (\varphi_b(x, s_b), u_b) \in E(x)$

$\exists g : \varphi_b(x, s_b) = \rho(\varphi_a(x, s_a), g) = \varphi_a(x, s_a g) \Leftrightarrow s_b = \varphi_{ba}(x) s_a g$

$u_b = r(g^{-1}) u_a = r(s_b^{-1} \varphi_{ba}(x) s_a) u_a$

Each fiber (F(x),g(x)) has a Clifford algebra structure Cl(F(x),g(x)) isomorphic to Cl(TM)(x). There is a family $(O_a, T_a)_{a \in A}$ of Clifford isomorphism : $T_a(x) : Cl(F(x), g(x)) \to Cl(\mathbb{R}, r, s)$ defined by identifying the bases : $T_a(x)(\varepsilon_{ai}(x)) = \varepsilon_i$ and on $x \in O_a \cap O_b : T_a(x)(\varepsilon_{ai}(x)) = T_b(x)(\varepsilon_{bi}(x)) = \varepsilon_i = Ad_{\varphi_{ba}(x)} T_b(x)(\varepsilon_{ai}(x))$

$\forall W_x \in Cl(F(x), g(x)) : T_b(x)(W_x) = Ad_{\varphi_{ba}(x)} T_a(x)(W_x)$

The action R(x) is defined by the family $(O_a, R_a)_{a \in A}$ of maps:

$R(x) : Cl(F(x), g(x)) \times E(x) \to E(x) :: R_a(x)(W_x)(\varphi_a(x, s_a), u_a) = (\varphi_a(x, s_a), r(s_a^{-1} \cdot T_a(x)(W_x) \cdot s_a) u_a)$

The definition is consistent and does not depend on the trivialization:

For $x \in O_a \cap O_b$:

$R_b(x)(W_x)(U_x) = (\varphi_b(x, s_b), r(s_b^{-1} \cdot T_b(x)(W_x) \cdot s_b) u_b)$

$\sim (\varphi_a(x, s_a), r(s_a^{-1} \varphi_{ab}(x) s_b) r(s_b^{-1} \cdot T_b(x)(W_x) \cdot s_b) u_b)$

$= (\varphi_a(x, s_a), r(s_a^{-1} \varphi_{ab}(x) s_b) r(s_b^{-1} \cdot T_b(x)(W_x) \cdot s_b) r(s_b^{-1} \varphi_{ba}(x) s_a) u_a)$

$= (\varphi_a(x, s_a), r(s_a^{-1} \cdot \varphi_{ab}(x) \cdot T_b(x)(W_x) \cdot \varphi_{ba}(x) \cdot s_a) u_a)$

$= (\varphi_a(x, s_a), r(s_a^{-1} \cdot Ad_{\varphi_{ab}(x)} T_b(x)(W_x) \cdot s_a) u_a) = R_a(x)(W_x)(U_x)$ ∎

25 JETS

Physicists are used to say that two functions f(x),g(x) are "closed at the rth order" in the neighborhood of a if $|f(x) - g(x)| < k |x - a|^r$ which translates as the rth derivatives are equal at a : $f^k(a) = g^k(a), k \leq r$. If we take differentiable maps between manifolds the equivalent involves partial derivatives. This is the starting point of the jets framework. It is useful in numerical analysis, however it has found a fruitful outcome in the study of fiber bundles and functional analysis. Indeed whenever some operator on a manifold involves differentials, that is most often, we have to deal with structures which contains some parts expressed as jets : a linear combination of differentials, related to the objects defined on the manifolds. Whereas in differential geometry one strives to get "intrinsic formulations", without any reference to the coordinates, they play a central role in jets.

This section comes here because jets involve fiber bundles, and one common definition of connections on fiber bundle involve jets. Jets are exensively used in Functional Analysis.

25.1 Jets on a manifold

In this section the manifolds will be assumed real finite dimensional and smooth.

25.1.1 Definition of a jet

We follow Kolar Chap.IV

Definition 2011 *Two paths $P_1, P_2 \in C_\infty(\mathbb{R}; M)$ on a real finite dimensional smooth manifold M, going through $p \in M$ at t=0, are said to have a **r-th order contact** at p if for any smooth real function f on M the function $(f \circ P_1 - f \circ P_2)$ has derivatives equal to zero for all order*

$0 \leq k \leq r : k=0,...r :(f \circ P_1 - f \circ P_2)(0)^{(r)} = 0$

Definition 2012 *Two maps $f, g \in C_r(M; N)$ between the real finite dimensional, smooth manifolds M,N are said to have a r-th order contact at $p \in M$ if for any smooth path P on M going through p then the paths $f \circ P, g \circ P$ have a r-th order contact at p.*

*A **r-jet** at p is a class of equivalence in the relation "to have a r-th order contact at p" for maps in $C_r(M; N)$.*

Two maps belonging to the same r-jet at p have same value f(p), and same derivatives at all order up to r at p.

$f, g \in C_r(M; N) : f \sim g \Leftrightarrow 0 \leq s \leq r : f^{(s)}(p) = g^{(s)}(p)$

The **source** is p and the **target** is f(p)=g(p)=q

Notation 2013 j_p^r *is a r-jet at the source* $p \in M$
$j_p^r f$ *is the class of equivalence of* $f \in C_r(M; N)$ *at p*

Any s order derivative of a map f is a s symmetric linear map $z_s \in \mathcal{L}_S^s(T_pM; T_qN)$, or equivalently a tensor $\odot^s T_pM^* \otimes T_qN$ where $\odot$ denote the symmetric tensorial product (see Algebra).

So a r jet at p in M denoted j_p^r with target q can be identified with a set :
$\{q, z_s \in \mathcal{L}_S^s(T_pM; T_qN), s = 1...r\} \equiv \{q, z_s \in \odot^s T_pM^* \otimes T_qN, s = 1...r\}$
and the set $J_p^r(M, N)_q$ of all r jets j_p^r with target q can be identified with the set $(\oplus_{s=1}^r \odot^s T_pM^*) \otimes T_qN$

We have in particular : $J_p^1(M, N)_q = T_qN \otimes T_pM^* = T_pM^* \otimes T_qN$. Indeed the 1 jet is just f'(p) and this is the set of linear map from T_pM to T_qN.

The tensorial notation is usually the most illuminating.

Notation 2014 $J_p^r(M, N)_q$ *is the set of r order jets at* $p \in M$ *(the source), with value* $q \in N$ *(the target)*
$J_p^r(M, N)_q = \{z_s \in \odot^s T_pM^* \otimes T_qN, s = 1...r\}$
$J_p^r(M, N) = \cup_q J_p^r(M, N)_q = \{z_s \in \odot^s T_pM^* \otimes TN, s = 1...r\}$
$J^r(M, N)_q = \cup_p J_p^r(M, N)_q = \{z_s \in \odot^s TM^* \otimes T_qN, s = 1...r\}$
$J^r(M, N) = \{z_s \in \odot^s TM^* \otimes TN, s = 1...r\}$
$J^0(M; N) = M \times N$ *Conventionally*

Definition 2015 *the **r jet prolongation** of a map* $f \in C_r(M; N)$ *is the map* $: J^r f : M \to J^r(M, N) :: (J^r f)(p) = j_p^r f$

Notation 2016 π_s^r *is the projection* $\pi_s^r : J^r(M, N) \to J^s(M, N) :: \pi_s^r(j^r) = j^s$ *where we drop the s+1...r terms*
π_0^r *is the projection* $\pi_0^r : J^r(M, N) \to M \times N :: \pi_0^r(j_p^r f) = (p, f(p))$
π^r *is the projection* $\pi^r : J^r(M, N) \to M : \pi_0^r(j_p^r f) = p$

25.1.2 Structure of the space of r-jets

In the following we follow Krupka (2000).

The key point is that in a r-jet we "forget" the map f to keep only the value of f(p) and the value of the derivarives $f^{(s)}(p)$

Coordinates expression

Any map $f \in C_r(M; N)$ between the manifolds M,N with atlas $(\mathbb{R}^m, (O_a, \varphi_a)_{a \in A})$, $(\mathbb{R}^n, (Q_b, \psi_b)_{b \in B})$ is represented in coordinates by the functions:
$F = \varphi_a \circ f \circ \psi_b^{-1} : \mathbb{R}^m \to \mathbb{R}^n :: \zeta^i = F^i(\xi^1, ..., \xi^m)$
The partial derivatives of f expressed in the holonomic bases of the atlas are the same as the partial derivatives of f, with respect to the coordinates. So :

Theorem 2017 *A r-jet $j_p^r \in J_p^r(M,N)_q$,in atlas of the manifolds, is represented by a set of scalars :*

$\left(\zeta_{\alpha_1...\alpha_s}^i, 1 \leq \alpha_k \leq m, i=1..n, s=1..r \right)$ *with m=dimM, n=dim N where the* $\zeta_{\alpha_1...\alpha_s}^i$ *are symmetric in the lower indices*

These scalars are the components of the symmetric tensors : $\{z_s \in \odot^s T_p M^* \otimes T_q N, s = 1...r\}$ in the holonomic bases.(dx^α) of $T_p M^*$, (∂y_i) of $T_q N$, :

$Z = \sum_{s=1}^r \sum_{(\alpha_1,...\alpha_s)} \sum_{i=1}^n \zeta_{\alpha_1...\alpha_s}^i dx^{\alpha_1} \otimes ... \otimes dx^{\alpha_s} \otimes \partial y_i \in \oplus_{s=1}^r \left(\odot^s T_p M^* \otimes T_q N \right)$

Structure with p,q fixed

The source and the target are the same for all maps in $J_p^r(M,N)_q$

The set of scalars:

$\left(\zeta_{\alpha_1...\alpha_s}^i, 1 \leq \alpha_k \leq m, i=1..n, s=1..r \right)$ with m=dimM, n=dim N where the $\zeta_{\alpha_1...\alpha_s}^i$ are symmetric in the lower indices

is a $J = n \left(C_{m+r}^m - 1 \right)$ dimensional vector subspace of the vector space $\mathbb{R}^{mn \frac{r(r+1)}{2}}$

It can be seen as a smooth real J dimensional manifold $L_{m,n}^r$ embedded in $\mathbb{R}^{mn \frac{r(r+1)}{2}}$

A point Z has for coordinates : $\left(\zeta_{\alpha_1...\alpha_s}^i, 1 \leq \alpha_1 \leq \alpha_2.. \leq \alpha_s \leq m, i = 1..n, s = 1..r \right)$ to account for the symmetries

and charts : $\Phi_{a,b}^s : L_{m,n}^s \to \mathbb{R}^J :: \sigma_s \in \mathfrak{S}(s) : z_{\alpha_{\sigma(1)}...\alpha_{\sigma(s)}}^i = \zeta_{\alpha_1...\alpha_s}^i$

with the basis $(dx^{\alpha_1} \otimes ... \otimes dx^{\alpha_s} \otimes \partial y_i, i = 1..n, \alpha_k = 1...m, s = 1..r)$

Remark : A r-jet is represented by a set Z of scalars, symmetric and with non ordered indices. If we account for the symmetries, we have fewer coordinates for Z, with ordered set of indices.

Structure of $J^r(M,N)$

If we do not specify p and q, we have $J^r(M,N) = \{p, q, z_s \in \odot^s TM^* \otimes TN, s = 1...r\}$. We have to account for p and q.

$J^r(M,N)$ has the structure of a smooth N+m+n=nC_{m+r}^m+m real manifold with charts $\left(O_a \times Q_b \times L_{m,n}^r, \left(\varphi_a, \psi_b, \Phi_{a,b}^s \right) \right)$ and coordinates

$\left(\xi^\alpha, \zeta^i, \zeta_{\alpha_1...\alpha_s}^i, 1 \leq \alpha_1 \leq \alpha_2.. \leq \alpha_s \leq m, i = 1..n, s = 1..r \right)$

The projections $\pi_s^r : J^r(M,N) \to J^s(M,N)$ are smooth submersions and translate in the chart by dropping the terms s>r

For each p the set $J_p^r(M,N)$ is a submanifold of $J^r(M,N)$

For each q the set $J^r(M,N)_q$ is a submanifold of $J^r(M,N)$

$J^r(M,N)$ is a smooth fibered manifold $J^r(M,N)(M \times N, \pi_0^r)$

$J^r(M,N)$ is an affine bundle over $J^{r-1}(M,N)$ with the projection π_{r-1}^r

(Krupka p.66).$J^r(M,N)$ is an associated fiber bundle $GT_m^r(M)[T_n^r(N), \lambda]$ - see below

Associated polynomial map

Theorem 2018 *To any r-jet $j_p^r \in J_p^r(M,N)_q$ is associated a symmetric polynomial of order r*

This is the polynomial :
i=1..n : $P^i\left(\xi^1,...,\xi^m\right) = \zeta_0^i + \sum_{s=1}^r \sum_{\alpha_1...\alpha_s=1}^m \zeta_{\alpha_1...\alpha_s}^i \left(\xi^{\alpha_1} - \xi_0^{\alpha_1}\right)...\left(\xi^{\alpha_s} - \xi_0^{\alpha_s}\right)$
with : $\left(\xi_0^1,...,\xi_0^m\right) = \varphi_a(p)$, $\left(\zeta_0^1,...,\zeta_0^n\right) = \psi_b(q)$
At the transitions between open subsets of M,N we have local maps :
$\varphi_{aa'}(p) \in \mathcal{L}\left(\mathbb{R}^m;\mathbb{R}^m\right), \psi_{bb'}(q) \in \mathcal{L}\left(\mathbb{R}^n;\mathbb{R}^n\right)$
$(\zeta - \zeta_0) = \psi_{bb'}(q)(\zeta' - \zeta_0')$
$(\xi - \xi_0) = \varphi_{aa'}(p)(\xi' - \xi_0')$
$(\zeta' - \zeta_0')$
$= \sum_{s=1}^r \sum_{\alpha_1...\alpha_s=1}^m P_{\alpha_1...\alpha_s}^i \sum_{\beta_1..\beta_s=1}^m [\varphi_{aa'}(p)]_{\beta_1}^{\alpha_1}\left(\xi'^{\beta_1} - \xi_0'^{\beta_1}\right)...[\varphi_{aa'}(p)]_{\beta_s}^{\alpha_s}\left(\xi'^{\alpha_s} - \xi_0'^{\alpha_s}\right)$

$P_{\alpha_1...\alpha_s}'^i = \left[\psi_{bb'}(q)^{-1}\right]_j^i P_{\alpha_1...\alpha_s}^j \sum_{\beta_1..\beta_s=1}^m [\varphi_{aa'}(p)]_{\beta_1}^{\alpha_1}...[\varphi_{aa'}(p)]_{\beta_s}^{\alpha_s}$

They transform according to the rules for tensors in $\odot^s T_p M^* \otimes T_q N$

It is good to keep in mind this representation of a r jet as a polynomial : they possess all the information related to a r-jet and so can be used to answer some subtle questions about r-jet, which are sometimes very abstract objects.

25.1.3 Jets groups

The composition of maps and the chain rule give the possibility to define the product of r-jets, and then invertible elements and a group structure.

Composition of jets :

Definition 2019 *The composition of two r jets is given by the rule :*
$\circ : J^r(M;N) \times J^r(N;P) \to J^r(M;P) :: j_x^r(g \circ f) = \left(j_{f(x)}^r g\right) \circ (j_x^r f)$

The definition makes sense because :
For $f_1, f_2 \in C_r(M;N), g_1, g_2 \in C_r(N;P)$, $j_x^r f_1 = j_x^r f_2, j_y^r g_1 = j_y^r g_2, y = f_1(x) = f_2(x)$
$\Rightarrow j_x^r(g_1 \circ f_1) = j_x^r(g_2 \circ f_2)$

Theorem 2020 *The composition of r jets is associative and smooth*

In coordinates the map $L_{m,n}^r \times L_{n,p}^r \to L_{m,p}^r$ can be obtained by the product of the polynomials P^i, Q^i and discarding all terms of degree $>r$.
$\left(\sum_{s=1}^r \sum_{\alpha_1,...,\alpha_s} a_{\alpha_1...\alpha_s}^i t_{\alpha_1}...t_{\alpha_s}\right) \times \left(\sum_{s=1}^r \sum_{\alpha_1,...,\alpha_s} b_{\alpha_1...\alpha_s}^i t_{\alpha_1}...t_{\alpha_s}\right)$
$= \left(\sum_{s=1}^{2r} \sum_{\alpha_1,...,\alpha_s} c_{\alpha_1...\alpha_s}^i t_{\alpha_1}...t_{\alpha_s}\right)$
$c_{\alpha_1...\alpha_s}^j = \sum_{k=1}^s \sum a_{\beta_1...\beta_k}^j b_{I_1}^{\beta_1}...b_{I_k}^{\beta_k}$ where $(I_1, I_2,..I_k) =$ any partition of $(\alpha_1,...\alpha_s)$

For r= 2 : $c_\alpha^i = a_\beta^i b_\beta^\beta; c_{\alpha\beta}^i = a_{\lambda\mu}^i b_\alpha^\lambda b_\beta^\mu + a_\gamma^i b_{\alpha\beta}^\gamma 1$ and the coefficients b for the inverse are given by : $\delta_\alpha^i = a_\beta^i b_\beta^\beta; a_{\alpha\beta}^i = -b_{\lambda\mu}^j a_\alpha^\lambda a_\beta^\mu a_j^i$
(see Krupka for more values)

Invertible r jets :

Definition 2021 *If dimM=dimN=n a r-jet* $X \in J_p^r(M;N)_q$ *is said **invertible** (for the composition law) if :*

$\exists Y \in J_q^r(N;M)_p : X \circ Y = Id_M, Y \circ X = Id_N$

and then it will be denoted X^{-1}.
X is inversible iff $\pi_1^r X$ is invertible.
The set of invertible elements of $J_p^r(M,N)_q$ is denoted $GJ_p^r(M,N)_q$

Definition 2022 *The **rth differential group** (or r jet group) is the set, denoted $GL^r(\mathbb{R}, n)$, of invertible elements of $L_{n,n}^r = J_0^r(\mathbb{R}^n, \mathbb{R}^n)_0$*

Theorem 2023 *(Kolar p.129) The set $GL^r(\mathbb{R}, n)$ is a Lie group, with Lie algebra $L^r(\mathbb{R}, n)$ given by the vector space $\{j_0^r X, X \in C_r(\mathbb{R}^n; \mathbb{R}^n), X(0) = 0\}$ and bracket :* $[j_0^r X, j_0^r Y] = -j_0^r([X,Y])$. *The exponential mapping is :* $\exp j_0^r X = j_0^r \Phi_X$

For r=1 we have $GL^1(\mathbb{R}, n) = GL(\mathbb{R}, n)$.
The canonical coordinates of $G \in GL^r(\mathbb{R}, n)$ are $\left\{g_{\alpha_1...\alpha_s}^i, i, \alpha_j = 1...n, s = 1...r\right\}$
The group $GL^r(\mathbb{R}, n)$ has many special properties (see Kolar IV.13) which can be extended to representations.

Functor J^r:
The functor : $J^r : \mathfrak{M}_m \times \mathfrak{M} \mapsto \mathfrak{FM}$ where $\mathfrak{M}_m$ is the category of m dimensional manifolds with local diffeormorphisms, associates to each couple :
$(M \times N) \mapsto J^r(M, N)$
$(f \in C_r(M_1, N_1), g \in C_r(M_2, N_2)) \mapsto \hom(J^r(M_1, N_1), J^r(M_2, N_2)) :: J^r(f,g)(X) = (j_q^r g) \circ X \circ (j_p^r f)^{-1}$ where $X \in J^r(M_1, N_1)$ is such that : $q = X(p)$

25.1.4 Velocities and covelocities

So far we have seen jets between manifolds. To emphasize the properties which are more specific to a given manifold, we define velocities and covelocities

Velocities

Definition 2024 *The space of **k velocities** to a manifold M is the set* $T_k^r(M) = J_0^r(\mathbb{R}^k, M)$

$T_k^r(M) = \left\{q, z_s \in \odot^s \mathbb{R}^{k*} \otimes TM, s = 1...r\right\}$
$= \left\{q, z_{\alpha_1...\alpha_s}^i e^{\alpha_1} \otimes .. \otimes e^{\alpha_k} \otimes \partial x_i, \alpha_j = 1...k, s = 1...r, i = 1..m\right\}$ where the $z_{\alpha_1...\alpha_s}^i$ are symmetric in the lower indices. Notice that it includes the target in M.
For k=1,r=1 we have just the tangent bundle $T_1^1(M) = TM = \left\{q, z^i \partial x_i\right\}$

Theorem 2025 *(Kolar p.120)* $T_k^r(M)$ *is* :

    a smooth mC_{k+r}^k *dimensional manifold,*

    a smooth fibered manifold $T_k^r(M) \to M$

    a smooth fibered bundle $T_k^r(M)(M, L_{km}^r, \pi^r)$

    $T_k^r(M)$ *can be canonically identified with the associated fiber bundle* $GT_m^r(M)[L_{mn}^r, \lambda]$ *with* $\lambda = $ *the left action of* $GL^r(\mathbb{R}, m)$ *on* L_{mn}^r

    $GL^r(\mathbb{R}, k)$ acts on the right on $T_k^r(M)$:

    $\rho : T_k^r(M) \times GL^r(\mathbb{R}, k) \to T_k^r(M) :: \rho(j_0^r f, j_0^r \varphi) = j_0^r(f \circ \varphi)$ with $\varphi \in Diff_r(\mathbb{R}^k; \mathbb{R}^k)$,

Theorem 2026 *If G is a Lie group then* $T_k^r(G)$ *is a Lie group with multiplication* : $(j_0^r f) \cdot (j_0^r g) = j_0^r(f \cdot g), f, g \in C_r(\mathbb{R}^k; G)$

Definition 2027 *The functor* : $T_k^r : \mathfrak{M} \mapsto \mathfrak{FM}$ *associates to each* :

    manifold $M : T_k^r(M)$

    map $f \in C_r(M; N) \mapsto T_r^k f : T_k^r(M) \to T_k^r(N) :: T_k^r f(j_0^r g) = j_0^r(f \circ g)$

This functor preserves the composition of maps. For k=r=1 we have the functor : $f \to f'$

Definition 2028 *(Kolar p.122) The **bundle of r-frames** over a m dimensional manifold M is the set of* $GT_m^r(M)$ *the jets of r differentiable frames on M, or equivalently the set of all inversible jets of* $T_m^r(M)$. *This is a principal fiber bundle over* $M : GT_m^r(M)(M, GL^r(\mathbb{R}, m), \pi^r)$

For r=1 we have the usual linear frame bundle.

For any local r diffeomorphism $f : M \to N$ (with dimM=dimN=m) the map $GT_m^r f : GT_m^r(M) \to GT_m^r(N) :: GT_m^r f(j_0^r \varphi) = j_0^r(f \circ \varphi)$ is a morphism of principal bundles.

Covelocities

Definition 2029 *The space of **k covelocities** from a manifold M is the set* : $T_k^{r*}(M) = J^r(M, \mathbb{R}^k)_0$

    $T_k^{r*}(M) = \{p, z_{\alpha_1 \ldots \alpha_s}^i dx^{\alpha_1} \otimes \ldots \otimes dx^{\alpha_s} \otimes e_i, s = 1 \ldots r, \alpha_j = 1 \ldots m, i = 1 \ldots k\}$ where the $z_{\alpha_1 \ldots \alpha_s}^i$ are symmetric in the lower indices.

    $T_k^{r*}(M)$ is a vector bundle : for $f, g \in C_r(\mathbb{R}^k; M) : j_0^r(\lambda f + \mu g) = \lambda j_0^r f + \mu j_0^r g$

    For k=r=1 we have the cotangent bundle : $T_1^{1*}(M) = TM^*$

    The projection : $\pi_{r-1}^r : T_k^{r*}(M) \to T_k^{r-1*}(M)$ is a linear morphism of vector bundles

    If k=1 : $T_1^{r*}(M)$ is an algebra with multiplication of maps : $j_0^r(f \times g) = (j_0^r f) \times (j_0^r g)$

    There is a canonical bijection between $J_p^r(M, N)_q$ and the set of algebras morphisms hom$(J_p^r(M, \mathbb{R})_0, J_q^r(N, \mathbb{R})_0)$

    T_k^{r*} is a functor $\mathfrak{M}_m \mapsto \mathfrak{VM}$

25.2 Jets on fiber bundles

Sections on fiber bundles are defined as maps $M \to E$ so the logical extension of r jets to fiber bundles is to start from sections seen as maps. We follow Krupka (2000) and Kolar.

25.2.1 Jets on fibered manifolds

Definitions

Definition 2030 *The **r-jet prolongation** of a fibered manifold E, denoted $J^r E$, is the space of r-jets of sections over E.*

A section on a fibered manifold E is a local map S from an open O of M to E such that $\pi \circ S = \pi$.

The set $J^r(M, E)$ of r jets of maps $M \to E$ contains all the sections, and $J^r E \subset J^r(M, E)$

Fibered manifold structure

Theorem 2031 *(Kolar p.124) The r jet prolongation $J^r E$ of a fibered manifold $E(M, \pi)$ is a closed submanifold of $J^r(M, E)$ and a fibered submanifold of $J^r(M, E)(M \times E, \pi_0^r)$*

The dimension of $J^r E$ is $nC_{m+r}^m + m$ with dim(E)=m+n
Notice that $J\left(J^{r-1}E\right) \neq J^r E$

Notation 2032 *The projections are defined as :*
$\pi^r : J^r E \to M : \pi^r\left(j_x^r S\right) = x$
$\pi_0^r : J^r E \to E : \pi_0\left(j_x^r S\right) = S(x)$
$\pi_s^r : J^r E \to J^s E : \pi_s^r\left(j_x^r S\right) = j_x^s$

Sections

Notation 2033 $\mathfrak{X}\left(J^r E\right)$ *is the set of sections of the fibered manifold $J^r E$*

Theorem 2034 *A class r section $S \in \mathfrak{X}_r(E)$ induces a section called its **r jet prolongation** denoted $J^r S \in \mathfrak{X}\left(J^r E\right)$ by : $J^r S(x) = j_x^r S$*

But conversely a section of $\mathfrak{X}\left(J^r E\right)$ does not need to come from a section of $S \in \mathfrak{X}_r(E)$.

The map : $J^r : \mathfrak{X}_r(E) \to \mathfrak{X}\left(J^r E\right)$ is neither injective or surjective. So the image of M by $J^r S$ is a subset of $J^r E$.

If E is a single manifold M then $J^r M \equiv M$. So the r jet prolongation $J^r S$ of a section S is a morphism of fibered manifolds : $M \to J^r M$

Coordinates

Let $(O_a, \psi_a)_{a \in A}$ be an adapted atlas of the finite dimensional fibered manifold E. Then : $\psi_a : O_a \to \mathbb{R}^m \times \mathbb{R}^n :: \psi_a(p) = (\xi^\alpha, y^i)$ where $(\xi^\alpha)_{\alpha=1}^m$ are the coordinates of x, and y^i the coordinates of a point u in some manifold V representing the fiber. A section S on E has for coordinates in this chart : $(\xi^\alpha, y^i(\xi))$ where $y^i : \mathbb{R}^m \to \mathbb{R}^n$. Its defines a set of numbers :

$\left(\xi^\alpha, y^i, \frac{\partial^s y^i}{\partial \xi^{\alpha_1}...\partial \xi^{\alpha_s}}, s = 1...r, 1 \leq \alpha_k \leq m, i = 1..n\right)$

A point Z of $J^r E$ as for coordinates :

$(\xi^\alpha, y^i, y^i_{\alpha_1..\alpha_s}, s = 1...r, 1 \leq \alpha_k \leq \alpha_{k+1} \leq m, i = 1..n)$

For a change a chart $(\xi^\alpha, y^i) \to \left(\widetilde{\xi}^\alpha, \widetilde{y}^i\right)$ in the manifold E we have on $J^r E$

:

$\widetilde{y}^i_{\beta \alpha_1...\alpha_s} = \sum_\gamma \frac{\partial \xi^\gamma}{\partial \widetilde{\xi}^\beta} d_\gamma y^i_{\alpha_1...\alpha_s}$ where $d_\gamma = \frac{\partial}{\partial \xi^\gamma} + \sum_{s=1}^r \sum_{\beta_1 \leq ... \leq \beta_s} y^i_{\gamma \beta_1...\beta_s} \frac{\partial}{\partial y^i_{\beta_1...\beta_s}}$

is the total differential (see below)

Prolongation of morphisms of fibered manifolds

Definition 2035 *The **r jet prolongation of a fibered manifold morphism** $(F,f) : E_1(M_1, \pi_1) \to E_2(M_2, \pi_2)$ with f is a smooth local diffeomorphism, is the morphism of fibered manifolds $(J^r F, J^r f) : J^r E_1 \to J^r E_2$*

(F,f) is such that : $F : E_1 \to E_2, f : M_1 \to M_2$ and : $\pi_2 \circ F = f \circ \pi_1$
$(J^r F, J^r f)$ is defined as :
$j^r f : J^r M_1 \to J^r M_2 :: j^r f(x_1) = j^r_{x_1} f$
$J^r F : J^r E_1 \to J^r E_2 :: \forall S \in \mathfrak{X}(E_1) : J^r F(j^r S(x)) = j^r_{f(x)}\left(F \circ S \circ f^{-1}\right)$
$\pi^r_s(J^r F) = J^s(\pi^r_s), \pi^r(J^r F) = f(\pi^r)$
If F is surjective or injective then so is its prolongation.

Total differential

Definition 2036 *(Kolar p.388) The **total differential of a base preserving morphism** $F : J^r E \to \Lambda_p T M^*$ from a fibered manifold $E(M, \pi)$ is the map: $\mathfrak{D}F : J^{r+1} E \to \Lambda_{p+1} T M^*$ defined as : $\mathfrak{D}F\left(j^{r+1}_x S\right) = d\left(F \circ j^r X\right)(x)$ for any local section S on E*

The total differential is also called formal differential.
F reads : $F = \sum_{\{\beta_1...\beta_p\}} \varpi_{\{\beta_1...\beta_k\}}\left(\xi^\alpha, \eta^i, \eta^i_\alpha, ...\eta^i_{\alpha_1...\alpha_r}\right) d\xi^{\beta_1} \wedge ... \wedge d\xi^{\beta_p}$
then : $\mathfrak{D}F = \sum_{\alpha=1}^m (d_\alpha F) d\xi^\alpha \wedge d\xi^{\beta_1} \wedge ... \wedge d\xi^{\beta_p}$
with : $d_\alpha F = \frac{\partial F}{\partial \xi^\alpha} + \sum_{s=1}^r \sum_{\beta_1 \leq .. \leq \beta_s} \frac{\partial F}{\partial u^i_{\beta_1...\beta_s}} \eta^i_{\alpha \beta_1...\beta_s}$

Theorem 2037 *The total differential of morphisms has the properties :*
$\mathfrak{D} \circ d = d \circ \mathfrak{D}$
$\mathfrak{D}(\varpi \wedge \mu) = \mathfrak{D}\varpi \wedge \mu + \varpi \wedge \mathfrak{D}\mu,$
$\mathfrak{D}(dx^\alpha) = 0$
$d_\gamma\left(du^i_{\alpha_1...\alpha_s}\right) = du^i_{\gamma \alpha_1...\alpha_s}$

Definition 2038 *The **total differential of a function** $f : J^r E \to \mathbb{R}$ is a map* : $\mathfrak{D} : J^r E \to \Lambda_1 TM^* :: \mathfrak{D} f = \sum_\alpha (d_\alpha f) \, d\xi^\alpha \in \Lambda_1 TM$ *with* :

$d_\alpha f = \frac{\partial f}{\partial \xi^\alpha} + \sum_{s=1}^r \sum_{\beta_1 \leq \ldots \leq \beta_s} \frac{\partial f}{\partial u^i_{\beta_1 \ldots \beta_s}} \eta^i_{\alpha \beta_1 \ldots \beta_s}$

Its definition comes from the fact that if we take the derivative $\frac{df}{d\xi^\alpha}$ when f is evaluated along a section X(x) then each component : $\frac{d\eta^i_{\beta_1 \ldots \beta_s}}{d\xi^\alpha} = \eta^i_{\alpha \beta_1 \ldots \beta_s}$

Prolongation of a projectable vector field

A projectable vector field W on fibered manifold $E(M, \pi)$ is such that : $\exists Y \in \mathfrak{X}(TM) : \forall p \in E : \pi'(p) W(p) = Y(\pi(p))$

Its components read : $W = W^\alpha \partial x_\alpha + W^i \partial u_i$ where W^α does not depend on u. It defines, at least locally, with any section X on E a one parameter group of base preserving morphisms on E through its flow : $U(t) : \mathfrak{X}(E) \to \mathfrak{X}(E) :: U(t) X(x) = \Phi_W(X(\Phi_Y(x, -t)), t)$

Definition 2039 *The **r-jet prolongation of the one parameter group** associated to a projectable vector field W on a fibered manifold E is the one parameter group of morphisms* :

$J^r \Phi_W : J^r E \to J^r E :: J^r \Phi_W(j^r X(x), t) = j^r_{\Phi_Y(x,t)}(\Phi_W(X(\Phi_Y(x, -t)), t))$
with any section X on E

*The **r-jet prolongation of the vector field** W is the vector field $J^r W$ on $J^r E$* :

$J^r W(j^r X(x)) = \frac{\partial}{\partial t} J^r \Phi_W(j^r X(x), t)|_{t=0}$

Its expression is the following (Krupka 2001 p.18) in an adapted chart:
$J^r W = W^\alpha \partial x_\alpha + W^i \partial u_i + \sum_{s=1}^r \sum_{\alpha_1 \leq .. \leq \alpha_s} W^i_{\alpha_1 \ldots \alpha_s} \partial u^{\alpha_1 \ldots \alpha_s}_i$ with $W^i_{\beta \alpha_1 \ldots \alpha_s} = d_\beta W^i_{\alpha_1 \ldots \alpha_s} - \sum_\gamma \eta^i_{\gamma \alpha_1 \ldots \alpha_s} \partial u^\gamma_i$ and the total differential d_β

For r=1 :
$J^1 W = \sum_\alpha W^\alpha \partial x_\alpha + \sum_i W^i \partial u_i + \sum_{i\alpha} \left(\frac{\partial W^i}{\partial \xi^\alpha} + \frac{\partial W^i}{\partial u^j} y^j_\alpha - \frac{\partial W^\beta}{\partial \xi^\alpha} y^i_\beta \right) \partial y^\alpha_i$

The r jet prolongation of a vertical vector field ($W^\alpha = 0$) is a vertical vector field :
$J^r W = W^i \partial u_i + \sum_{s=1}^r \sum_{\alpha_1 \leq \ldots \leq \alpha_s} \left(d_{\alpha_1} d_{\alpha_2} \ldots d_{\alpha_s} u^i \right) \partial u^{\alpha_1 \ldots \alpha_s}_i$
So by construction :

Theorem 2040 *The r jet prolongation of the one parameter group of morphisms induced by the projectable vector field W on E is the one parameter group of morphisms induced by the r jet prolongation $J^r W$ on $J^r E$.*

And it implies that :

Theorem 2041 *The Lie derivative $\pounds_{J^r W} Z$ of a section of $J^r E$ is a section of the vertical bundle $V J^r E$.*

As $J^r U(t)$ is fiber preserving $\pi^{r\prime}(\pounds_{J^r W} Z) = \pi^r(Z)$.
Moreover for s<r :

$$\pi_s^r \left(J^r U \left(t \right) Z \left(x \right) \right) = \pi_s^r \left(j_{\Phi_Y \left(x, t \right)}^r \left(\Phi_W \left(X \left(\Phi_Y \left(x, -t \right) \right), t \right) \right) \right)$$
$$= j_{\Phi_Y \left(x, t \right)}^s \left(\Phi_W \left(X \left(\Phi_Y \left(x, -t \right) \right), t \right) \right) = J^s U \left(t \right) \pi_s^r \left(Z \left(x \right) \right)$$
So : $\frac{d}{dt} \pi_s^r \left(J^r U \left(t \right) Z \left(x \right) \right) |_{t=0} = \frac{d}{dt} \left(J^s U \left(t \right) \pi_s^r \left(Z \left(x \right) \right) \right) |_{t=0}$
$$\pi_s^{r\prime} \left(Z \right) \pounds_{J^r W} Z = \pounds_{J^s W} \pi_s^r \left(Z \right)$$

25.2.2   Jets on fiber bundles

The definition of the jet prolongation of a fiber bundle is the same as for fibered manifolds (fiber bundles are fibered manifolds). But we can be more explicit.

Fiber bundle structure

Theorem 2042 *(Krupka 2000 p.75) The r jet prolongation $J^r E$ of the fiber bundle $E(M,V,\pi)$ is a fiber bundle $J^r E \left(M, J_0^r \left(\mathbb{R}^{\dim M}, V \right), \pi^r \right)$ and a vector bundle $J^r E \left(E, J_0^r \left(\mathbb{R}^{\dim M}, V \right)_0, \pi_0^r \right)$*

Let E(M,V,π) be a smooth fiber bundle. with trivializations $\left(O_a, \varphi_a \right)_{a \in A}$

A class r section S on E is defined by a family $\left(\sigma_a \right)_{a \in A}$ of class r maps $\sigma_a : O_a \to V$ such that : $S \left(x \right) = \varphi_a \left(x, \sigma_a \left(x \right) \right)$ and which meets the transition conditions : $\forall a, b \in A, O_a \cap O_b \neq \varnothing, \forall x \in O_a \cap O_b : \sigma_b \left(x \right) = \varphi_{ba} \left(x, \sigma_a \left(x \right) \right)$

Two sections are in the same r-jet at x if the derivatives $\sigma_a^{(s)} \left(x \right), 0 \leq s \leq r$ have same value.

The definition of $J^r E$ is purely local, so the transition conditions are not involved.

Let $Z \in J^r E$. This is the class of equivalence of sections $S \in \mathfrak{X}_r \left(E \right)$ which have the same projection $\pi \left(S \right) = x$ and same target p=S(x). Given any trivialization $\left(O_a, \varphi_a \right)_{a \in A}$ of E we have the same value for $\sigma_a \left(x \right)$

So the projection are :
$\pi^r : J^r E \to M : \pi^r \left(Z \right) = x$
$\pi_0^r : J^r E \to E : \pi^r \left(Z \right) = p$

And for r > 1 $Z \in J^r E$ can be identified with $\{x, p\} \times J_x^r \left(M, V \right)_{\sigma(x)}$

In order to give a definition of this space which is independant of x it suffices to see that
$J_x^r \left(M, V \right) \simeq J_0^r \left(\mathbb{R}^{\dim M}; V \right) \simeq TV \otimes \left(\oplus_{s=0}^r \odot^s \left(\mathbb{R}^{\dim M} \right)^* \right)$ with $\odot$ the symmetric tensorial product.

If we keep M as base of the fibered manifold $J^r E$ we have to account for p$\in E$ and thus include the target, the fiber $J_x^r E$ is modelled on $J_0^r \left(\mathbb{R}^{\dim M}; V \right)$.

If we take E as base of the fibered manifold $J^r E$ the fiber $J_p^r E$ is modelled on $J_0^r \left(\mathbb{R}^{\dim M}; V \right)_0$ which is a vector space, isomorphic to $TV \otimes \left(\oplus_{s=0}^r \odot^s \left(\mathbb{R}^{\dim M} \right)^* \right)$ with $\odot$ the symmetric tensorial product.

A point of $J_0^r \left(\mathbb{R}^{\dim M}, V \right)_0$ is a set $z = \left(u_{\alpha_1 \ldots \alpha_s}, s = 1 \ldots r, 1 \leq \alpha_k \leq m \right)$ where
$$u_{\alpha_1 \ldots \alpha_s} = z_{\alpha_1 \ldots \alpha_s}^i e^{\alpha_1} \otimes .. \otimes e^{\alpha_k} \otimes \partial x_i \in TV \otimes \left(\oplus_{s=1}^r \odot^s \left(\mathbb{R}^{\dim M} \right)^* \right)$$
The trivialization of $J^r E \left(M, J_0^r \left(\mathbb{R}^{\dim M}, V \right), \pi^r \right)$ is :

$\Phi_a\left(x,\left(u,z\right)\right)=\left(\varphi_a\left(x,u\right),z\right)$

The trivialization of $J^r E\left(E,J_0^r\left(\mathbb{R}^{\dim M},V\right),\pi^r\right)$ is :

$\widetilde{\Phi}_a\left(p,z\right)=\left(z\right)$

Coordinates

E is a fibered manifold, and it has, as a manifold, an adapted atlas $\left(O_a,\psi_a\right)_{a\in A}$ such that :

$\psi_a:O_a\to\mathbb{R}^m\times\mathbb{R}^n::\psi_a\left(p\right)=\left(\xi^\alpha,y^i\right)$ where $\left(\xi^\alpha\right)_{\alpha=1}^m$ are the coordinates of x, and y^i the coordinates of u in V

In an atlas of V, we have : $\phi_a:O_a'\to\mathbb{R}^n::\phi_a\left(u\right)=\eta_i$

So : $\left(y^i\right)_{i=1}^n=\left(\eta^i\right)_{i=1}^n=\phi_a\left(u\right)$

The coordinates in $J^r E$ are : $\left(\xi^\alpha,\eta^i,\eta_{\alpha_1..\alpha_s}^i,s=1...r,1\le\alpha_k\le\alpha_{k+1}\le m,i=1..n\right)$

Theorem 2043 *The fiber bundle defined by the projection :* $\pi_{r-1}^r:J^r E\to J^{r-1}E$ *is an affine bundle modelled on the vector bundle* $J^{r-1}E\left(E,J_0^r\left(\mathbb{R}^{\dim M},V\right)_0,\pi_0^r\right)$

$J^1 E$ is an affine bundle over E, modelled on the vector bundle $TM^*\otimes VE\to E$

Sections of $J^r E$

Definition 2044 *A section* $J^r Z$ *of* $J^r E\left(M,J_0^r\left(\mathbb{R}^{\dim M},V\right),\pi^r\right)$ *is a map* $M\to J^r E$

with the trivializations : $\Phi_a\left(x,\left(u\left(x\right),z\left(x\right)\right)\right)=\left(\varphi_a\left(x,u\left(x\right)\right),z\left(x\right)\right)$

A r class section $S\in\mathfrak{X}_r\left(E\right)$ gives rise to a section $J^r S\in\mathfrak{X}\left(J^r E\right):J^r S\left(x\right)=j_x^r S$ with components : $z_{a\alpha_1...\alpha_s}^i=D_{\alpha_1...\alpha_s}\sigma_a^i$

Notice that sections on $J^r E$ are defined independantly of sections on E. In the coordinate expression of $J^r z$ we *do not have* necessarily the relations between the components :

$u_{a\,\alpha_1...\alpha_s\beta}^i\left(x\right)=\frac{\partial}{\partial\xi^\beta}u_{a\alpha_1...\alpha_s}^i\left(x\right)$

Vectors on the tangent space $T_Z J^r E$

A base of $T_Z J^r E$ is a set of vectors :

$\left(\partial x_\alpha,\partial u_i,\partial u_i^{\alpha_1...\alpha_{\alpha_s}},s=1..r,i=1...n,1\le\alpha_1\le\alpha_2..\le\alpha_m\le m\right)$

We need to take an ordered set of indices in order to account for the symmetry of the coordinates.

$\left(\partial x_\alpha,\partial u_i\right)$ is a basis of TE.

A vector on the tangent space $T_Z J^r E$ of $J^r E$ reads :

$W_Z=w_p+\sum_{s=1}^r\sum_{1\le\alpha_1\le..\le\alpha_s\le m}w_{\alpha_1...\alpha_s}^i\partial u_i^{\alpha_1...\alpha_{\alpha_s}}$

with $w_p=w^\alpha\partial x_\alpha+w^i\partial u_i\in T_p E$

The projections give :

$\pi^{r\prime}\left(Z\right):T_Z J^r E\to T_{\pi^r(Z)}M:\pi^{r\prime}\left(Z\right)W_Z=w^\alpha\partial\xi_\alpha$ Notice the difference : $\partial x_\alpha=\varphi_x'\left(x,u\right)\partial\xi_\alpha$

$\pi_0^{r\prime}\left(Z\right):T_Z J^r E\to T_{\pi_0^r(Z)}E:\pi^{r\prime}\left(Z\right)W_Z=w_p$

We have the dual basis :
$$\left(dx^\alpha, du^i, du^i_{\alpha_1...\alpha_{\alpha_s}}, s = 1..r, i = 1...n, 1 \le \alpha_1 \le \alpha_2.. \le \alpha_m \le m\right)$$

25.2.3   r jet prolongations of fiber bundles

Prolongation of a vector bundle

Theorem 2045 *The r jet prolongation $J^r E$ of a vector bundle $E(M, V, \pi)$ is the vector bundle $J^r E\left(M, J_0^r\left(\mathbb{R}^{\dim M}, V\right), \pi^r\right)$*

Take a holonomic basis $(e_i(x))_{i=1}^n$ of E and a holonomic basis $(dx^\alpha)_{\alpha=1}^m$ of TM^* then a vector of $J^r E$ reads :
$Z(x) = Z_0^i(x) e_i(x) + \sum_{s=1}^r \sum_{\alpha_1..\alpha_s=1}^m \sum_{i=1}^n Z_{\alpha_1...\alpha_s}^i e_i^{\alpha_1...\alpha_m}(x)$ where $Z_{\alpha_1...\alpha_s}^i$ are symmetric in the subscripts

A section $Z \in \mathfrak{X}(J^r E)$ can also conveniently be defined as a map : $M \to V \times \prod_{s=1}^r \left\{\mathcal{L}_S^s\left(\mathbb{R}^{\dim M}, V\right)\right\}$

Prolongation of a principal bundle

Theorem 2046 *(Kolar p.150) If $P(M, G, \pi)$ is a principal fiber bundle with M m dimensional, then its r jet prolongation is the principal bundle $W^r P(M, W_m^r G, \pi^r)$ with $W^r P = GT_m^r(M) \times_M J^r P, W_m^r G = GL^r(\mathbb{R}, m) \rtimes T_m^r(G)$*

The product in $W_m^r G$ is :
$(A_1, \theta_1), (A_2, \theta_2) \in GL^r(\mathbb{R}, m) \times T_m^r(G) : (A_1, \theta_1) \times (A_2, \theta_2) = (A_1 \times A_2, (\theta_1 \times A_2) \cdot \theta_2)$
A section of P reads :$S(x) = \varphi(x, \sigma(x))$ so the coordinates in $J^r P$ are :
$\left(\xi^\alpha, \sigma^i, \sigma^i_{\beta_1...\beta_s}, \beta_1 \le .. \le \beta_s, s = 1...r\right)$

Prolongation of an associated bundle

Theorem 2047 *(Kolar p.152) If $P(M, G, \pi)$ is a principal fiber bundle with G n dimensional, $P[V, \lambda]$ an associated bundle, then the r-jet prolongation of $P[V, \lambda]$ is $W^r P[T_n^r V, \Lambda]$ with the action defined by :*
$j_x^r(\psi_a(x, u)) \sim j_x^r \psi_a\left(\rho(p_a(x), g), \lambda(g^{-1}, u)\right)$

25.2.4   Infinite order jet

(Kolar p.125)
We have the inverse sequence :
$E \overset{\pi_0^1}{\leftarrow} J^1 E \overset{\pi_1^2}{\leftarrow} J^2 E \overset{\pi_2^3}{\leftarrow} ....$

If the base M and the fiber V are smooth the infinite prolongation $J^\infty E$ is defined as the minimal set such that the projections :

$\pi_r^\infty : J^\infty E \to J^r E, \pi^\infty : J^\infty E \to M, \pi_0^\infty : J^\infty E \to E$ are submersions and follow the relations : $\pi_r^\infty = \pi_r^s \circ \pi_s^\infty$

This set exists and, provided with the inductive limit topology, all the projections are continuous and $J^\infty E$ is a paracompact Fréchet space. It is not a manifold according to our definition : it is modelled on a Fréchet space.

26 CONNECTIONS

With fiber bundles we can extend the scope of mathematical objects defined over a manifold. When dealing with them the frames are changing with the location making the comparisons and their derivation more complicated. Connections are the tool for this job : they "connect" the frames at different point. So this is an of extension of the parallel transport. To do so we need to distinguish between transport in the base manifold, which becomes "horizontal transport", and transport along the fiber manifold, which becomes "vertical transport". Indeed vertical transports are basically changes in the frame without changing the location. So we can split a variation in an object between what can be attributed to a change of location, and what comes from a simple change of the frame.

As for fiber bundles, we have general connections on general fiber bundles. They are presented first, with their general properties, and the many objects: covariant derivative, curvature, exterior covariant derivative,... which are linked to connections, Then, for each of the 3 classes of fiber bundles : vector bundles, principal bundles, associated bundles, we have connections which takes advantage of the special feature of the respective bundle. For the most part it is an adaptation of the general framework.

26.1 General connections

Connections on a fiber bundle can be defined in a purely geometrical way, without any reference to coordinates, or through jets. Both involve Christoffel symbols. The first is valid in a general context (whatever the dimension and the field of the manifolds), the second restricted to the common case of finite dimensional real fiber bundles. They give the same results and the choice is mainly a matter of personnal preference. We will follow the geometrical way,as it is less abstract and more intuitive.

26.1.1 Definitions

Geometrical definition

The tangent space $T_p E$ at any point p to a fiber bundle $E(M, V, \pi)$ has a preferred subspace : the vertical space corresponding to the kernel of $\pi'(p)$, which does not depend on the trivialization. And any vector v_p can be decomposed between a part $\varphi'_x(x, u) v_x$ related to $T_x M$ and another part $\varphi'_u(x, u) v_u$ related to $T_u V$. However this decomposition is not unique, so it makes sense to look for a way to split v_p . That is to define a projection of v_p onto the vertical space.

Definition 2048 *A **connection** on a fiber bundle $E(M, V, \pi)$ is a 1-form Φ on E valued in the vertical bundle, which is a projection $TE \rightarrow VE$: $\Phi \in \wedge_1(E; VE) :: \Phi \circ \Phi = \Phi, \Phi(TE)) = VE$*

So Φ acts on vectors of the tangent bundle TE, and the result is in the vertical bundle.

Φ has constant rank, and $\ker \Phi$ is a vector subbundle of TE.

Definition 2049 *The* **horizontal bundle** *of the tangent bundle TE is the vector subbundle $HE=\ker \Phi$*

The tangent bundle TE is the direct sum of two vector bundles : $TE = HE \oplus VE$:

$V_p E = \ker \pi'(p)$

$H_p E = \ker \Phi(p)$

$\forall p \in E : T_p E = H_p E \oplus V_p E$

The horizontal bundle can be seen as "displacements along the base M(x)" and the vertical bundle as "displacements along V (u)". The key point here is that, even if the vertical space VE is well identified without the need of a connection (it depends only on π), to define a projection on VE we need some direction (think to an "orthogonal projection") in order to have a unique result.

Christoffel form

Theorem 2050 *A connection Φ on a fiber bundle $E(M,V,\pi)$ with atlas $(O_a, \varphi_a)_{a \in A}$ is uniquely defined by a family of maps $(O_a, \Gamma_a)_{a \in A}$, $\Gamma_a \in C\left(\pi^{-1}(O_a); TM^* \otimes TV\right)$ called the* **Christoffel forms** *of the connection.*

$\Gamma_a \in C\left(\pi^{-1}(O_a); TM^* \otimes TV\right) :: \Phi(p)v_p = \varphi'_a(x,u)(0, v_u + \Gamma_a(p)v_x)$

Proof. $\varphi_a : O_a \times V \to \pi^{-1}(O_a) \Rightarrow \varphi'_a : T_x O_a \times T_u V \to T_p E :: v_p = \varphi'_a(x,u)(v_x, v_u)$

and $\varphi'_a(x, u_a)$ is invertible.

$\Phi(p)v_p \in V_p E \Rightarrow \exists w_u \in T_u V : \Phi(p)v_p = \varphi'_a(x,u)(0, w_u)$

$\Phi(p)v_p = \Phi(p)\varphi'_a(x,u)(0, v_u) + \Phi(p)\varphi'(x,u)(v_x, 0) = \varphi'_a(x,u)(0, w_u)$

$\varphi'_a(x,u)(0, v_u) \in V_p E \Rightarrow \Phi(p)\varphi'_a(x,u)(0, v_u) = \varphi'_a(x,u)(0, v_u)$

So $\Phi(p)\varphi'(x,u)(v_x, 0) = \varphi'_a(x,u)(0, w_u - v_u)$ depends linearly on $w_u - v_u$

Let us define : $\Gamma_a : \pi^{-1}(O_a) \to \mathcal{L}(T_x M; T_u V) :: \Gamma_a(p)v_x = w_u - v_u$

So : $\Phi(p)v_p = \varphi'_a(x,u)(0, v_u + \Gamma_a(p)v_x)$ ∎

Γ is a map, defined on E (it depends on p) and valued in the tensorial product $TM^* \otimes TV : \Gamma(p) = \sum \Gamma(p)^i_\alpha \, d\xi^\alpha \otimes \partial u_i$ with a dual holonomic basis $d\xi^\alpha$ of TM* [4]. This is a 1-form acting on vectors of $T_{\pi(p)}M$ and valued dans $T_u V$

Remark : there are different conventions regarding the sign in the above expression. I have taken a sign which is consistent with the common definition of affine connections on the tangent bundle of manifolds, as they are the same objects.

---

[4] I will keep the notations $\partial x_\alpha, dx^\alpha$ for the part of the basis on TE related to M, and denote $\partial \xi_\alpha, d\xi^\alpha$ for holonomic bases in TM,TM*

Theorem 2051 *On a fiber bundle $E\left(M,V,\pi\right)$ a family of maps $\Gamma_a \in C\left(\pi^{-1}\left(O_a\right); TM^* \otimes TV\right)$ defines a connection on E iff it satisfies the transition conditions : $\Gamma_b\left(p\right) = \varphi'_{ba}\left(x,u_a\right) \circ \left(-Id_{TM}, \Gamma_a\left(p\right)\right)$ in an atlas $\left(O_a, \varphi_a\right)_{a \in A}$ of E.*

Proof. At the transitions between charts (see Tangent space in Fiber bundles) :

$x \in O_a \cap O_b : v_p = \varphi'_a\left(x,u_a\right)v_x + \varphi'_{au}\left(x,u_a\right)v_{au} = \varphi'_b\left(x,u_b\right)v_x + \varphi'_{bu}\left(x,u_b\right)v_{bu}$

we have the identities :

$\varphi'_{ax}\left(x,u_a\right) = \varphi'_{bx}\left(x,u_b\right) + \varphi'_{bu}\left(x,u_b\right)\varphi'_{bax}\left(x,u_a\right)$

$\varphi'_{au}\left(x,u_a\right) = \varphi'_{bu}\left(x,u_b\right)\varphi'_{bau}\left(x,u_a\right)$

$v_{bu} = \varphi'_{ba}\left(x,u_a\right)\left(v_x,v_{au}\right)$

i) If there is a connection : $\Phi(p)v_p = \varphi'_a(x,u_a)\left(0, v_{au} + \Gamma_a\left(p\right)v_x\right) = \varphi'_b(x,u_b)\left(0, v_{bu} + \Gamma_b\left(p\right)v_x\right)$

$v_{bu} + \Gamma_b\left(p\right)v_x = \varphi'_{ba}\left(x,u_a\right)\left(0, v_{au} + \Gamma_a\left(p\right)v_x\right) = \varphi'_{bau}\left(x,u_a\right)v_{au} + \varphi'_{bau}\left(x,u_a\right)\Gamma_a\left(p\right)v_x$

$= \varphi'_{bax}\left(x,u_a\right)v_x + \varphi'_{bau}\left(x,u_a\right)v_{au} + \Gamma_b\left(p\right)v_x$

$\Gamma_b\left(p\right)v_x = \varphi'_{bau}\left(x,u_a\right)\Gamma_a\left(p\right)v_x - \varphi'_{bax}\left(x,u_a\right)v_x = \varphi'_{ba}\left(x,u_a\right)\left(-v_x, \Gamma_a\left(p\right)v_x\right)$

ii) Conversely let be a set of maps $\left(\Gamma_a\right)_{a \in A}, \Gamma_a \in C\left(O_a; TM^* \otimes TV\right)$ such that $\Gamma_b\left(p\right) = \varphi'_{ba}\left(x,u_a\right) \circ \left(-Id_{TM}, \Gamma_a\left(p\right)\right)$

define $\Phi_a(p)v_p = \varphi'_a(x,u_a)\left(0, v_{au} + \Gamma_a\left(p\right)v_x\right)$

let us show that $\Phi_b(p)v_p = \varphi'_b(x,u_b)\left(0, v_{bu} + \Gamma_b\left(p\right)v_x\right) = \Phi_a(p)v_p$

$v_{bu} = \varphi'_{ba}\left(x,u_a\right)\left(v_x,v_{au}\right); \Gamma_b\left(p\right)v_x = -\varphi'_{bax}\left(x,u_a\right)v_x + \varphi'_{bax}\left(x,u_a\right)\Gamma_a\left(p\right)v_x$

$\varphi'_b(x,u_b) = \varphi'_{bu}\left(x, \varphi_{bau}\left(x,u_a\right)\right)$

$\Phi_b(p)v_p = \varphi'_{bu}(x,u_b)\varphi'_{bax}\left(x,u_a\right)v_x + \varphi'_{bu}(x,u_b)\varphi'_{bau}\left(x,u_a\right)v_{au}$

$-\varphi'_{bu}(x,u_b)\varphi'_{bax}\left(x,u_a\right)v_x + \varphi'_{bu}(x,u_b)\varphi'_{bax}\left(x,u_a\right)\Gamma_a\left(p\right)v_x$

$= \varphi'_{bu}(x,u_b)\varphi'_{bau}\left(x,u_a\right)\left(v_{au} + \Gamma_a\left(p\right)v_x\right) = \varphi'_{au}(x,u_a)\left(v_{au} + \Gamma_a\left(p\right)v_x\right)$ ∎

Jet definition

The 1-jet prolongation of E is an affine bundle J^1E over E, modelled on the vector bundle $TM^* \otimes VE \to E$. So a section of this bundle reads : $j^i_\alpha\left(p\right)d\xi^\alpha \otimes \partial u^i$

On the other hand $\Gamma\left(p\right) \in \mathcal{L}\left(T_xM; T_uV\right)$ has the coordinates in charts : $\Gamma\left(p\right)^i_\alpha d\xi^\alpha \otimes \partial u_i$ and we can define a connection through a section of the 1-jet prolongation J^1E of E.

The geometric definition is focused on Φ and the jet definition is focused on $\Gamma\left(p\right)$.

Pull back of a connection

Theorem 2052 *For any connection Φ on a fiber bundle $E\left(M,V,\pi\right)$, N smooth manifold and $f : N \to M$ smooth map, the pull back $f^*\Phi$ of Φ is a connection on $f^*E : \left(f^*\Phi\right)\left(y,p\right)\left(v_y, v_p\right) = \left(0, \Phi\left(p\right)v_p\right)$*

Proof. If $f : N \to M$ is a smooth map, then the pull back of the fiber bundle is a fiber bundle $f^*E\left(N,V,f^*\pi\right)$ such that :

Base : N, Standard fiber : V

Total space : $f^*E = \{\left(y,p\right) \in N \times E : f\left(y\right) = \pi\left(p\right)\}$

Projection : $\widetilde{\pi} : f^*E \to N :: \widetilde{\pi}\,(y,p) = y$

So q=$(y,p) \in f^*E \to v_q = (v_y, v_p) \in T_q f^*E$

$\widetilde{\pi}'\,(q)\,v_q = \widetilde{\pi}'\,(q)\,(v_y, v_p) = (v_y, 0)$

Take : $(f^*\Phi)\,(q)\,v_q = (0, \Phi\,(p)\,v_p)$

$\widetilde{\pi}'\,(q)\,(0, \Phi\,(p)\,v_p) = 0 \Leftrightarrow (0, \Phi\,(p)\,v_p) \in V_q f^*E$ ∎

26.1.2   Covariant derivative

The common covariant derivative is a map which transforms vector fields, $\otimes^1 TM$ tensors, into $\otimes^1_1 TM$. So it acts on sections of the tangent bundle. Similarly the covariant derivative associated to a connection acts on sections of E.

Definition 2053 *The covariant derivative ∇ associated to a connection Φ on a fiber bundle $E\,(M, V, \pi)$ is the map : $\nabla : \mathfrak{X}\,(E) \to \Lambda_1\,(M; VE) :: \nabla S = S^*\Phi$*

So the covariant derivative along a vector field X on M is : $\nabla_X S(x) = \Phi(S(x))(S'(x)X) \in \mathfrak{X}\,(VE)$

If $S\,(x) = \varphi_a\,(x, \sigma_a\,(x)) : \nabla_X S(x) = \sum_{\alpha i}(\partial_\alpha \sigma^i_a + \Gamma_a(S\,(x))^i_\alpha)X^\alpha_a \partial u_i$

Notice the difference : the connection Φ acts on sections of TE, the covariant derivative ∇ acts on sections of E.

The covariant derivative is also called the absolute differential.

∇S is linear with respect to the vector field : $\nabla_{X+Y} S(x) = \nabla_X S(x) + \nabla_Y S(x), \nabla_{kX} S(x) = k \nabla_X S(x)$ but we cannot say anything about linearity with respect to S in a general fiber bundle.

Definition 2054 *A section S is said to be an integral of a connection on a fiber bundle with covariant derivative ∇ if $\nabla S = 0$.*

Theorem 2055 *(Giachetta p.33) For any global section S on a fiber bundle $E\,(M, V, \pi)$ there is always a connection such that S is an integral*

26.1.3   Lift to a fiber bundle

A connection is a projection on the vertical bundle. Similarly we can define a projection on the horizontal bundle.

Horizontal form

Definition 2056 *The **horizontal form** of a connection Φ on a fiber bundle $E\,(M, V, \pi)$ is the 1 form*

$\chi \in \wedge_1\,(E; HE) : \chi\,(p) = Id_{TE} - \Phi\,(p)$

A connection can be equivalently defined by its horizontal form $\chi \in \wedge_1\,(E; HE)$ and we have :

$\chi \circ \chi = \chi;$

$\chi(\Phi) = 0;$

$V_p E = \ker \chi;$

$$\chi\left(p\right)v_{p} = \varphi_{a}'(x,u)\left(v_{x}, v_{u}\right) - \varphi_{a}'(x,u)\left(0, v_{u} + \Gamma\left(p\right)v_{x}\right) = \varphi_{a}'(x,u)\left(v_{x}, -\Gamma\left(p\right)v_{x}\right) \in$$
$H_{p}E$

The horizontal form is directly related to TM, as we can see in the formula above which involves only v_{x} . So we can "lift" any object defined on TM onto TE by "injecting" $v_{x} \in T_{x}M$ in the formula.

Horizontal lift of a vector field

Definition 2057 *The **horizontal lift** of a vector field on M by a connection Φ on a fiber bundle $E\left(M, V, \pi\right)$ with trivialization φ is the map : $\chi_{L} : \mathfrak{X}\left(TM\right) \rightarrow \mathfrak{X}\left(HE\right) :: \chi_{L}\left(p\right)\left(X\right) = \varphi'(x,u)\left(X\left(x\right), -\Gamma\left(p\right)X\left(x\right)\right)$*

$\chi_{L}\left(p\right)\left(X\right)$ is a horizontal vector field on TE, which is projectable on X.

Theorem 2058 *(Kolar p.378) For any connection Φ on a fiber bundle $E\left(M, V, \pi\right)$ with covariant derivative ∇ , the horizontal lift $\chi_{L}\left(X\right)$ of a vector field X on M is a projectable vector field on E and for any section $S \in \mathfrak{X}\left(E\right)$: $\nabla_{X}S = \pounds_{\chi_{L}(X)}S$*

Horizontal lift of a curve
  By lifting the tangent to a curve we can lift the curve itself.

Theorem 2059 *(Kolar p. 80) For any connection Φ on a fiber bundle $E\left(M, V, \pi\right)$ and path c:$[a,b] \rightarrow M$ in M, with $0 \in [a,b], c\left(0\right) = x, A \in \pi^{-1}\left(x\right)$ there is a neighborhood $n(x)$ and a smooth map : $P : n\left(x\right) \rightarrow E$ such that :*
  i) $P(c(t))$ is a unique path in E with $\pi\left(P\left(c\left(t\right)\right)\right) = c\left(t\right)$, $P(x)=A$
  ii) $\Phi\left(\frac{dP}{dt}\right) = 0$ when defined
  iii) the curve is inchanged in a smooth change of parameter
  iv) if c depends also smoothly of other parameters, then P depends smoothly on those parameters

  P is defined through the equation : $\Phi\left(\frac{dP}{dt}\right) = 0$ that is equivalent to :
  $\Phi(P\left(c\left(t\right)\right)\left(\frac{dP}{dt}\right) = \Phi(P\left(c\left(t\right)\right)\left(P'\frac{dc}{dt}\right) = \nabla_{c'(t)}P\left(c\left(t\right)\right) = 0$

Theorem 2060 *(Kolar p.81) A connection is said to be **complete** if the lift is defined along any path on M. Each fiber bundle admits complete connections.*

  A complete connection is sometimes called an Ehresmann connection.
  Remarks : the lift is sometime called "parallel transport" (Kolar), but there are significant differences with what is defined usually on a simple manifold.
  i) the lift transports curves on the base manifold to curves on the fiber bundle, whereas the parallel transport transforms curves in the same manifold.
  ii) a vector field on a manfold can be parallel transported, but there is no "lift" of a section on a vector bundle. But a section can be parallel transported by the flow of a projectable vector field.
  iii) there is no concept of geodesic on a fiber bundle. It would be a curve such that its tangent is parallel transported along the curve, which does not apply

here. Meanwhile on a fiber bundle there are horizontal curves : $C : [a, b] \rightarrow E$ such that $\Phi\left(C\left(t\right)\right)\left(\frac{dC}{dt}\right) = 0$ so its tangent is a horizontal vector. Given any curve c(t) on M there is always a horizontal curve which projects on c(t), this is just the lift of c.

Holonomy group

If the path c is a loop in M : $c : [a, b] \rightarrow M :: c(a) = c(b) = x$, the lift goes from a point A$\in E\left(x\right)$ to a point B in the same fiber E(x) over x, so we have a map in V : $A = \varphi\left(x, u_A\right) \rightarrow B = \varphi\left(x, u_B\right) :: u_B = \phi\left(u_A\right)$. This map has an inverse (take the opposite loop with the reversed path) and is a diffeomorphism in V. The set of all these diffeomorphisms has a group structure : this is the **holonomy group** at x $H\left(\Phi, x\right)$. If we restrict the loops to loops which are homotopic to a point we have the restricted holonomy group $H_0\left(\Phi, x\right)$.

26.1.4 Curvature

There are several objects linked to connections which are commonly called curvature. The following is the curvature of the connection Φ.

Definition 2061 *The **curvature** of a connection Φ on the fiber bundle $E\left(M, V, \pi\right)$ is the 2-form $\Omega \in \wedge_2\left(E; VE\right)$ such that for any vector field X,Y on E : $\Omega(X, Y) = \Phi([\chi X, \chi Y]_{TE})$ where χ is the horizontal form of Φ*

Theorem 2062 *The local components of the curvature are given by the Maurer-Cartan formula :*

$$\varphi_a^* \Omega = \sum_i \left(-d_M \Gamma^i + [\Gamma, \Gamma]_V^i\right) \otimes \partial u_i$$

$$\Omega = \sum_{\alpha\beta} \left(-\partial_\alpha \Gamma_\beta^i + \Gamma_\alpha^j \partial_j \Gamma_\beta^i\right) dx^\alpha \wedge dx^\beta \otimes \partial u_i$$

the holonomic basis of TE is $(\partial x_\alpha, \partial u_i)$

Notice that the curvature is a 2-form on E (and not M), in the bracket we have χX and not $\chi_L X$,whence the notation $dx^\alpha \wedge dx^\beta, \partial u_i$.The bracket is well defined for vector fields on the tangent bundle. See below the formula.

The curvature is zero if one of the vector X,Y is vertical (because then $\chi X = 0$) so the curvature is an horizontal form, valued in the vertical bundle.

Proof. in an atlas (O_a, φ_a) of E:

$\chi\left(p\right) X_p = \varphi_a'(x, u)\left(v_x, -\Gamma\left(p\right) v_x\right) = v_x^\alpha \partial x_\alpha - \Gamma\left(p\right)_\alpha^i v_x^\alpha \partial u_i$

$\chi\left(p\right) Y_p = \varphi_a'(x, u)\left(w_x, -\Gamma\left(p\right) w_x\right) = w_x^\alpha \partial x_\alpha - \Gamma\left(p\right)_\alpha^i w_x^\alpha \partial u_i$

$[\left(v_x, -\Gamma\left(p\right) v_x\right), \left(w_x, -\Gamma\left(p\right) w_x\right)]$

$= \sum_\alpha \left(\sum_\beta v_x^\beta \partial_\beta w_x^\alpha - w_x^\beta \partial_\beta v_x^\alpha + \sum_j \left(-\Gamma_\beta^j v_x^\beta \partial_j w_x^\alpha + \Gamma_\beta^j w_x^\beta \partial_j v_x^\alpha\right)\right) \partial x_\alpha$

$+ \sum_i (\sum_\alpha v_x^\alpha \partial_\alpha \left(-\Gamma_\beta^i w_x^\beta\right) - w_x^\alpha \partial_\alpha \left(-\Gamma_\beta^i v_x^\beta\right)$

$+ \sum_j \left(-\Gamma_\alpha^j v_x^\alpha\right) \partial_j \left(-\Gamma_\beta^i w_x^\beta\right) - \left(-\Gamma_\alpha^j w_x^\alpha\right) \partial_j \left(-\Gamma_\beta^i v_x^\beta\right) \partial u_i)$

$\Phi\left[\left(v_x, -\Gamma\left(p\right) v_x\right), \left(w_x, -\Gamma\left(p\right) w_x\right)\right]$

$$= \sum_i \left(\sum_\alpha -v_x^\alpha \partial_\alpha \left(\Gamma_\beta^i w_x^\beta \right) + w_x^\alpha \partial_\alpha \left(\Gamma_\beta^i v_x^\beta \right) + \sum_j \Gamma_\alpha^j v_x^\alpha \partial_j \left(\Gamma_\beta^i w_x^\beta \right) - \Gamma_\alpha^j w_x^\alpha \partial_j \left(\Gamma_\beta^i v_x^\beta \right) \right) \partial u_i$$

$$+ \sum_\alpha \Gamma_\alpha^i \left(\sum_\beta v_x^\beta \partial_\beta w_x^\alpha - w_x^\beta \partial_\beta v_x^\alpha \right) \partial u_i$$

$$= \sum \{ -v_x^\alpha w_x^\beta \partial_\alpha \Gamma_\beta^i - \Gamma_\beta^i v_x^\alpha \partial_\alpha w_x^\beta + w_x^\alpha v_x^\beta \partial_\alpha \Gamma_\beta^i + \Gamma_\beta^i w_x^\alpha \partial_\alpha v_x^\beta + \Gamma_\alpha^j v_x^\alpha w_x^\beta \partial_j \Gamma_\beta^i +$$

$$\Gamma_\alpha^j \Gamma_\beta^i v_x^\alpha \partial_j w_x^\beta$$

$$- \Gamma_\alpha^j v_x^\beta w_x^\alpha \partial_j \Gamma_\beta^i - \Gamma_\alpha^j \Gamma_\beta^i w_x^\alpha \partial_j v_x^\beta + \Gamma_\alpha^i \left(v_x^\beta \partial_\beta w_x^\alpha - w_x^\beta \partial_\beta v_x^\alpha \right) \} \partial u_i$$

$$= \sum \{ \left(w_x^\alpha v_x^\beta - v_x^\alpha w_x^\beta \right) \partial_\alpha \Gamma_\beta^i + \left(v_x^\alpha w_x^\beta - v_x^\beta w_x^\alpha \right) \Gamma_\alpha^j \partial_j \Gamma_\beta^i$$

$$+ \Gamma_\beta^i \left(w_x^\alpha \partial_\alpha v_x^\beta - v_x^\alpha \partial_\alpha w_x^\beta \right) + \Gamma_\alpha^i \left(v_x^\beta \partial_\beta w_x^\alpha - w_x^\beta \partial_\beta v_x^\alpha \right) \} \partial u_i$$

$$= \sum \left(v_x^\alpha w_x^\beta - v_x^\beta w_x^\alpha \right) \left(-\partial_\alpha \Gamma_\beta^i + \Gamma_\alpha^j \partial_j \Gamma_\beta^i \right) \partial u_i$$

$$+ \Gamma_\alpha^i \left(w_x^\beta \partial_\beta v_x^\alpha - v_x^\beta \partial_\beta w_x^\alpha \right) + \Gamma_\alpha^i \left(v_x^\beta \partial_\beta w_x^\alpha - w_x^\beta \partial_\beta v_x^\alpha \right) \partial u_i$$

$$= \sum \left(v_x^\alpha w_x^\beta - v_x^\beta w_x^\alpha \right) \left(-\partial_\alpha \Gamma_\beta^i + \Gamma_\alpha^j \partial_j \Gamma_\beta^i \right) \partial u_i$$

$$\Omega = -\partial_\alpha \Gamma_\beta^i dx^\alpha \otimes dx^\beta \otimes \partial u_i + \partial_\alpha \Gamma_\beta^i dx^\beta \otimes dx^\alpha \otimes \partial u_i + \Gamma_\alpha^j \partial_j \Gamma_\beta^i dx^\alpha \otimes dx^\beta \otimes$$

$$\partial u_i - \Gamma_\alpha^j \partial_j \Gamma_\beta^i dx^\beta \otimes dx^\alpha \otimes \partial u_i$$

$$\Omega = \sum_{\alpha\beta} \left(-\partial_\alpha \Gamma_\beta^i + \Gamma_\alpha^j \partial_j \Gamma_\beta^i \right) dx^\alpha \wedge dx^\beta \otimes \partial u_i$$

The sign - on the first term comes from the convention in the definition of Γ. ∎

Theorem 2063 *For any vector fields X, Y on M :*

$$\nabla_X \circ \nabla_Y - \nabla_Y \circ \nabla_X = \nabla_{[X,Y]} + \pounds_{\Omega(\chi_L(X), \chi_L(Y))}$$

$[\chi_L(X), \chi_L(Y)]_{TE}$ *is a projectable vector field :* $\pi_* \left([\chi_L(X), \chi_L(Y)]_{TE} \right) (\pi(p)) =$
$[\pi_* \chi_L(X), \pi_* \chi_L(Y)] = [X, Y]$

$$\Omega(p)(\chi_L(X), \chi_L(Y)) = [\chi_L(X), \chi_L(Y)]_{TE} - \chi_L([X,Y]_{TM})$$

Proof. $[\chi_L(X), \chi_L(Y)] = [\varphi_a'(x,u)(X(x), -\Gamma(p) X(x)), \varphi_a'(x,u)(Y(x), -\Gamma(p) Y(x))]$

The same computation as above gives :

$$[\chi_L(X), \chi_L(Y)] = \sum_\alpha \left(\sum_\beta X^\beta \partial_\beta Y^\alpha - Y^\beta \partial_\beta X^\alpha \right) \partial x_\alpha$$

$$+ \sum_i (\sum_{\alpha\beta} X^\alpha \partial_\alpha \left(-\Gamma_\beta^i Y^\beta \right) - Y^\alpha \partial_\alpha \left(-\Gamma_\beta^i X^\beta \right)$$

$$+ \sum_{j\alpha\beta} \left(-\Gamma_\alpha^j X^\alpha \right) \partial_j \left(-\Gamma_\beta^i Y^\beta \right) - \left(-\Gamma_\alpha^j Y^\alpha \right) \partial_j \left(-\Gamma_\beta^i X^\beta \right) \partial u_i)$$

$$= \sum_\alpha \left(\sum_\beta X^\beta \partial_\beta Y^\alpha - Y^\beta \partial_\beta X^\alpha \right) \partial x_\alpha$$

$$+ \sum_i (\sum_{\alpha\beta} -X^\alpha Y^\beta \partial_\alpha \Gamma_\beta^i - X^\alpha \Gamma_\beta^i \partial_\alpha Y^\beta + Y^\alpha X^\beta \partial_\alpha \Gamma_\beta^i + \Gamma_\beta^i Y^\alpha \partial_\alpha X^\beta$$

$$+ \sum_j \Gamma_\alpha^j X^\alpha Y^\beta \partial_j \Gamma_\beta^i - \Gamma_\alpha^j Y^\alpha X^\beta \partial_j \Gamma_\beta^i \partial u_i)$$

$$= \sum_\alpha \left(\sum_\beta X^\beta \partial_\beta Y^\alpha - Y^\beta \partial_\beta X^\alpha \right) \left(\partial x_\alpha - \sum_i \Gamma_\alpha^i \partial u_i \right)$$

$$+ \sum_i \sum_{\alpha\beta} \left(X^\alpha Y^\beta - Y^\alpha X^\beta \right) \left(-\partial_\alpha \Gamma_\beta^i + \sum_j \Gamma_\alpha^j \partial_j \Gamma_\beta^i \right) \partial u_i$$

$$= \chi_L([X,Y]) + \Omega(\chi_L(X), \chi_L(Y))$$

Moreover for any projectable vector fields W,U : $\pounds_{[W,U]} S = \pounds_W \circ \pounds_U S - \pounds_U \circ \pounds_W S$

So : $\pounds_{\Omega(\chi_L(X), \chi_L(Y)) + \chi_L(p)([X,Y]_{TM})} = \pounds_{\chi_L(X)} \circ \pounds_{\chi_L(Y)} - \pounds_{\chi_L(X)} \circ \pounds_{\chi_L(Y)} = \pounds_{\Omega(\chi_L(X), \chi_L(Y))} + \pounds_{\chi_L(p)([X,Y]_{TM})}$

$\Omega \left(\chi_L \left(X \right) , \chi_L \left(Y \right) \right)$ is a vertical vector field, so projectable in 0

$\nabla_X \circ \nabla_Y - \nabla_Y \circ \nabla_X = \nabla_{[X,Y]} + \mathcal{L}_{\Omega(\chi_L(X),\chi_L(Y))}$ ∎

So $\Omega \circ \chi_L$ is a 2-form on M :

$\Omega \circ \chi_L : TM \times TM \to TE :: \Omega \left(p \right) \left(\chi_L \left(X \right) , \chi_L \left(Y \right) \right) = \left[\chi_L \left(X \right) , \chi_L \left(Y \right) \right]_{TE} - \chi_L \left(p \right) \left(\left[X, Y \right]_{TM} \right)$

which measures the obstruction against lifting the commutator of vectors.

The horizontal bundle HE is integrable if the connection has a null curvature (Kolar p.79).

26.2   Connections on vectors bundles

The main feature of vector bundles is the vector space structure of E itself, and the linearity of the transition maps. So the connections specific to vector bundles are linear connections.

26.2.1   Linear connection

Each fiber E(x) has the structure of a vector space and in an atlas $(O_a, \varphi_a)_{a \in A}$ of E the chart φ_a is linear with respect to u.

The tangent bundle of the vector bundle $E \left(M, V, \pi \right)$ is the vector bundle $TE \left(TM, V \times V, \pi \times \pi' \right)$. With a holonomic basis $(e_i \left(x \right))_{i \in I}$ a vector $v_p \in T_p E$ reads : $v_p = \sum_{\alpha \in A} v_x^\alpha \partial x^\alpha + \sum_{i \in I} v_u^i e_i \left(x \right)$ and the vertical bundle is isomorphic to ExE : $VE \left(M, V, \pi \right) \simeq E \times_M E$.

A connection $\Phi \in \wedge_1 \left(E; VE \right)$ is defined by a family of Chrisfoffel forms $\Gamma_a \in C \left(\pi^{-1} \left(O_a \right) ; TM^* \otimes V \right) : \Gamma_a \left(p \right) \in \Lambda_1 \left(M; V \right)$

Φ reads in this basis : $\Phi(p) \left(v_x^\alpha \partial x_\alpha + v_u^i e_i \left(x \right) \right) = \sum_i \left(v_u^i + \sum_\alpha \Gamma \left(p \right)_\alpha^i v_x^\alpha \right) e_i \left(x \right)$

Definition 2064 *A **linear connection** Φ on a vector bundle is a connection such that its Christoffel forms are linear with respect to the vector space structure of each fiber :*

$\forall L \in \mathcal{L} \left(V; V \right) , v_x \in T_x M : \Gamma_a \left(\varphi_a \left(x, L \left(u_a \right) \right) \right) v_x = L \circ \left(\Gamma_a \left(\varphi \left(x, u_a \right) \right) v_x \right)$

A linear connection can then be defined by maps with domain in M.

Theorem 2065 *A linear connection Φ on a vector bundle $E \left(M, V, \pi \right)$ with atlas $(O_a, \varphi_a)_{a \in A}$ is uniquely defined by a family $(O_a, \Gamma_a)_{a \in A}$ of 1-form on M valued in $\mathcal{L}(V; V) : \Gamma_a \in \Lambda_1 \left(O_a; \mathcal{L} \left(V; V \right) \right)$ such that at the transitions $\Gamma_b \left(x \right) = -\varphi'_{ba} \left(x \right) \varphi_{ab} \left(x \right) + \Gamma_a \left(x \right)$*

Proof. i) Let Φ be a linear connection :

$\Phi(p) \left(v_x^\alpha \partial x_\alpha + v_u^i e_i \left(x \right) \right) = \sum_i \left(v_u^i + \sum_\alpha \widetilde{\Gamma} \left(p \right)_\alpha^i v_x^\alpha \right) e_i \left(x \right)$

At the transitions :

$x \in O_a \cap O_b : p = \varphi_b \left(x, u_b \right) = \varphi_a \left(x, u_a \right)$

$\widetilde{\Gamma}_b\left(p\right) = \varphi'_{ba}\left(x, u_a\right) \circ \left(-Id_{TM}, \widetilde{\Gamma}_a\left(p\right)\right) = -\varphi'_{ba}\left(x\right) u_a + \varphi_{ba}\left(x\right) \widetilde{\Gamma}_a\left(p\right)$

$\widetilde{\Gamma}_b\left(\varphi_b\left(x, u_b\right)\right) = -\varphi'_{ba}\left(x\right) u_a + \varphi_{ba}\left(x\right) \widetilde{\Gamma}_a\left(\varphi_a\left(x, u_a\right)\right)$

$= -\varphi'_{ba}\left(x\right) u_a + \widetilde{\Gamma}_a\left(\varphi_a\left(x, \varphi_{ba}\left(x\right) u_a\right)\right) = -\varphi'_{ba}\left(x\right) u_a + \widetilde{\Gamma}_a\left(\varphi_a\left(x, u_b\right)\right)$

Define : $\Gamma_a \in \Lambda_1\left(O_a; \mathcal{L}\left(V; V\right)\right) :: \Gamma_a\left(x\right)\left(v_x\right)\left(u\right) = \widetilde{\Gamma}_a\left(\varphi_a\left(x, u\right)\right) v_x$

$\widetilde{\Gamma}_b\left(\varphi_b\left(x, u_b\right)\right) = \Gamma_b\left(x\right)\left(u_b\right) = -\varphi'_{ba}\left(x\right) \varphi_{ab}\left(x\right) u_b + \Gamma_a\left(x\right)\left(u_b\right)$

$\Gamma_b\left(x\right) = -\varphi'_{ba}\left(x\right) \varphi_{ab}\left(x\right) + \Gamma_a\left(x\right)$

ii) Conversely if there is a family $\Gamma_a \in \Lambda_1\left(O_a; \mathcal{L}\left(V; V\right)\right) :: \Gamma_b\left(x\right) = -\varphi'_{ba}\left(x\right) \varphi_{ab}\left(x\right) + \Gamma_a\left(x\right)$

$\widetilde{\Gamma}_a\left(\varphi_a\left(x, u\right)\right) v_x = \Gamma_a\left(x\right)\left(v_x\right)\left(u\right)$ defines a family of linear Christoffel forms

At the transitions :

$\widetilde{\Gamma}_b\left(\varphi_b\left(x, u_b\right)\right) v_x = \Gamma_b\left(x\right)\left(v_x\right)\left(u_b\right) = -\varphi'_{ba}\left(x\right) v_x \varphi_{ab}\left(x\right) u_b + \left(\Gamma_a\left(x\right) v_x\right) u_b = -\varphi'_{ba}\left(x\right) v_x u_a + \left(\Gamma_a\left(x\right) v_x\right) \varphi_{ba}\left(x\right) u_a$

$= \left(\varphi_{ba}\left(x\right) u_a\right)' \left(-v_x, \widetilde{\Gamma}_a\left(\varphi_a\left(x, u_a\right)\right) v_x\right)$ ∎

In a holonomic basis $e_{ai}\left(x\right) = \varphi_a\left(x, e_i\right)$ the maps Γ still called Christoffel forms are :

$\Gamma\left(x\right) = \sum_{ij\alpha} \Gamma^i_{\alpha j}\left(x\right) d\xi^\alpha \otimes e_i\left(x\right) \otimes e^j\left(x\right) \leftrightarrow \widetilde{\Gamma}\left(\varphi\left(x, u\right)\right) = \sum_{ij\alpha} \Gamma^i_{\alpha j}\left(x\right) u^j d\xi^\alpha \otimes e_i\left(x\right)$

$\Phi(\varphi\left(x, \sum_{i \in I} u^i e_i\right))\left(v_x^\alpha \partial x_\alpha + v_u^i e_i\left(x\right)\right) = \sum_i \left(v_u^i + \sum_{j\alpha} \Gamma^i_{\alpha j}\left(x\right) u^j v_x^\alpha\right) e_i\left(x\right)$

The horizontal form of a linear connection Φ on a vector bundle $E\left(M, V, \pi\right)$ is the 1 form

$\chi \in \wedge_1\left(E; HE\right) : \chi\left(p\right) = Id_{TE} - \Phi\left(p\right)$

$\chi\left(p\right)\left(\varphi\left(x, \sum_{i \in I} u^i e_i\right)\right)\left(v_x^\alpha \partial x_\alpha + v_u^i e_i\left(x\right)\right) = \sum_i \left(v_u^i + \sum_{j\alpha} \Gamma^i_{\alpha j}\left(x\right) u^j v_x^\alpha\right) e_i\left(x\right)$

The horizontal lift of a vector on TM is :

$\chi_L\left(\varphi\left(x, u\right)\right)\left(v_x\right) = \varphi'\left(x, u\right)\left(v_x, -\Gamma\left(p\right) v_x\right) = \sum_\alpha v_x^\alpha \partial x^\alpha - \sum_{i\alpha} \Gamma^i_{j\alpha}\left(x\right) u^j v_x^\alpha e_i\left(x\right)$

26.2.2 Curvature

Definition 2066 *For any linear connection Φ defined by the Christoffel form Γ on the vector bundle E, there is a 2-form $\Omega \in \Lambda_2\left(E; E \otimes E^*\right) : \Omega\left(x\right) = \sum_{\alpha\beta} \sum_{j \in I} \left(-\partial_\alpha \Gamma^i_{j\beta}\left(x\right) + \sum_{k \in I} \Gamma^k_{j\alpha}\left(x\right) \Gamma^i_{k\beta}\left(x\right)\right) dx^\alpha \wedge dx^\beta \otimes e_i\left(x\right) \otimes e^j\left(x\right)$ in a holonomic basis of E such that the curvature form $\widetilde{\Omega}$ of Φ : $\widetilde{\Omega}\left(\sum_i u^i e_i\left(x\right)\right) = \sum_j u^j \Omega_j\left(x\right)$*

Proof. The Cartan formula gives for the curvature with $\partial u_i = e_i\left(x\right)$ in a holonomic basis :

$\widetilde{\Omega}\left(p\right) = \sum_{\alpha\beta} \left(-\partial_\alpha \widetilde{\Gamma}^i_\beta + \widetilde{\Gamma}^j_\alpha \partial_j \widetilde{\Gamma}^i_\beta\right) dx^\alpha \wedge dx^\beta \otimes e_i\left(x\right)$

So with a linear connection :

$\widetilde{\Gamma}\left(\varphi\left(x, \sum_{i \in I} u^i e_i\right)\right) = \sum_{i, j \in I} \Gamma^i_{\beta j}\left(x\right) u^j d\xi^\beta \otimes e_i\left(x\right)$

$\partial_\alpha \widetilde{\Gamma}^i_\beta\left(\varphi\left(x, \sum_{i \in I} u^i e_i\right)\right) = \sum_{k \in I} u^k \partial_\alpha \Gamma^i_{\beta k}\left(x\right)$

$\partial_j \widetilde{\Gamma}^i_\beta\left(\varphi\left(x, \sum_{i \in I} u^i e_i\right)\right) = \Gamma^i_{\beta j}\left(x\right)$

$\widetilde{\Omega}\left(p\right)$

$= \sum_{\alpha\beta} \left(-\sum_{j \in I} u^j \partial_\alpha \Gamma^i_{j\beta}\left(x\right) + \sum_{j, k \in I} u^k \Gamma^j_{k\alpha}\left(x\right) \Gamma^i_{j\beta}\left(x\right)\right) dx^\alpha \wedge dx^\beta \otimes e_i\left(x\right)$

$$= \sum_{\alpha\beta} \sum_{j\in I} u^j \left(-\partial_\alpha \Gamma^i_{j\beta}(x) + \sum_{k\in I} \Gamma^k_{j\alpha}(x) \Gamma^i_{k\beta}(x) \right) dx^\alpha \wedge dx^\beta \otimes e_i(x)$$

$$\widetilde{\Omega}(\varphi(x,u)) = \sum_j u^j \Omega_j(x) \quad \blacksquare$$

26.2.3   Covariant derivative

Covariant derivative of sections

The covariant derivative of a section X on a vector bundle $E(M,V,\pi)$ is a map : $\nabla X \in \Lambda_1(M;E) \simeq TM^* \otimes E$. It is independant of the trivialization. It has the following coordinates expression for a linear connection:

Theorem 2067 *The covariant derivative ∇ associated to a linear connection Φ on a vector bundle $E(M,V,\pi)$ with Christoffel form Γ is the map : $\nabla : \mathfrak{X}(E) \to \Lambda_1(M;E) :: \nabla X(x) = \sum_{\alpha i}(\partial_\alpha X^i + X^j \Gamma^i_{j\alpha}(x)) d\xi^\alpha \otimes e_i(x)$ in a holonomic basis of E*

Proof. A section on E is a vector field : $X : M \to E :: X(x) = \sum_i X^i(x) e_i(x)$ with $\pi(X) = x$ .Its derivative is

$X'(x) : T_x M \to T_{X(x)} E :: X'(x) = \sum_{i,\alpha} \left(\partial_\alpha X^i(x) \right) e_i(x) \otimes d\xi^\alpha \in T_x M^* \otimes E(x)$ with a dual holonomic basis $d\xi^\alpha \in T_x M^*$ $\blacksquare$

The covariant derivative of a section *on E* along a vector field *on M* is a *section on E* which reads :

$\nabla_Y X(x) = \sum_{\alpha i}(\partial_\alpha X^i + X^j \Gamma^i_{j\alpha}(x)) Y^\alpha e_i(x) \in \mathfrak{X}(E)$

With the tangent bundle TM of a manifold we get back the usual definition of the covariant derivative (which justifies the choice of the sign before Γ).

Covariant derivative of tensor fields

Tensorial functors $\mathfrak{T}^r_s$ can extend a vector bundle $E(M,V,\pi)$ to a tensor bundle $\mathfrak{T}^r_s E(M,\mathfrak{F}^r_s V,\pi)$. There are connections defined on these vector bundles as well, but we can extend a linear connection from E to $\mathfrak{T}^r_s E$.

If we look for a derivation D on the algebra $\mathfrak{T}^r_s E(x)$ at some fixed point x, with the required properties listed in Differential geometry (covariant derivative), we can see, by the same reasonning, that this derivation is determined by its value over a basis:

$\nabla e_i(x) = \sum_{i,j,\alpha} \Gamma^j_{i\alpha}(x) d\xi^\alpha \otimes e_j(x) \in T_x M^* \otimes E(x)$

$\nabla e^i(x) = -\sum_{i,j,\alpha} \Gamma^i_{j\alpha} d\xi^\alpha \otimes e^j(x) \in T_x M^* \otimes E^*(x)$

with the dual bases $e^i(x), d\xi^\alpha$ of $e_i(x), \partial \xi_\alpha$ in the fiber E(x) and the tangent space $T_x M$.

Notice that we have $d\xi^\alpha \in T_x M^*$ because the covariant derivative acts on vector fields on M. So the covariant derivative is a map : $\nabla : \mathfrak{T}^r_s E \to TM^* \otimes \mathfrak{T}^r_s E$

The other properties are the same as the usual covariant derivative on the tensorial bundle $\otimes^r_s M$

Linearity (with scalars k,k') :

$\forall S,T \in \mathfrak{T}^r_s E, k, k' \in K : \nabla(kS + k'T) = k\nabla S + k'\nabla T$

$\forall Y,Z \in VM, k, k' \in K : \nabla_{kY+k'Z} S = k\nabla_Y S + k'\nabla_Z S$

Leibnitz rule with respect to the tensorial product (because it is a derivation):

$\nabla (S \otimes T) = (\nabla S) \otimes T + S \otimes (\nabla T)$

If $f \in C_1(M; \mathbb{R}), Y \in \otimes^1 E : \nabla_X f Y = df(X) Y + f \nabla_X Y$

Commutative with the trace operator :

$\nabla(Tr(T)) = Tr(\nabla T)$

The formulas of the covariant derivative are :

for a section on E :

$X = \sum_{i \in I} X^i(x) e_i(x) \rightarrow \nabla_Y X(x) = \sum_{\alpha i}(\partial_\alpha X^i + X^j \Gamma^i_{j\alpha}(x)) Y^\alpha e_i(x)$

for a section of E* :

$\varpi = \sum_{i \in I} \varpi_i(x) e^i(x) \rightarrow \nabla_Y \varpi = \sum_{\alpha i} \left(\partial_\alpha \varpi_i - \Gamma^j_{i\alpha} \varpi_j \right) Y^\alpha \otimes e^i(x)$

for a mix tensor, section of $\mathfrak{T}^r_s E$:

$T(x) = \sum_{i_1...i_r} \sum_{j_1....j_s} T^{i_1...i_r}_{j_1...j_s}(x) e_{i_1}(x) \otimes .. \otimes e_{i_r}(x) \otimes e^{j_1}(x) \otimes ... \otimes e^{j_s}(x)$

$\nabla T = \sum_{i_1...i_r} \sum_{j_1....j_s} \sum_\alpha \widehat{T}^{i_1...i_r}_{\alpha j_1...j_s} d\xi^\alpha \otimes e_{i_1}(x) \otimes .. \otimes e_{i_r}(x) \otimes e^{j_1}(x) \otimes ... \otimes e^{j_s}(x)$

$\widehat{T}^{i_1...i_r}_{\alpha j_1...j_s} = \partial_\alpha T^{i_1...i_r}_{j_1...j_s} + \sum_{k=1}^{r} \Gamma^{i_k}_{m\alpha} T^{i_1.i_{k-1}mi_{k+1}..i_r}_{j_1...j_s} - \sum_{k=1}^{s} \Gamma^{m}_{j_k\alpha} T^{i_1...i_r}_{j_1..j_{k-1}mj_{k+1}..j_s}$

Horizontal lift of a vector field

The horizontal lift of a vector field $X \in \mathfrak{X}(TM)$ on a vector bundle $E(M, V, \pi)$ by a linear connection on HE is the linear map :

$\chi_L : \mathfrak{X}(TM) \rightarrow \mathfrak{X}(HE) :: \chi_L(e_i(x))(X) = \sum_{i\alpha} \left(\partial x_\alpha - \Gamma^j_{\alpha i}(x) e_j(x) \right) X^\alpha(x)$

in a holonomic basis of TE

This a projectable vector field $\pi'(e_i(x))(\chi_L(e_i(x))(X)) = X$ and for any $S \in \mathfrak{X}(E) : \nabla_X S = \pounds_{\chi_L(X)} S$

Notice that this is a lift to HE and not E. Indeed a lift of a vector field X in TM on E is given by the covariant derivative $\nabla_X S$ of a section S on E along X.

Lift of a curve

Theorem 2068 *For any path $c \in C_1([a; b]; M)$ with $0 \in [a; b]$ there is a unique path $C \in C_1([a; b]; E)$ with $C(0)=\pi^{-1}(c(0))$ lifted on the vector bundle $E(M, V, \pi)$ by the linear connection with Christoffel form Γ such that : $\nabla_{c'(t)} C = 0, \pi(C(t)) = c(t)$*

Proof. C is defined in a holonomic basis by :

$C(t) = \sum_i C^i(t) e_i(c(t))$

$\nabla_{c'(t)} C = \sum_i \left(\frac{dC^i}{dt} + \sum_{ij\alpha} \Gamma^i_{j\alpha}(c(t)) C^j(t) \frac{dc}{dt} \right) e_i(c(t)) = 0$

$\forall i : \frac{dC^i}{dt} + \sum_{j\alpha} \Gamma^i_{j\alpha}(c(t)) C^j(t) \frac{dc}{dt} = 0$

$C(0)=C_0 = \sum_i C^i(t) e_i(c(0))$

This is a linear ODE which has a solution. ∎

Notice that this is a lift of a curve on M to a curve in E (and not TE).

26.2.4 Exterior covariant derivative

Exterior covariant derivative

The formulas above apply to any tensorial section on E. In differential geometry one defines the exterior covariant derivative of a r-form ϖ on M valued in the tangent bundle TM : $\varpi \in \wedge_r (M; TM)$. There is an exterior covariant derivative of r forms *on M* valued *in a vector bundle E* (and not TE). The original formula, which needs both the commutator of vector fields and the exterior differential of forms that are not defined on E, cannot be implemented. But there is an alternate definition which gives the same results.

A r-form ϖ *on M* valued *in the vector bundle E* : $\varpi \in \wedge_r (M; E)$ reads in an holonomic basis of M and a basis of E:

$\varpi = \sum_{i\{\alpha_1...\alpha_r\}} \varpi^i_{\alpha_1...\alpha_r} (x) d\xi^{\alpha_1} \wedge d\xi^{\alpha_2} \wedge ... \wedge d\xi^{\alpha_r} \otimes e_i (x)$

Definition 2069 *The exterior covariant derivative ∇_e of r-forms ϖ on M valued in the vector bundle $E (M, V, \pi)$, is a map : $\nabla_e : \Lambda_r (M; E) \to \Lambda_{r+1} (M; E)$. For a linear connection with Christoffel form Γ it is given in a holonomic basis by the formula:* $\nabla_e \varpi = \sum_i \left(d\varpi^i + \left(\sum_j \left(\sum_\alpha \Gamma^i_{j\alpha} d\xi^\alpha \right) \wedge \varpi^j \right) \right) \otimes e_i (x)$

the exterior differential $d\varpi^i$ is taken on M.

Theorem 2070 *(Kolar p.112) The exterior covariant derivative ∇_e on a vector bundle $E (M, V, \pi)$ with linear covariant derivative ∇ has the following properties :*

i) if $\varpi \in \Lambda_0 (M; E) : \nabla_e \varpi = \nabla \varpi$ (we have the usual covariant derivative of a section on E)

ii) the exterior covariant derivative is the only map : $\nabla_e : \Lambda_r (M; E) \to \Lambda_{r+1} (M; E)$ such that :

$\forall \mu_r \in \Lambda_r (M; \mathbb{R}), \forall \varpi_s \in \Lambda_s (M; E) : \nabla_e (\mu_r \wedge \varpi_s) = (d\mu_r) \wedge \varpi_s + (-1)^r \mu_r \wedge \nabla_e \varpi_s$

iii) if $f \in C_\infty (N; M), \varpi \in \Lambda_r (N; f^*E) : \nabla_e (f^*\varpi) = f^* (\nabla_e \varpi)$

Accounting for the last property, we can implement the covariant exterior derivative for $\varpi \in \Lambda_r (N; E)$, meaning when the base of E is not N.

Riemann curvature

Definition 2071 *The **Riemann curvature** of a linear connection Φ on a vector bundle $E (M, V, \pi)$ with the covariant derivative ∇ is the map :*

$\mathfrak{X} (TM)^2 \times \mathfrak{X} (E) \to \mathfrak{X} (E) :: R(Y_1, Y_2)X = \nabla_{Y_1} \nabla_{Y_2} X - \nabla_{Y_2} \nabla_{Y_1} X - \nabla_{[Y_1, Y_2]} X$

The formula makes sense : $\nabla : \mathfrak{X} (E) \to \Lambda_1 (M; E)$ so $\nabla_Y X \in \mathfrak{X} (E)$ and $\nabla_{Y_1} (\nabla_{Y_2} X) \in \mathfrak{X} (E)$

If $X = \partial \xi_\alpha, Y = \partial \xi_\beta$ then $[\partial \xi_\alpha, \partial \xi_\beta] = 0$ and $R(\partial \xi_\alpha, \partial \xi_\beta, X) = \left(\nabla_{\partial \xi_\alpha} \nabla_{\partial \xi_\beta} - \nabla_{\partial \xi_\beta} \nabla_{\partial \xi_\alpha} \right) X$ so R is a measure of the obstruction of the covariant derivative to be commutative. The name is inspired by the corresponding object on manifolds.

Theorem 2072 *The Riemann curvature is a tensor valued in the tangent bundle :* $R \in \wedge_2 (M; E \otimes E')$

$R = \sum_{\{\alpha\beta\}} \sum_{ij} R^i_{j\alpha\beta} d\xi^\alpha \wedge d\xi^\beta \otimes e^j(x) \otimes e_i(x)$ *and* $R^i_{j\alpha\beta} = \partial_\alpha \Gamma^i_{j\beta} - \partial_\beta \Gamma^i_{j\alpha} + \sum_k \left(\Gamma^i_{k\alpha} \Gamma^k_{j\beta} - \Gamma^i_{k\beta} \Gamma^k_{j\alpha} \right)$

Proof. This is a straightforward computation similar to the one given for the curvature in Differential Geometry ■

Theorem 2073 *For any r-form ϖ on M valued in the vector bundle $E(M, V, \pi)$ endowed with a linear connection and covariant derivative ∇ :* $\nabla_e (\nabla_e \varpi) = R \wedge \varpi$ *where R is the Riemann curvature tensor*

More precisely in a holonomic basis :

$\nabla_e (\nabla_e \varpi) = \sum_{ij} \left(\sum_{\alpha\beta} R^i_{j\alpha\beta} d\xi^\alpha \wedge d\xi^\beta \right) \wedge \varpi^j \otimes e_i(x)$

Where $R = \sum_{\alpha\beta} \sum_{ij} R^i_{j\alpha\beta} d\xi^\alpha \wedge d\xi^\beta \otimes e^j(x) \otimes e_i(x)$ and $R^i_{j\alpha\beta} = \partial_\alpha \Gamma^i_{j\beta} + \sum_k \Gamma^i_{k\alpha} \Gamma^k_{j\beta}$

Proof. $\nabla_e \varpi = \sum_i \left(d\varpi^i + \sum_j \Omega^i_j \wedge \varpi^j \right) \otimes e_i(x)$ with $\Omega^i_j = \sum_\alpha \Gamma^i_{j\alpha} d\xi^\alpha$

$\nabla_e (\nabla_e \varpi) = \sum_i \left(d(\nabla_e \varpi)^i + \sum_j \Omega^i_j \wedge (\nabla_e \varpi)^j \right) \otimes e_i(x)$

$= \sum_i \left(d \left(d\varpi^i + \sum_j \Omega^i_j \wedge \varpi^j \right) + \sum_j \Omega^i_j \wedge \left(d\varpi^j + \sum_k \Omega^j_k \wedge \varpi^k \right) \right) \otimes e_i(x)$

$= \sum_{ij} \left(d\Omega^i_j \wedge \varpi^j - \Omega^i_j \wedge d\varpi^j + \Omega^i_j \wedge d\varpi^j + \Omega^i_j \wedge \sum_k \Omega^j_k \wedge \varpi^k \right) \otimes e_i(x)$

$= \sum_{ij} \left(d\Omega^i_j \wedge \varpi^j + \sum_k \Omega^i_k \wedge \Omega^k_j \wedge \varpi^j \right) \otimes e_i(x)$

$\nabla_e (\nabla_e \varpi) = \sum_{ij} \left(d\Omega^i_j + \sum_k \Omega^i_k \wedge \Omega^k_j \right) \wedge \varpi^j \otimes e_i(x)$

$d\Omega^i_j + \sum_k \Omega^i_k \wedge \Omega^k_j = d \left(\sum_\alpha \Gamma^i_{j\alpha} d\xi^\alpha \right) + \sum_k \left(\sum_\alpha \Gamma^i_{k\alpha} d\xi^\alpha \right) \wedge \left(\sum_\beta \Gamma^k_{j\alpha} d\xi^\beta \right)$

$= \sum_{\alpha\beta} \left(\partial_\beta \Gamma^i_{j\alpha} d\xi^\beta \wedge d\xi^\alpha + \sum_k \Gamma^i_{k\alpha} \Gamma^k_{j\beta} d\xi^\alpha \wedge d\xi^\beta \right)$

$= \sum_{\alpha\beta} \left(\partial_\alpha \Gamma^i_{j\beta} + \sum_k \Gamma^i_{k\alpha} \Gamma^k_{j\beta} \right) d\xi^\alpha \wedge d\xi^\beta$ ■

Theorem 2074 *The commutator of vector fields on M lifts to E iff R=0.*

Proof. We have for any connection with $Y_1, Y_2 \in \mathfrak{X}(TM), X \in \mathfrak{X}(E)$

$\nabla_{Y_1} \circ \nabla_{Y_2} X - \nabla_{Y_2} \circ \nabla_{Y_1} X = \nabla_{[Y_1, Y_2]} X + \mathcal{L}_{\Omega(\chi_L(Y_1), \chi_L(Y_2))} X$

So : $R(Y_1, Y_2)X = \mathcal{L}_{\Omega(\chi_L(X), \chi_L(Y))} X =$

$R \left(\varphi \left(x, \sum_i u^i e_i \right) \right)(Y_1, Y_2)X = \sum_i u^i \mathcal{L}_{\widehat{\Omega}(\chi_L(Y_1), \chi_L(Y_2)) e_i(x)} X$

$R(e_i(x))(Y_1, Y_2)X = \mathcal{L}_{\widehat{\Omega}(\chi_L(Y_1), \chi_L(Y_2)) e_i(x)} X$ ■

Theorem 2075 *The exterior covariant derivative of ∇ is :* $\nabla_e (\nabla X) = \sum_{\{\alpha\beta\}} R^i_{j\alpha\beta} X^j d\xi^\alpha \wedge d\xi^\beta \otimes e_i(x)$

Proof. The covariant derivative ∇ is a 1 form valued in E, so we can compute its exterior covariant derivative :

$\nabla X = \sum_{\alpha i}(\partial_\alpha X^i + X^j \Gamma^i_{j\alpha}(x))d\xi^\alpha \otimes e_i(x)$

$\nabla_e(\nabla X) = \sum_i (d\left(\sum_\alpha (\partial_\alpha X^i + X^j \Gamma^i_{j\alpha}(x))d\xi^\alpha\right)$

$+ \sum_j \left(\sum_\beta \Gamma^i_{j\beta}d\xi^\beta\right) \wedge \left(\sum_{\alpha i}(\partial_\alpha X^j + X^k \Gamma^j_{k\alpha}(x))d\xi^\alpha\right)) \otimes e_i(x)$

$= \sum \left(\partial^2_{\beta\alpha}X^i + \Gamma^i_{j\alpha}\partial_\beta X^j + X^j \partial_\beta \Gamma^i_{j\alpha} + \Gamma^i_{j\beta}\partial_\alpha X^j + X^k \Gamma^i_{j\beta}\Gamma^j_{k\alpha}\right) d\xi^\beta \wedge d\xi^\alpha \otimes$
$e_i(x)$

$= \sum \left(\partial^2_{\beta\alpha}X^i\right) d\xi^\beta \wedge d\xi^\alpha \otimes e_i(x) + \left(\Gamma^i_{j\alpha}\partial_\beta X^j + \Gamma^i_{j\beta}\partial_\alpha X^j\right) d\xi^\beta \wedge d\xi^\alpha \otimes e_i(x)$

$+ X^j \left(\partial_\beta \Gamma^i_{j\alpha} + \Gamma^i_{k\beta}\Gamma^k_{j\alpha}\right) d\xi^\beta \wedge d\xi^\alpha \otimes e_i(x)$

$\nabla_e(\nabla X) = \sum X^j \left(\partial_\beta \Gamma^i_{j\alpha} + \Gamma^i_{k\beta}\Gamma^k_{j\alpha}\right) d\xi^\beta \wedge d\xi^\alpha \otimes e_i(x)$

$\nabla_e(\nabla X) = \sum \left(X^j \partial_\beta \Gamma^i_{j\alpha}d\xi^\beta \wedge d\xi^\alpha + X^j \Gamma^i_{k\beta}\Gamma^k_{j\alpha}d\xi^\beta \wedge d\xi^\alpha\right) \otimes e_i(x)$

$\nabla_e(\nabla X) = \sum \left(X^j \partial_\alpha \Gamma^i_{j\beta}d\xi^\alpha \wedge d\xi^\beta + X^j \Gamma^i_{k\alpha}\Gamma^k_{j\beta}d\xi^\alpha \wedge d\xi^\beta\right) \otimes e_i(x)$

$\nabla_e(\nabla X) = \sum X^j \left(\partial_\alpha \Gamma^i_{j\beta} + \Gamma^i_{k\alpha}\Gamma^k_{j\beta}\right) d\xi^\alpha \wedge d\xi^\beta \otimes e_i(x)$

$\nabla_e(\nabla X) = \sum R^i_{j\alpha\beta}X^j d\xi^\alpha \wedge d\xi^\beta \otimes e_i(x)$ ∎

Remarks :
The curvature of the connection reads :
$\Omega = -\sum_{\alpha\beta}\sum_{j\in I}\left(\partial_\alpha \Gamma^i_{j\beta} + \sum_{k\in I}\Gamma^i_{k\alpha}\Gamma^k_{j\beta}\right) dx^\alpha \wedge dx^\beta \otimes e_i(x) \otimes e^j(x)$
The Riemann curvature reads :
$R = \sum_{\alpha\beta}\sum_{ij}\left(\partial_\alpha \Gamma^i_{j\beta} + \sum_k \Gamma^i_{k\alpha}\Gamma^k_{j\beta}\right) d\xi^\alpha \wedge d\xi^\beta \otimes e^j(x) \otimes e_i(x)$

i) the curvature is defined for any fiber bundle, the Riemann curvature is defined only for vector bundle

ii) Both are valued in $E \otimes E'$ but the curvature is a two horizontal form on TE and the Riemann curvature is a two form on TM

iii) the formulas are not identical but opposite of each other

26.2.5 Metric connection

Definition 2076 *A linear connection with covariant derivative ∇ on a vector bundle $E(M, V, \pi_E)$ endowed with a scalar product g is said to be **metric** if $\nabla g = 0$*

We have the general characterization of such connexions :

Theorem 2077 *A linear connection with Christoffel forms Γ on a complex or real vector bundle $E(M, V, \pi_E)$ endowed with a scalar product g is metric if :*
$\forall \alpha, i, j : \partial_\alpha g_{ij} = \sum_k \left(\Gamma^k_{\alpha i}g_{kj} + \Gamma^k_{\alpha j}g_{ik}\right) \Leftrightarrow [\partial_\alpha g] = [\Gamma_\alpha]^t [g] + [g][\Gamma_\alpha]$

Proof. Real case :
the scalar product is defined by a tensor $g \in \mathfrak{X}(\odot_2 E)$
At the transitions : $g_{bij}(x) = \sum_{kl} \overline{[\varphi_{ab}(x)]}^k_i [\varphi_{ab}(x)]^l_j g_{akl}(x)$

The covariant derivative of g reads with the Christoffel forms $\Gamma(x)$ of the connection :

$\nabla g = \sum_{\alpha ij} \left(\partial_\alpha g_{ij} - \sum_k \left(\Gamma^k_{\alpha i} g_{kj} + \Gamma^k_{\alpha j} g_{ik}\right)\right) e^i_a(x) \otimes e^j_a(x) \otimes d\xi^\alpha$

So : $\forall \alpha, i, j : \partial_\alpha g_{ij} = \sum_k \left(\Gamma^k_{\alpha i} g_{kj} + \Gamma^k_{\alpha j} g_{ik}\right)$

$[\partial_\alpha g]^i_j = \sum_k \left([\Gamma_\alpha]^k_i [g]^k_j + [\Gamma_\alpha]^k_j [g]^i_k\right) \Leftrightarrow [\partial_\alpha g] = [\Gamma_\alpha]^t [g] + [g] [\Gamma_\alpha]$

Complex case :

the scalar product is defined by a real structure σ and a tensor $g \in \mathfrak{X}(\otimes_2 E)$ which is not symmetric.

The covariant derivative of g is computed as above with the same result. ■

Remarks :

i) The scalar product of two covariant derivatives, which are 1-forms on M valued in E, has no precise meaning, so it is necessary to go through the tensorial definition to stay rigorous. And there is no clear equivalent of the properties of metric connections on manifolds : preservation of the scalar product of transported vector field, or the formula : $\forall X, Y, Z \in \mathfrak{X}(TM) : R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0$.

ii) A scalar product γ on V induces a scalar product on E iff the transition maps preserve the scalar product : $[\gamma] = [\varphi_{ba}]^* [\gamma] [\varphi_{ba}]$.

If E is real then this scalar product defines a tensor $g_{aij}(x) = g_{aij}(x)(e_i(x), e_j(x)) = \gamma(e_i, e_j)$ which is metric iff $\forall \alpha : [\Gamma_{a\alpha}]^t [\gamma] + [\gamma][\Gamma_{a\alpha}] = 0$ and it is easy t check that if it is met for $[\Gamma_{a\alpha}]$ it is met for $[\Gamma_{b\alpha}] = [\Gamma_{a\alpha}] - [\partial_\alpha \varphi_{ba}][\varphi_{ab}]$.

If E is complex there is a tensor associated to the scalar product iff there is a real structure on E. A real structure σ on V induces a real structure on E iff the transition maps are real : $\varphi_{ab}(x) \circ \sigma = \sigma \circ \varphi_{ab}(x)$

The connection is real iff : $\widetilde{\Gamma}(\varphi(x, \sigma(u))) = \sigma\left(\widetilde{\Gamma}(\varphi(x, u))\right) \Leftrightarrow \Gamma(x) = \overline{\Gamma(x)}$

Then the condition above reads : $[\Gamma_\alpha]^t [\gamma] + [\gamma][\Gamma_\alpha] = 0 \Leftrightarrow [\Gamma_\alpha]^* [\gamma] + [\gamma][\Gamma_\alpha] = 0$

26.3 Connections on principal bundles

The main feature of principal bundles is the right action of the group on the bundle. So connections specific to principal bundle are connections which are equivariant under this action.

26.3.1 Principal connection

Definition

The tangent bundle TP of a principal fiber bundle $P(M, G, \pi)$ is a principal bundle $TP(TM, TG, \pi \times \pi')$. Any vector $v_p \in T_p P$ can be written $v_p = \varphi'_{ax}(x, g)v_x + \zeta(u_g)(p)$ where ζ is a fundamental vector field and $u_g \in T_1 G$: $\zeta(u_g)(p) = \rho'_g(p, 1) X = \varphi'_{ag}(x, g)(L'_g 1) u_g$. The vertical bundle VP is a trivial vector bundle over P : $VP(P, T_1 G, \pi) \simeq P \times T_1 G$.

So a connection Φ on P reads in an atlas $(O_a, \varphi_a)_{a \in A}$ of P:

$\Phi(p)(\varphi'_{ax}(x, g)v_x + \zeta(u_g)(p)) = \varphi'_{ag}(x, g)\left((L'_g 1) u_g + \Gamma_a(p) v_x\right)$

$$= \varphi'_{ag}(x,g) \left(L'_g 1\right) \left(u_g + L'_{g^{-1}}(g)\,\Gamma_a(p)\,v_x\right) = \zeta \left(u_g + L'_{g^{-1}}(g)\,\Gamma_a(p)\,v_x\right)(p)$$

As we can see in this formula the Lie algebra T_1G will play a significant role in a connection : this is a fixed vector space, and through the fundamental vectors formalism any equivariant vector field on VP can be linked to a fixed vector on T_1G .

Definition 2078 *A connection on a principal bundle $P(M,G,\pi)$ is said to be* **principal** *if it is equivariant under the right action of G :*

Theorem 2079 *(Kolar p.101) Any finite dimensional principal bundle admits principal connections*

Theorem 2080 *A connection is principal iff $\Gamma_a(\varphi_a(x,g)) = R'_g(1)\Gamma_a(\varphi_a(x,1))$*

Proof. It is principal iff :

$\forall g \in G, p \in P : \rho(p,g)_*\,\Phi(p) = \rho'_p(p,g)\,\Phi(p) \Leftrightarrow \Phi(\rho(p,g))\rho'_p(p,g)\,v_p = \rho'_p(p,g)\,\Phi(p)v_p$

Take $p = p_a(x) = \varphi_a(x,1)$

$v_p = \varphi'_{ax}(x,1)v_x + \zeta(u_g)(p)$

$\rho'_p(p,g)\,v_p = \varphi'_{ax}(x,1)v_x + \rho'_p(p,g)\,\zeta(u_g)(p) = \varphi'_{ax}(x,g)v_x + \zeta\left(Ad_{g^{-1}}u_g\right)(\rho(p,g))$

$\Phi(\rho(p,g))\rho'_p(p,g)\,v_p = \zeta\left(Ad_{g^{-1}}u_g + L'_{g^{-1}}(g)\,\Gamma_a(\rho(p,g))\,v_x\right)(\rho(p,g))$

$\rho'_p(p,g)\,\Phi(p)v_p = \zeta\left(Ad_{g^{-1}}\left(u_g + \Gamma_a(p)\,v_x\right)\right)(\rho(p,g))$

$Ad_{g^{-1}}u_g + L'_{g^{-1}}(g)\,\Gamma_a(\rho(p,g))\,v_x = Ad_{g^{-1}}\left(u_g + \Gamma_a(p)\,v_x\right)$

$\Gamma_a(\rho(p,g)) = L'_g(1)\,Ad_{g^{-1}}\Gamma_a(p) = L'_g(1)\,L'_{g^{-1}}(g)R'_g(1)\Gamma_a(p)$

So : $\Gamma_a(\varphi_a(x,g)) = R'_g(1)\Gamma_a(\varphi_a(x,1))$ ∎

To use a convenient and usual notation and denomination in physics :

Notation 2081 $\grave{A}_a = \Gamma_a(\varphi_a(x,1)) \in \Lambda_1(O_a; T_1G)$ *is the* **potential** *of the connection*

Theorem 2082 *A principal connection Φ on a principal bundle fiber bundle $P(M,G,\pi)$ with atlas $(O_a, \varphi_a)_{a\in A}$ is uniquely defined by a family $\left(O_a, \grave{A}_a\right)_{a\in A}$ of maps $\grave{A}_a \in \Lambda_1(O_a; T_1G)$ such that : $\grave{A}_b(x) = Ad_{g_{ba}}\left(\grave{A}_a(x) - L'_{g_{ba}^{-1}}(g_{ba})g'_{ba}(x)\right)$ by : $\Phi(\varphi_a(x,g))(\varphi'_{ax}(x,g)v_x + \zeta(u_g)(p)) = \zeta\left(u_g + Ad_{g^{-1}}\grave{A}_a(x)\,v_x\right)(p)$*

So : $\Gamma_a(\varphi_a(x,g)) = R'_g(1)\grave{A}_a(x)$

Proof. i) If Φ is a principal connection :

$$\Phi\left(\varphi_a\left(x,g\right)\right)\left(\varphi'_{ax}(x,g)v_x + \zeta\left(u_g\right)(p)\right) = \zeta\left(u_g + L'_{g^{-1}}\left(g\right)\Gamma_a\left(\varphi_a\left(x,g\right)\right)v_x\right)(p) =$$

$$\zeta\left(u_g + L'_{g^{-1}}\left(g\right)R'_g\left(1\right)\Gamma_a\left(\varphi_a\left(x,1\right)\right)v_x\right)(p)$$

Define : $\grave{A}_a\left(x\right) = \Gamma_a\left(\varphi_a\left(x,1\right)\right) \in \Lambda_1\left(O_a; T_1 G\right)$

$$\Phi\left(\varphi_a\left(x,g\right)\right)\left(\varphi'_{ax}(x,g)v_x + \zeta\left(u_g\right)(p)\right) = \zeta\left(u_g + L'_{g^{-1}}\left(g\right)R'_g\left(1\right)\grave{A}_a\left(x\right)v_x\right)(p) =$$

$$\zeta\left(u_g + Ad_{g^{-1}}\grave{A}_a\left(x\right)v_x\right)(p)$$

$\Gamma_a\left(\varphi_a\left(x,g_a\right)\right) = R'_{g_a}\left(1\right)\grave{A}_a\left(x\right)$

In a transition $x \in O_a \cap O_b : p = \varphi_a\left(x,g\right) = \varphi_b\left(x,g_b\right) \Rightarrow \Gamma_b\left(p\right) = \varphi'_{ba}\left(x,g_a\right)\circ$
$\left(-Id_{TM},\Gamma_a\left(p\right)\right)$

$\Gamma_b\left(p\right) = \left(g_{ba}\left(x\right)\left(g_a\right)\right)'\circ\left(-Id_{TM},\Gamma_a\left(p\right)\right) = -R'_{g_a}\left(g_{ba}(x)\right)g'_{ba}\left(x\right)+L'_{g_{ba}(x)}\left(g_a\right)\Gamma_a\left(p\right)$

with the general formula : $\frac{d}{dx}\left(g\left(x\right)h\left(x\right)\right) = R'_{h(x)}\left(g(x)\right)\circ g'\left(x\right)+L'_{g(x)}\left(h(x)\right)\circ$
$h'\left(x\right)$

$R'_{g_b}\left(1\right)\grave{A}_b\left(x\right) = -R'_{g_a}\left(g_{ba}(x)\right)g'_{ba}\left(x\right) + L'_{g_{ba}(x)}\left(g_a\right)R'_{g_a}\left(1\right)\grave{A}_a\left(x\right)$

$\grave{A}_b\left(x\right) = -R'_{g_b^{-1}}\left(g_b\right)R'_{g_a}\left(g_{ba}\right)g'_{ba}\left(x\right) + R'_{g_b^{-1}}\left(g_b\right)L'_{g_{ba}}\left(g_a\right)R'_{g_a}\left(1\right)\grave{A}_a\left(x\right)$

$\grave{A}_b\left(x\right) = -Ad_{g_b}L'_{g_b^{-1}}\left(g_b\right)R'_{g_{ba}g_a}\left(1\right)R'_{g_{ba}^{-1}}\left(g_{ba}\right)g'_{ba}\left(x\right)+R'_{g_b^{-1}}\left(g_b\right)L'_{g_{ba}g_a}\left(1\right)L'_{g_a^{-1}}\left(g_a\right)R'_{g_a}\left(1\right)\grave{A}_a\left(x\right)$

$\grave{A}_b\left(x\right) = -Ad_{g_b}L'_{g_b^{-1}}\left(g_b\right)R'_{g_b}\left(1\right)R'_{g_{ba}^{-1}}\left(g_{ba}\right)g'_{ba}\left(x\right)+R'_{g_b^{-1}}\left(g_b\right)L'_{g_b}\left(1\right)L'_{g_a^{-1}}\left(g_a\right)R'_{g_a}\left(1\right)\grave{A}_a\left(x\right)$

$\grave{A}_b\left(x\right) = -Ad_{g_b}Ad_{g_b^{-1}}R'_{g_{ba}^{-1}}\left(g_{ba}\right)g'_{ba}\left(x\right) + Ad_{g_b}Ad_{g_a^{-1}}\grave{A}_a\left(x\right)$

$\grave{A}_b\left(x\right) = -R'_{g_{ba}^{-1}}\left(g_{ba}\right)g'_{ba}\left(x\right) + Ad_{g_{ba}}\grave{A}_a\left(x\right)$

$\grave{A}_b\left(x\right) = -Ad_{g_{ba}}L'_{g_{ba}^{-1}}\left(g_{ba}\right)g'_{ba}\left(x\right) + Ad_{g_{ba}}\grave{A}_a\left(x\right)$

ii) If there is a family of maps :

$\grave{A}_a \in \Lambda_1\left(O_a; T_1 G\right)$ such that : $\grave{A}_b\left(x\right) = Ad_{g_{ba}}\left(\grave{A}_a\left(x\right) - L'_{g_{ba}^{-1}}(g_{ba})g'_{ba}\left(x\right)\right)$

Define : $\Gamma_a\left(\varphi_a\left(x,g_a\right)\right) = R'_{g_a}\left(1\right)\grave{A}_a\left(x\right)$

$\Gamma_b\left(\varphi_b\left(x,g_b\right)\right) = R'_{g_b}\left(1\right)\grave{A}_b\left(x\right) = R'_{g_b}\left(1\right)Ad_{g_{ba}}\left(\grave{A}_a\left(x\right) - L'_{g_{ba}^{-1}}(g_{ba})g'_{ba}\left(x\right)\right)$

$= R'_{g_b}\left(1\right)Ad_{g_{ba}}R'_{g_a^{-1}}\left(g_a\right)\Gamma_a\left(\varphi_a\left(x,g_a\right)\right)-R'_{g_a}\left(g_{ba}\right)R'_{g_{ba}}\left(1\right))R'_{g_{ba}^{-1}}\left(g_{ba}\right)L'_{g_{ba}}\left(1\right)L'_{g_{ba}^{-1}}(g_{ba})g'_{ba}\left(x\right)$

$= R'_{g_b}\left(1\right)Ad_{g_{ba}}Ad_{g_a}L'_{g_a^{-1}}\left(g_a\right)\Gamma_a\left(p\right) - R'_{g_a}\left(g_{ba}\right)g'_{ba}\left(x\right)$

$= R'_{g_b}\left(1\right)Ad_{g_b}L'_{g_a^{-1}}\left(g_a\right)\Gamma_a\left(p\right) - R'_{g_a}\left(g_{ba}\right)g'_{ba}\left(x\right)$

$= L'_{g_b}\left(1\right)L'_{g_a^{-1}}\left(g_a\right)\Gamma_a\left(p\right)-R'_{g_a}\left(g_{ba}\right)g'_{ba}\left(x\right) = L'_{g_{ba}}\left(g_a\right)\Gamma_a\left(p\right)-R'_{g_a}\left(g_{ba}\right)g'_{ba}\left(x\right)$

$\Gamma_b\left(p\right) = \left(g_{ba}\left(x\right)\left(g_a\right)\right)'\circ\left(-Id_{TM},\Gamma_a\left(p\right)\right)$

So the family defines a connection, and it is principal. ∎

Theorem 2083 *The fundamental vectors are invariant by a principal connection:* $\forall X \in T_1 G : \Phi(p)\left(\zeta\left(X\right)(p)\right) = \zeta\left(X\right)(p)$

Proof. $\Phi(p)\left(\zeta\left(X\right)(p)\right) = \zeta\left(X + Ad_{g^{-1}}\grave{A}_a\left(x\right)0\right)(p) = \zeta\left(X\right)(p)$ ∎

Connexion form

The **connexion form** is the form : $\widehat{\Phi}(p) : TP \rightarrow T_1G :: \Phi(p)(v_p) = \zeta\left(\widehat{\Phi}(p)(v_p)\right)(p)$

$\widehat{\Phi}(\varphi_a(x,g))(\varphi'_{ax}(x,g)v_x + \zeta(u_g)(p)) = u_g + Ad_{g^{-1}}\grave{A}_a(x)v_x$

It has the property for a principal connection : $\rho(p,g)_* \widehat{\Phi}(p) = Ad_{g^{-1}}\widehat{\Phi}(p)$

Proof. $\rho(p,g)_* \Phi(p) = \rho'_p(p,g)\Phi(p)$

$\rho'_p(p,g)\Phi(p) = \zeta\left(Ad_{g^{-1}}\widehat{\Phi}(p)\right)(\rho(p,g))$ ∎

This is a 1-form on TP valued in a fixed vector space. For any X$\in T_1G$ the fundamendal vector field $\zeta(X)(p) \in \mathfrak{X}(VP) \sim \mathfrak{X}(P \times T_1G)$. So it makes sens to compute the Lie derivative, in the usual meaning, of the 1 form $\widehat{\Phi}$ along a fundamental vector field and (Kolar p.100):

$\pounds_{\zeta(X)}\widehat{\Phi} = -ad(X)\left(\widehat{\Phi}\right) = \left[\widehat{\Phi}, X\right]_{T_1G}$

The bundle of principal connections

Theorem 2084 *(Kolar p.159) There is a bijective correspondance between principal connections on a principal fiber bundle P(M,G,π) and the equivariant sections of the first jet prolongation J^1P given by :*

$\Gamma : P \rightarrow J^1P:: \quad \Gamma(p) = \left(\Gamma(p)^i_\alpha\right) \quad such\ that\ \Gamma(\rho(p,g)) = \left(R'_g1\right)\Gamma(p)$

The set of potentials $\left\{\grave{A}(x)^i_\alpha\right\}$ on a principal bundle has the structure of an affine bundle, called the **bundle of principal connections**.

Its structure is defined as follows :

The adjoint bundle of P is the vector bundle E=$P[T_1G, Ad]$ associated to P.

QP=J^1P is an affine bundle over E, modelled on the vector bundle $TM^* \otimes VE \rightarrow E$

$A = A^i_\alpha dx^\alpha \otimes \varepsilon_{ai}(x)$

With :

$QP(x) \times QP(x) \rightarrow TM^* \otimes VE :: (A,B) = B - A$

At the transitions :

$(A_b, B_b) = B_b - A_b = Ad_{g_{ba}}(B_a - A_a)$

Holonomy group

Theorem 2085 *(Kolar p.105) The holonomy group Hol(Φ, p) on a principal fiber bundle $P(M,G,\pi)$ with principal connection Φ is a Lie subgroup of G, and $Hol_0(\Phi, p)$ is a connected Lie subgroup of G and $Hol(\Phi, \rho(p,g)) = Conj_{g^{-1}}Hol(\Phi, p)$*

The Lie algebra hol(Φ,p) of Hol(Φ,p) is a subalgebra of T_1G, linearly generated by the vectors of the curvature form $\widehat{\Omega}(v_p, w_p)$

The set $P_c(p)$ of all curves on P, lifted from curves on M and going through a point p in P, is a principal fiber bundle, with group $Hol(\Phi, p)$, subbundle of P. The pullback of Φ on this bundle is still a principal connection. P is foliated by $P_c(p)$. If the curvature of the connection $\Omega = 0$ then $Hol(\Phi, p) = \{1\}$ and each $P_c(p)$ is a covering of M.

26.3.2 Curvature

For a principal connection the curvature is equivariant : $\Omega(\rho(p, g))\rho'_p(p, g) = Ad_{g^{-1}}\Omega(p)$

As usual with principal fiber bundle it is convenient to relate the value of Ω to the Lie algebra.

Theorem 2086 *The curvature form of a principal connection with potentiel $\grave{A}$ on a principal bundle $P(V, G, \pi)$ is the 2 form $\widehat{\Omega} \in \Lambda_2(P; T_1G)$ such that :* $\Omega = \zeta\left(\widehat{\Omega}\right)$. *It has the following expression in an holonomic basis of TP and basis (ε_i) of T_1G*

$$\widehat{\Omega}(p) = -Ad_{g^{-1}}\sum_i\sum_{\alpha\beta}\left(\partial_\alpha\grave{A}^i_\beta + \left[\grave{A}_a, \grave{A}_\beta\right]^i_{T_1G}\right)dx^\alpha\Lambda dx^\beta\otimes\varepsilon_i \in \Lambda_2(P; T_1G)$$

The bracket $\left[\grave{A}_a, \grave{A}_\beta\right]^i_{T_1G}$ is the Lie bracket in the Lie algebra T_1G.

Proof. The curvature of a princial connection reads :

$\Omega = \sum_i\sum_{\alpha\beta}\left(-\partial_\alpha\Gamma^i_\beta + [\Gamma_\alpha, \Gamma_\beta]^i_{TG}\right)dx^\alpha\Lambda dx^\beta \otimes \partial g_i \in \Lambda_2(P; VP)$

The commutator is taken on TG

$\Gamma(p) = \left(R'_g1\right)\grave{A}_a(x)$ so Γ is a right invariant vector field on TG and :

$[\Gamma_\alpha, \Gamma_\beta]^i_{TG} = \left[\left(R'_g1\right)\grave{A}_\alpha, \left(R'_g1\right)\grave{A}_\beta\right]^i_{TG} = -\left(R'_g1\right)\left[\grave{A}_a, \grave{A}_\beta\right]^i_{T_1G}$

$u_1 \in T_1G : \zeta(u_1)\left(\left(\varphi_a(x, g)\right)\right) = \varphi'_{ag}(x, g)\left(L'_g1\right)u_1$

$u_g \in T_gG : \zeta\left(\left(L'_{g^{-1}}g\right)u_g\right)\left(\left(\varphi_a(x, g)\right)\right) = \varphi'_{ag}(x, g)u_g$

$\Omega(p) = -\sum_{\alpha\beta}dx^\alpha\Lambda dx^\beta\otimes\zeta\left(\left(L'_{g^{-1}}g\right)\left(R'_g1\right)\sum_i\left(\partial_\alpha\grave{A}^i_\beta + \left[\grave{A}_a, \grave{A}_\beta\right]^i_{T_1G}\right)\varepsilon_i\right)(p)$

$\Omega(p) = -\sum_{\alpha\beta}dx^\alpha\Lambda dx^\beta\otimes\zeta\left(Ad_{g^{-1}}\sum_i\left(\partial_\alpha\grave{A}^i_\beta + \left[\grave{A}_a, \grave{A}_\beta\right]^i_{T_1G}\right)\varepsilon_i\right)(p)$ where

(ε_j) is a basis of T_1G ∎

At the transitions we have : $\widehat{\Omega}(\varphi_a(x, 1)) = \widehat{\Omega}(\rho(\varphi_b(x, 1), g_{ba}(x)) = Ad_{g_{ba}^{-1}}\widehat{\Omega}(\varphi_b(x, 1))$

Theorem 2087 *The **strength of a principal connection** with potentiel $\grave{A}$ on a principal bundle $P(V, G, \pi)$ with atlas $(O_a, \varphi_a)_{a\in A}$ is the 2 form $\mathcal{F} \in \Lambda_2(M; T_1G)$ such that : $\mathcal{F}_a = -p_a^*\widehat{\Omega}$ where $p_a(x) = \varphi_a(x, 1)$. It has the following expression in an holonomic basis of TP and basis (ε_i) of T_1G :*

$$\mathcal{F}(x) = \sum_i\left(d\grave{A}^i + \sum_{\alpha\beta}\left[\grave{A}_a, \grave{A}_\beta\right]^i_{T_1G}d\xi^\alpha\Lambda d\xi^\beta\right)\otimes\varepsilon_i$$

At the transitions : $x \in O_a \cap O_b : \mathcal{F}_b(x) = Ad_{g_{ba}} \mathcal{F}_a(x)$

Proof. For any vector fields X,Y on M, their horizontal lifts $\chi_L(X), \chi_L(Y)$ is such that :

$\Omega(p)(\chi_L(X), \chi_L(Y)) = [\chi_L(X), \chi_L(Y)]_{TE} - \chi_L(p)([X,Y]_{TM})$

$\chi_L(\varphi_a(x,g))\left(\sum_\alpha X_x^\alpha \partial \xi^\alpha\right) = \left(\sum_\alpha X_x^\alpha \partial x^\alpha - \zeta\left(Ad_{g^{-1}}\grave{A}(x)X\right)\right)(\varphi_a(x,g))$

If we denote :

$\mathcal{F}(x) = \sum_i \left(d_M \grave{A}^i + \sum_{\alpha\beta}\left[\grave{A}_a, \grave{A}_\beta\right]_{T_1G}^i d\xi^\alpha \Lambda d\xi^\beta\right) \otimes \varepsilon_i \in \Lambda_2(M; T_1 G)$

then : $\widehat{\Omega}(p)(\chi_L(X), \chi_L(Y)) = -Ad_{g^{-1}}\mathcal{F}(x)(X,Y)$

We have with the sections $p_a(x) = \varphi_a(x,1)$:

$\mathcal{F}_a(x)(X,Y) = -\widehat{\Omega}(p_a(x))(p'_{ax}(x,1)X, p'_{ax}(x,1)Y)$

$\mathcal{F}_a = -p_a^*\widehat{\Omega}$ ■

In a change of gauge : $j(x) : \mathcal{F}(x) \to Ad_{j(x)^{-1}}\mathcal{F}(x)$

$\mathcal{F}(\pi(p))(\pi'(p)u_p, \pi'(p)v_p) = \mathcal{F}(x)(u_x, v_x) \Leftrightarrow \mathcal{F} = \pi_*\widehat{\Omega}$

26.3.3 Covariant derivative

Covariant derivative

Theorem 2088 *The covariant derivative of a section S on the principal bundle $P(M,V,\pi)$ with atlas $(O_a, \varphi_a)_{a\in A}$ is expressed by :*

$S(x) = \varphi_a(x, \sigma_a(x)) : \nabla S = \zeta\left(L'_{\sigma_a^{-1}}(\sigma_a)\left(\sigma'_a(x) + R'_{\sigma_a}(1)\grave{A}_a(x)\right)\right)(S(x))$

Proof. The covariant derivative associated to the connection Φ is the map :

$\nabla : \mathfrak{X}(P) \to \Lambda_1(M; VP) :: \nabla S = S^*\Phi$

A section is defined by a family of maps : $\sigma_a : O_a \to G$

$S(x) = \varphi_a(x, \sigma_a(x)) : S'(x) = \varphi'_{ax}(x, \sigma_a(x)) + \varphi'_{ag}(x, \sigma_a(x))\sigma'_a(x) = \varphi'_{ax}(x, \sigma_a(x)) + \zeta\left(L'_{\sigma_a^{-1}}(\sigma_a)\sigma'_a(x)\right)(S(x))$

For a principal connection and a section S on P, and $Y \in \mathfrak{X}(M)$

$\nabla_Y S = S^*\Phi(Y) = \Phi(S(x))(S'(x)Y) = \zeta\left(\left(L'_{\sigma_a^{-1}}(\sigma_a)\sigma'_a(x) + Ad_{\sigma_a^{-1}}\grave{A}_a(x)\right)Y\right)(S(x))$

$\nabla_Y S = \zeta\left(L'_{\sigma_a^{-1}}(\sigma_a)\left(\sigma'_a(x) + R'_{\sigma_a}(1)\grave{A}_a(x)\right)Y\right)(S(x))$ ■

So with the holonomic map :$p_a = \varphi_a(x,1) : \nabla p_a = p_a^*\Phi = \Phi(p_a(x))(p'_a(x)) = \zeta\left(\grave{A}_a(x)\right)(p_a(x))$

At the transitions we can check that ∇S does not depend on the trivialization:

Proof. $x \in O_a \cap O_b : p = \varphi_a(x, g_a) = \varphi_b(x, g_b), g_b = g_{ba}(x)g_a$

$S(x) = \varphi_a(x, \sigma_a(x)) = \varphi_b(x, \sigma_b(x)), \sigma_b(x) = g_{ba}(x)\sigma_a(x)$

$\nabla S = \zeta\left(L'_{\sigma_b^{-1}}(\sigma_b)\left(\sigma'_b(x) + R'_{\sigma_b}(1)\grave{A}_b(x)\right)\right)(S(x))$

$L'_{\sigma_b^{-1}}(\sigma_b)\sigma'_b(x) + Ad_{\sigma_b^{-1}}\grave{A}_b(x) = L'_{\sigma_b^{-1}}(\sigma_b)\left(L'_{g_{ba}}(\sigma_a)\sigma'_a + R'_{\sigma_a}(g_{ba})g'_{ba}\right) + Ad_{\sigma_b^{-1}}Ad_{g_{ba}}\left(\grave{A}_a - L'_{g_{ba}^{-1}}(g_{ba})g'_{ba}\right)$

$$= \left(L_{\sigma_b^{-1}} \circ L_{g_{ba}} \circ L_{\sigma_a}(1)\right)' L'_{\sigma_a^{-1}}(\sigma_a)\,\sigma'_a + Ad_{\sigma_a^{-1}}\grave{A}_a + \left(R'_{\sigma_a}(g_{ba}) - R'_{g_a}(g_a)L'_{\sigma_a^{-1}}(1)L'_{g_{ba}^{-1}}(g_{ba})\right) g'_{ba}$$

$$= L'_{\sigma_a^{-1}}(\sigma_a)\,\sigma'_a(x) + Ad_{\sigma_a^{-1}}\grave{A}_a(x) + \left(R'_{\sigma_a}(g_{ba}) - \left(R_{\sigma_a} \circ L_{\sigma_a^{-1}} \circ L_{g_{ba}^{-1}}(g_{ba})\right)'\right) g'_{ba}$$

∎

Horizontal form

1. The horizontal form of a principal connection reads :

$$\chi \in \wedge_1(P; HP) : \chi(p)\left(\varphi'_{ax}(x,g)v_x + \zeta(u_g)(p)\right) = \varphi'_{ax}(x,g)v_x - \zeta\left(Ad_{g^{-1}}\grave{A}_a(x)\,v_x\right)(p)$$

2. Horizontal lift of a vector field on M

Theorem 2089 *The horizontal lift of a vector field on M by a principal connection with potential $\grave{A}$ on a principal bundle $P(M,G,\pi)$ with trivialization φ is the map : $\chi_L : \mathfrak{X}(TM) \to \mathfrak{X}(HP) :: \chi_L(p)(X) = \varphi'_x(x,g)X(x) - \zeta\left(Ad_{g^{-1}}\grave{A}_a(x)X(x)\right)(p)$*

Proof. From the general definition: $\chi_L(p)(X) = \varphi'(x,g)(X(x), -\Gamma(p)X(x))$

$$= \varphi'(x,g)\left(X(x), -R'_g(1)\grave{A}_a(x)X(x)\right) = \varphi'_x(x,g)X(x) - \varphi'_g(x,g)L'_g(1)L'_{g^{-1}}(g)R'_g(1)\grave{A}_a(x)X(x)$$

$$= \varphi'_x(x,g)X(x) - \zeta\left(Ad_{g^{-1}}\grave{A}_a(x)X(x)\right)(p) \quad ∎$$

$\chi_L(p)(X)$ is a horizontal vector field on TE, which is projectable on X.

For any section S on P we have : $S \in \mathfrak{X}(E): \nabla_X S = \pounds_{\chi_L(X)}S$

3. Horizontalization of a r-form on P:

Definition 2090 *The **horizontalization** of a r-form on a principal bundle $P(M,G,\pi)$ endowed with a principal connection is the map :*

$$\chi^* : \mathfrak{X}(\Lambda_r TP) \to \mathfrak{X}(\Lambda_r TP) :: \chi^*\varpi(p)(v_1,...,v_r) = \varpi(p)(\chi(p)(v_1),...,\chi(p)(v_r))$$

$\chi^*\varpi$ is a horizontal form : it is null whenever one of the vector v_k is vertical.

It reads in the holonomic basis of P:

$$\chi^*\varpi(p) = \sum_{\{\alpha_1..\alpha_r\}} \mu_{\alpha_1..\alpha_r} dx^{\alpha_1} \wedge ... \wedge dx^{\alpha_r}$$

Theorem 2091 *(Kolar p.103) The horizontalization of a r-form has the following properties :*

$$\chi^* \circ \chi^* = \chi^*$$

$$\forall \mu \in \mathfrak{X}(\Lambda_r TP), \varpi \in \mathfrak{X}(\Lambda_s TP) : \chi^*(\mu \wedge \varpi) = \chi^*(\mu) \wedge \chi^*(\varpi),$$

$$\chi^*\widehat{\Phi} = 0$$

$$X \in T_1 G : \chi^*\pounds\zeta(X) = \pounds\zeta(X) \circ \chi^*$$

Exterior covariant derivative

The exterior covariant derivative on a vector bundle E is a map: $\nabla_e : \Lambda_r(M;E) \to \Lambda_{r+1}(M;E)$. The exterior covariant derivative on a principal bundle is a map : $\nabla_e : \Lambda_r(P;V) \to \Lambda_{r+1}(P;V)$ where V is any fixed vector space.

Definition 2092 *(Kolar p.103) The **exterior covariant derivative** on a principal bundle $P(M, G, \pi)$ endowed with a principal connection is the map : $\nabla_e : \Lambda_r(P; V) \to \Lambda_{r+1}(P; V) :: \nabla_e \varpi = \chi^*(d\varpi)$ where V is any fixed vector space.*

Theorem 2093 *(Kolar p.103) The exterior covariant derivative ∇_e on a principal bundle $P(M, G, \pi)$ endowed with a principal connection has the following properties :*

    i) $\nabla_e \widehat{\Phi} = \widehat{\Omega}$

    ii) Bianchi identity : $\nabla_e \widehat{\Omega} = 0$

    iii) $\forall \mu \in \Lambda_r(P; \mathbb{R}), \varpi \in \Lambda_s(P; V) : \nabla_e(\mu \wedge \varpi) = (\nabla_e \mu) \wedge \chi^* \varpi + (-1)^r (\chi^* \mu) \wedge \nabla_e \varpi$

    iv) $X \in T_1 G : \pounds \zeta(X) \circ \nabla_e = \nabla_e \circ \pounds \zeta(X)$

    v) $\forall g \in G : \rho(.,g)^* \nabla_e = \nabla_e \rho(.,g)^*$

    vi) $\nabla_e \circ \pi^* = d \circ \pi^* = \pi^* \circ d$

    vii) $\nabla_e \circ \chi^* - \nabla_e = \chi^* i_\Omega$

    viii) $\nabla_e \circ \nabla_e = \chi^* \circ i_\Omega \circ d$

Theorem 2094 *The exterior covariant derivative ∇_e of a r-forms ϖ on P, horizontal, equivariant and valued in the Lie algebra is:*

    $\nabla_e \varpi = d\varpi + \left[\widehat{\Phi}, \varpi \right]_{T_1 G}$ *with the connexion form $\widehat{\Phi}$ of the connection*

Such a form reads :

$\varpi = \sum_{\{\alpha_1 \ldots \alpha_r\}} \sum_i \varpi^i_{\alpha_1 \ldots \alpha_r}(p) \, dx^{\alpha_1} \wedge .. \wedge dx^r \otimes \varepsilon_i$ with $\sum_i \varpi^i_{\alpha_1 \ldots \alpha_r}(\rho(p,g)) \varepsilon_i = Ad_{g^{-1}} \sum_i \varpi^i_{\alpha_1 \ldots \alpha_r}(p) \varepsilon_i$

the bracket is the bracket in the Lie algebra:

$\left[\widehat{\Phi}, \varpi \right]^i_{T_1 G} = \left[\sum_{j\beta} \widehat{\Phi}^j_\beta dx^\beta \otimes \varepsilon_j, \sum_k \varpi^k_{\alpha_1 \ldots \alpha_r} dx^{\alpha_1} \wedge .. \wedge dx^r \otimes \varepsilon_k \right]^i_{T_1 G}$

$= \sum_{\beta\{\alpha_1 \ldots \alpha_r\}} \left[\sum_j \widehat{\Phi}^j_\beta \varepsilon_j, \sum_k \varpi^k_{\alpha_1 \ldots \alpha_r} \varepsilon_k \right]^i dx^\beta \wedge dx^{\alpha_1} \wedge .. \wedge dx^r \otimes \varepsilon_i$

The algebra of r-forms on P, horizontal, equivariant and valued in the Lie algebra on one hand, and the algebra of r-forms on P, horizontal, equivariant and valued in VP on the other hand, are isomorphic :

$\widehat{\varpi} = \sum_{\{\alpha_1 \ldots \alpha_r\}} \sum_i \varpi^i_{\alpha_1 \ldots \alpha_r}(p) \, dx^{\alpha_1} \wedge .. \wedge dx^r \otimes \varepsilon_i \longleftrightarrow \varpi = \sum_{\{\alpha_1 \ldots \alpha_r\}} \sum_i \varpi^i_{\alpha_1 \ldots \alpha_r}(p) \, dx^{\alpha_1} \wedge .. \wedge dx^r \otimes \zeta(\varepsilon_i)(p)$

So : $\zeta(\nabla_e \widehat{\varpi}) = \left[\Phi, \zeta\left(\widehat{\Phi}\right) \right]$

26.4 Connection on associated bundles

We now combine all the previous elements.

26.4.1 Connection on general associated bundles

Connection

Theorem 2095 *A principal connexion Φ with potentiels $\left(\grave{A}_a\right)_{a\in A}$ on a principal bundle $P\left(M,G,\pi\right)$ with atlas $\left(O_a,\varphi_a\right)_{a\in A}$ induces on any associated bundle $E=P\left[V,\lambda\right]$ a connexion defined by the set of Christoffel forms : $\Gamma_a\left(\varphi_a\left(x,1\right),u_a\right)=\lambda'_g\left(1,u_a\right)\grave{A}_a\left(x\right)$ with the transition rule : $\Gamma_b\left(q\right)=\lambda'\left(g_{ba},u_a\right)\left(-Id,\Gamma_a\left(p\right)\right).$ and $\Psi\left(\psi_a\left(x,u_a\right)\right)\left(v_{ax},v_{au}\right)\sim\left(v_{au}+\Gamma_a\left(q\right)v_{ax}\right)\in T_uV$*

Proof. i) $P(M,G,\pi)$ is a principal bundle with atlas $\left(O_a,\varphi_a\right)_{a\in A}$ and transition maps :$g_{ba}\left(x\right)\in G$

E is a fiber bundle $E\left(M,V,\pi_E\right)$ with atlas $\left(O_a,\psi_a\right)_{a\in A}$:

trivializations $\psi_a:O_a\times V\to E::\psi_a\left(x,u\right)=\mathrm{Pr}\left(\left(\varphi_a(x,1),u\right)\right)$

transitions maps : $\psi_a\left(x,u_a\right)=\psi_b\left(x,u_b\right)\Rightarrow u_b=\lambda\left(g_{ba}\left(x\right),u_a\right)$

Define the maps : $\Gamma_a\in\Lambda_1\left(\pi_E^{-1}\left(O_a\right);TM^*\otimes TV\right)::\Gamma_a\left(\psi_a\left(x,u_a\right)\right)=\lambda'_g\left(1,u_a\right)\grave{A}_a\left(x\right)$

$$T_xM\to\overset{\grave{A}_a(x)}{\to}\to T_1G\to\overset{\lambda'_g\left(1,u_a\right)}{\to}\to TV$$

At the transitions : $x\in O_a\cap O_b:q=\psi_a\left(x,u_{a()}\right)=\psi_b\left(x,u_b\right)$

$\grave{A}_b\left(x\right)=Ad_{g_{ba}}\left(\grave{A}_a\left(x\right)-L'_{g_{ba}^{-1}}(g_{ba})g'_{ba}\left(x\right)\right)$

$u_b=\lambda\left(g_{ba}\left(x\right),u_a\right)$

$\Gamma_b\left(q\right)=\lambda'_g(1,u_b)\grave{A}_b\left(x\right)=\lambda'_g(1,\lambda\left(g_{ba}\left(x\right),u_a\right))Ad_{g_{ba}}\left(\grave{A}_a\left(x\right)-L'_{g_{ba}^{-1}}(g_{ba})g'_{ba}\left(x\right)\right)$

$=\lambda'_g(g_{ba},u_a)R'_{g_{ba}}\left(1\right)Ad_{g_{ba}}\grave{A}_a-\lambda'_g(g_{ba},u_a)R'_{g_{ba}}\left(1\right)R'_{g_{ba}^{-1}}(g_{ba})L'_{g_{ba}}(1)L'_{g_{ba}^{-1}}(g_{ba})g'_{ba}\left(x\right)$

$=\lambda'_u(g_{ba},u_a)\lambda'_g(1,u_a)L'_{g_{ba}^{-1}}(g_{ba})R'_{g_{ba}}\left(1\right)Ad_{g_{ba}}\grave{A}_a-\lambda'_g(g_{ba},u_a)g'_{ba}\left(x\right)$

$=\lambda'_u(g_{ba},u_a)\lambda'_g(1,u_a)\grave{A}_a-\lambda'_g(g_{ba},u_a)g'_{ba}\left(x\right)$

$=\lambda'_u(g_{ba},u_a)\Gamma_a\left(p\right)-\lambda'_g(g_{ba},u_a)g'_{ba}\left(x\right)$

$=\lambda'\left(g_{ba},u_a\right)\left(-Id,\Gamma_a\left(p\right)\right)$

So the maps Γ_a define a connection $\Psi\left(q\right)$ on E.

ii) A vector of TE reads : $v_q=\psi'_{ax}\left(x,u_a\right)v_x+\psi'_{au}\left(x,u_a\right)v_{au}$ and :

$\Psi(q)v_q=\psi'_a(x,u_a)\left(0,v_{au}+\Gamma_a\left(q\right)v_{ax}\right)=\psi'_{au}(x,u_a)v_{au}+\psi'_{au}(x,u_a)\lambda'_g(1,u_a)\grave{A}_a\left(x\right)v_{ax}$

The vertical bundle is $\left\{\left(p_a\left(x\right),0\right)\times\left(u,v_u\right)\right\}\simeq P\times_G\left(TV\right)$

so : $\Psi(\psi_a\left(x,u_a\right))v_q\sim\left(v_{au}+\Gamma_a\left(q\right)v_{ax}\right)\in T_uV$ ∎

It can be shown (Kolar p.107) that $\Psi\left(p,u\right)\left(v_p,v_u\right)=\mathrm{Pr}\left(\Phi\left(p\right)v_p,v_u\right)$

$\Psi(\psi_a\left(x,u_a\right))v_q$ is a vertical vector which is equivalent to the fundamental

vector of E : $\frac{1}{2}Z\left(\lambda'_g\left(1,u_a\right)^{-1}v_{au}+\grave{A}_av_{ax}\right)\left(p_a,u_a\right)$

Conversely :

Theorem 2096 *(Kolar p.108) A connection Ψ on the associated bundle $P\left[V,\lambda\right]$ such that the map : $\zeta:T_1G\times P\to TP$ is injective is induced by a principal connection on P iff the Christoffel forms $\Gamma\left(q\right)$ of Ψ are valued in $\lambda'_g(1,u)T_1G$*

The horizontal form reads :

$\chi\in\wedge_1\left(TE;HE\right):\chi\left(p_a\left(x\right),u\right)\left(v_x,v_u\right)=\left(p_a\left(x\right),u\right)\left(v_x,-\lambda'_g(1,u)\grave{A}\left(x\right)v_x\right)$

Curvature

Theorem 2097 *The curvature $\Omega \in \wedge_2 (E; VE)$ of a connection Ψ induced on the associated bundle $E=P[V,\lambda]$ with atlas $(O_a, \psi_a)_{a \in \in A}$ by the connection Φ on the principal bundle P is :*

$$q=\psi_a (x, u), u_x, v_x \in T_x M : \Omega (q) (u_x, v_x) = \psi'_a (p_a (x), u) \left(0, \widehat{\Omega} (q) (u_x, v_x)\right)$$

with :

$\widehat{\Omega} (q) = -\lambda'_g(1, u)\mathcal{F}$ *where $\mathcal{F}$ is the strength of the connection Φ. In an holonomic basis of TM and basis (ε_i) of $T_1 G$:*

$$\widehat{\Omega} (q) = -\sum_i \left(d\grave{A}^i_a + \sum_{\alpha\beta} \left[\grave{A}_{aa}, \grave{A}_{a\beta}\right]^i_{T_1 G} d\xi^\alpha \Lambda d\xi^\beta\right) \otimes \lambda'_g(1, u) (\varepsilon_i)$$

Proof. The curvature of the induced connection is given by the Cartan formula

$$\psi^*_a \Omega = \sum_i \left(-d_M \Gamma^i + [\Gamma, \Gamma]^i_V\right) \otimes \partial u_i$$

$$\sum_i d_M \Gamma^i \partial u_i = \lambda'_g(1, u) d_M \grave{A}$$

$$\sum_i [\Gamma, \Gamma]^i_V \partial u_i = \sum_i \left[\lambda'_g(1, u)\grave{A}, \lambda'_g(1, u)\grave{A}\right]^i_V \partial u_i = -\lambda'_g(1, u) \left[\grave{A}, \grave{A}\right]_{T_1 G}$$

$$\psi^*_a \Omega = -\lambda'_g(1, u) \sum_i \left(d_M \grave{A} + \left[\grave{A}, \grave{A}\right]\right) = -\lambda'_g(1, u)\mathcal{F}$$

$$\widehat{\Omega} (q) = -\sum_i \left(d\grave{A}^i + \sum_{\alpha\beta} \left[\grave{A}_a, \grave{A}_\beta\right]^i_{T_1 G} d\xi^\alpha \Lambda d\xi^\beta\right) \otimes \lambda'_g(1, u) (\varepsilon_i) \quad \blacksquare$$

Covariant derivative

The covariant derivative $\nabla_X S$ of a section S on the associated bundle $P[V, \lambda]$ along a vector field X on M, induced by the connection of potential $\grave{A}$ on the principal bundle $P(M, G, \pi)$ is :

$$\nabla_X S = \Psi (S (x)) S' (x) X = \left(\lambda'_g(1, s (x))\grave{A} (x) X + s' (x) X\right) \in T_{s(x)}V$$

with $S (x) = (p_a (x), s (x))$

26.4.2 Connection on an associated vector bundle

Induced linear connection

Theorem 2098 *A principal connexion Φ with potentiels $\left(\grave{A}_a\right)_{a \in A}$ on a principal bundle $P(M, G, \pi)$ with atlas $(O_a, \varphi_a)_{a \in A}$ induces on any associated vector bundle $E=P[V, r]$ a linear connexion Ψ defined by the Christoffel forms : $\Gamma_a = r'_g(1)\grave{A}_a \in \Lambda_1 (O_a; \mathcal{L} (V; V))$ with the transition rule : $\Gamma_b (x) = Ad_{r(g_{ba})}\Gamma_a (x) - r'_g(1)R'_{g_{ba}^{-1}}(g_{ba})g'_{ba} (x)$. and in a holonomic basis $\left((e_i (x))_{i \in I}, (\partial x_\alpha)\right)$ of TE :*

$$\Psi \left(\sum_i u^i e_i (x)\right) \left(\sum_i v^i_u e_i (x) + \sum_\alpha v^\alpha_x \partial x_\alpha\right) = \sum_i \left(v^i_u + \sum_{ij\alpha} \Gamma^i_{j\alpha} (x) u^j v^\alpha_x\right) e_i (x)$$

In matrix notation :

Proof. i) $\widehat{\Gamma}_a\left(\varphi_a\left(x,1\right),u_a\right) = \lambda'_g(1,u_a)\dot{A}_a\left(x\right)$ reads : $\widehat{\Gamma}_a\left(\varphi_a\left(x,1\right),u_a\right) = r'_g(1)u_a\dot{A}_a\left(x\right)$

It is linear with respect to u. $\forall L \in \mathcal{L}\left(V;V\right) : \widehat{\Gamma}_a\left(\varphi_a\left(x,1\right),L\left(u_a\right)\right) = L\left(\widehat{\Gamma}_a\left(\varphi_a\left(x,1\right),u_a\right)\right)$

Denote : $\Gamma_a\left(x\right)u_a = \widehat{\Gamma}_a\left(\varphi_a\left(x,1\right),u_a\right)$
$\Gamma_a\left(x\right) = r'_g(1)\dot{A}_a\left(x\right) \in \Lambda_1\left(M;\mathcal{L}\left(V;V\right)\right)$

$$T_xM \overset{\dot{A}_a(x)}{\to} \to T_1G \overset{r'_g(1)}{\to} \to \mathcal{L}\left(V;V\right)$$

ii) For $x \in O_a \cap O_b : q = \left(\varphi_a\left(x,1\right),u_a\right) = \left(\varphi_b\left(x,1\right),u_b\right)$
$\Gamma_b\left(x\right) = r'_g(1)\dot{A}_b\left(x\right) = r'_g(1)Ad_{g_{ba}}\left(\dot{A}_a\left(x\right) - L'_{g_{ba}^{-1}}(g_{ba})g'_{ba}\left(x\right)\right)$
$Ad_{r(g)}r'\left(1\right) = r'\left(1\right)Ad_g$
$\Gamma_b\left(x\right) = Ad_{r(g_{ba})}r'\left(1\right)\dot{A}_a\left(x\right) - r'_g(1)Ad_{g_{ba}}L'_{g_{ba}^{-1}}(g_{ba})g'_{ba}\left(x\right)$
$\Gamma_b\left(x\right) = Ad_{r(g_{ba})}\Gamma_b\left(x\right) - r'_g(1)R'_{g_{ba}^{-1}}(g_{ba})L'_{g_{ba}}(1)L'_{g_{ba}^{-1}}(g_{ba})g'_{ba}\left(x\right)$
$\Gamma_b\left(x\right) = Ad_{r(g_{ba})}\Gamma_b\left(x\right) - r'_g(1)R'_{g_{ba}^{-1}}(g_{ba})g'_{ba}\left(x\right)$
iii) In a holonomic basis of E : $e_{ai}\left(x\right) = \left[\left(p_a\left(x\right),e_i\right)\right] : v_q \in T_qE :: v_q = \sum_i v_u^i e_i\left(x\right) + \sum_\alpha v_x^\alpha \partial x_\alpha$
The vertical bundle is $\left\{\left(p_a\left(x\right),0\right) \times \left(u,v_u\right)\right\} \simeq P \times_G \left(V \times V\right)$
$\Psi\left(p_a\left(x\right),u\right)\left(v_q\right) = \left(p_a\left(x\right),0\right) \times \left(u,r'\left(1\right)u\dot{A}\left(x\right)v_x + v_u\right)$
$\left(p_a\left(x\right),u\right) \in E$ reads $\sum_i u^i e_i\left(x\right)$
$\left(\sum_\alpha v_x^\alpha \partial x^\alpha\right)\left(p_a\left(x\right)\right)$ reads : $\sum_\alpha v_x^\alpha \partial x^\alpha$
$\left(u,r'\left(1\right)u\dot{A}\left(x\right)v_x\right)$ reads : $\sum_{ij\alpha}\Gamma_{j\alpha}^i\left(x\right)u^j v_x^\alpha e_i\left(x\right)$
So we can write :
$\Psi\left(\sum_i u^i e_i\left(x\right)\right)\left(\sum_i v_u^i e_i\left(x\right) + \sum_\alpha v_x^\alpha \partial x_\alpha\right) = \sum_i\left(v_u^i + \sum_{ij\alpha}\Gamma_{j\alpha}^i\left(x\right)u^j v_x^\alpha\right)e_i\left(x\right)$
∎

We have a linear connection on E, which has all the properties of a common linear connection on vector bundles, by using the holonomic basis : $\left(p_a\left(x\right),e_i\right) = e_{ai}\left(x\right)$.

Remark : there is always a complexified $E_{\mathbb{C}} = P\left[V_{\mathbb{C}},r_{\mathbb{C}}\right]$ of a real associated vector bundle with $V_{\mathbb{C}} = V \oplus iV$ and $r_{\mathbb{C}} : G \to V_{\mathbb{C}} :: r_{\mathbb{C}}\left(g\right)\left(u+iv\right) = r\left(g\right)u + ir\left(g\right)v$ so $\forall X \in T_1G : r'_{\mathbb{C}}\left(1\right)X \in GL\left(V_{\mathbb{C}};V_{\mathbb{C}}\right)$. A real connection on P has an extension on $E_{\mathbb{C}}$ by : $\Gamma_{\mathbb{C}} = r'_{\mathbb{C}}(1)\dot{A}_a \in GL\left(V_{\mathbb{C}};V_{\mathbb{C}}\right)$. However P, G and $\dot{A}$ stay the same. To define an extension to $\left(T_1G\right)_{\mathbb{C}}$ one needs an additional map : $\dot{A}_{\mathbb{C}a} = \dot{A}_a + i\,\text{Im}\,\dot{A}_a$

Curvature

Theorem 2099 *The curvature Ω of a connection Ψ induced on the associated vector bundle $P\left[V,r\right]$ by the connection Φ with potential $\dot{A}$ on the principal bundle P is :* $q=\psi_a\left(x,u\right),u_x,v_x \in T_xM : \Omega\left(q\right)\left(u_x,v_x\right) = \psi'_a\left(q\right)\left(0,\widehat{\Omega}\left(q\right)\left(u_x,v_x\right)\right)$

with : $\widehat{\Omega}\left(\psi_a\left(x,u\right)\right)=-r'\left(1\right)u\mathcal{F}$ where $\mathcal{F}$ is the strength of the connection Φ. In an holonomic basis of TM and basis (ε_i) of T_1G :

$$\widehat{\Omega}\left(q\right)=-\sum_i\left(d\dot{A}_a^i+\sum_{\alpha\beta}\left[\dot{A}_{aa},\dot{A}_{a\beta}\right]_{T_1G}^i d\xi^\alpha\wedge d\xi^\beta\right)\otimes r'_g\left(1\right)u\left(\varepsilon_i\right)$$

At the transitions : $x\in O_a\cap O_b:\widehat{\Omega}\left(p_b\left(x\right),e_i\right)=Ad_{r\left(g_{ba}\right)}\widehat{\Omega}\left(p_a\left(x\right),e_i\right)$

Proof. $e_{bi}\left(x\right)=\left(p_b\left(x\right),e_i\right)=r\left(g_{ba}\left(x\right)\right)e_{ai}\left(x\right)=r\left(g_{ba}\left(x\right)\right)\left(p_a\left(x\right),e_i\right)$

$\mathcal{F}_b=Ad_{g_{ba}}\mathcal{F}_a$

With the general identity : $Ad_{r\left(g_{ba}\right)}r'\left(1\right)=r'\left(1\right)Ad_{g_{ba}}$ and the adjoint map on $G\mathcal{L}(V;V)$

$\Omega\left(p_b\left(x\right),e_i\right)=-r'\left(1\right)e_i\mathcal{F}_b=-r'\left(1\right)e_iAd_{g_{ba}}\mathcal{F}_a=-Ad_{r\left(g_{ba}\right)}r'\left(1\right)e_i\mathcal{F}_a=Ad_{r\left(g_{ba}\right)}\Omega\left(p_a\left(x\right),e_i\right)$ ∎

The curvature is linear with respect to u.

The curvature form $\widetilde{\Omega}$ of Ψ : $\widetilde{\Omega}_i\left(x\right)=\widehat{\Omega}\left(p_a\left(x\right),e_i\right)$

Exterior covariant derivative

1. There are two different exterior covariant derivative :

- for r-forms on M valued on a vector bundle $E\left(M,V,\pi\right)$:

$$\widetilde{\nabla}_e:\Lambda_r\left(M;E\right)\rightarrow\Lambda_{r+1}\left(M;E\right)::\widetilde{\nabla}_e\varpi=\sum_i\left(d_M\varpi^i+\left(\sum_j\left(\sum_\alpha\widetilde{\Gamma}_{j\alpha}^i d\xi^\alpha\right)\wedge\varpi^j\right)\right)\otimes e_i\left(x\right)$$

with a linear connection. It is still valid on an associated vector bundle with the induced connection.

- for r-forms on a principal bundle $P(M,G,\pi)$, valued in a fixed vector space V:

$$\nabla_e:\Lambda_r\left(P;V\right)\rightarrow\Lambda_{r+1}\left(P;V\right)::\nabla_e\varpi=\chi^*\left(d_P\varpi\right)$$

with a principal connection

There is a relation between those two concepts .

Theorem 2100 *(Kolar p.112) There is a canonical isomorphism between the space $\Lambda_s\left(M;P\left[V,r\right]\right)$ of s-forms on M valued in the associated vector bundle $P\left[V,r\right]$ and the space $W_s\subset\Lambda_s\left(P;V\right)$ of horizontal, equivariant V valued s-forms on P.*

A s-form on M valued in the associated vector bundle $E=P\left[V,r\right]$ reads :

$\widetilde{\varpi}=\sum_{\{\alpha_1...\alpha_s\}}\sum_i\widetilde{\varpi}_{\alpha_1...\alpha_s}^i\left(x\right)d\xi^{\alpha_1}\wedge..\wedge d\xi^{\alpha_s}\otimes e_i\left(x\right)$

with a holonomic basis $(d\xi^\alpha)$ of TM*, and $e_i\left(x\right)=\psi_a\left(x,e_i\right)=\left(p_a\left(x\right),e_i\right)$

A s-form on P, horizontal, equivariant, V valued reads :

$\varpi=\sum_{\{\alpha_1...\alpha_s\}}\sum_i\varpi_{\alpha_1...\alpha_s}^i\left(p\right)dx^{\alpha_1}\wedge..\wedge dx^s\otimes e_i$ and : $\rho\left(.,g\right)^*\varpi=r\left(g^{-1}\right)\varpi$

with the dual basis of $\partial x_\alpha=\varphi'_a\left(x,g\right)\partial\xi_\alpha$ and a basis e_i of V

The equivariance means :

$\rho\left(.,g\right)^*\varpi\left(p\right)\left(v_1,..v_s\right)=\varpi\left(\rho\left(p,g\right)\right)\left(\rho'_p\left(p,g\right)v_1,...\rho'_p\left(p,g\right)v_s\right)=r\left(g^{-1}\right)\varpi\left(p\right)\left(v_1,..v_s\right)$

$\Leftrightarrow\sum_i\varpi_{\alpha_1...\alpha_s}^i\left(\rho\left(p,g\right)\right)e_i=r\left(g^{-1}\right)\left(\sum_i\varpi_{\alpha_1...\alpha_s}^i\left(p\right)e_i\right)$

So : $\sum_i\varpi_{\alpha_1...\alpha_s}^i\left(\varphi\left(x,g\right)\right)e_i=r\left(g^{-1}\right)\left(\sum_i\varpi_{\alpha_1...\alpha_s}^i\left(\varphi\left(x,1\right)\right)e_i\right)$

The isomorphism : $\theta:\Lambda_s\left(M;E\right)\rightarrow W_s$ reads :

$v_k\in T_pP:X_k=\pi'\left(p\right)v_k$

$\widetilde{\varpi}\left(x\right)\left(X_{1},...,X_{s}\right)=\left(p_{a}\left(x\right),\varpi\left(p_{a}\left(x\right)\right)\left(v_{1},...,v_{s}\right)\right)$

$\left(p_{a}\left(x\right),\varpi\left(\varphi_{a}\left(x,1\right)\right)\left(v_{1},...,v_{s}\right)\right)=\widetilde{\varpi}\left(x\right)\left(\pi'\left(p\right)v_{1},...,\pi'\left(p\right)v_{s}\right)$

$\theta\left(\widetilde{\varpi}\right)$ is called the **frame form** of $\widetilde{\varpi}$

Then : $\theta\circ\widetilde{\nabla}_{e}=\nabla_{e}\circ\theta:\Lambda_{r}\left(M;E\right)\rightarrow W_{r+1}$ where the connection on E is induced by the connection on P.

We have similar relations between the curvatures :

The Riemann curvature on the vector bundle $E=P\left[V,r\right]:R\in\Lambda_{2}\left(M;E^{*}\otimes E\right):$

$\widetilde{\nabla}_{e}\left(\widetilde{\nabla}_{e}\widetilde{\varpi}\right)=\sum_{ij}\left(\sum_{\alpha\beta}R^{i}_{j\alpha\beta}d\xi^{\alpha}\wedge d\xi^{\beta}\right)\wedge\widetilde{\varpi}^{j}\otimes e_{i}\left(x\right)$

The curvature form on the principal bundle : $\widehat{\Omega}\in\Lambda_{2}\left(P;T_{1}G\right):\widehat{\Omega}\left(p\right)=$

$-Ad_{g^{-1}}\sum_{i}\left(d_{M}\grave{A}^{i}+\left[\sum_{\alpha}\grave{A}_{a},\sum_{\beta}\grave{A}_{\beta}\right]^{i}_{T_{1}G}\right)\varepsilon_{i}$

We have :

$\theta\circ R=r'\left(1\right)\circ\widehat{\Omega}=\theta\circ\widetilde{\nabla}_{e}\widetilde{\nabla}_{e}=\nabla_{e}\nabla_{e}\circ\theta$

26.4.3 Metric connections

A connection on a linear vector bundle (where it is possible to define tensors) is a connection such that the covariant derivative of the metric is null.

General result

First a general result, which has a broad range of practical applications.

Theorem 2101 *If (V,γ) is a real vector space endowed with a scalar product γ, (V,r) an unitary representation of the group G, $P\left(M,G,\pi\right)$ a real principal fiber bundle endowed with a principal connection with potential $\grave{A}$, then $E=P[V,r]$ is a vector bundle endowed with a scalar product and a linear metric connection.*

Proof. i) γ is a symmetric form, represented in a base of V by the matrix $[\gamma]=[\gamma]^{t}$

ii) (V,r) is a unitary representation : $\forall u,v\in V,g\in G:\gamma\left(r\left(g\right)u,r\left(g\right)v\right)=\gamma\left(u,v\right)$ so (V,r'(1)) is a anti-symmetric representation of the Lie algebra $T_{1}G$: $\forall u,v\in V,\forall X\in T_{1}G:\gamma\left(r'(1)Xu,r'(1)Xv\right)=0\Leftrightarrow\left([\gamma]\left[r'_{g}(1)\right][X]\right)^{t}+[\gamma]\left[r'_{g}(1)\right][X]=0$

iii) (V,r) is representation of G so E=P[V,r] is a vector bundle.

iv) γ is preserved by r, so $g\left(\left(p_{a}\left(x\right),u\right),\left(p_{a}\left(x\right),v\right)\right)=\gamma\left(u,v\right)$ is a scalar product on E

v) the principal connection on P is real so its potential $\grave{A}$ is real

vi) It induces the linear connection on E with Christoffel form : $\Gamma\left(x\right)=r'_{g}(1)\grave{A}$

vii) the connection is metric because :

$\left([\gamma]\left[r'_{g}(1)\right]\left[\grave{A}_{\alpha}\right]\right)^{*}+[\gamma]\left[r'_{g}(1)\right]\left[\grave{A}_{\alpha}\right]=0$

$\Rightarrow\left[\grave{A}_{\alpha}\right]^{t}\left[r'_{g}(1)\right]^{*}[\gamma]+[\gamma]\left[r'_{g}(1)\right]\left[\grave{A}_{\alpha}\right]=0$

$\Leftrightarrow\left[\Gamma_{\alpha}\right]^{t}[\gamma]+[\gamma]\left[\Gamma_{\alpha}\right]=0$ ∎

Remark : the theorem still holds for the complex case, but we need a real structure on V with respect to which the transition maps are real, and a real connection.

Metric connections on the tangent bundle of a manifold

The following theorem is new. It makes a link between usual "affine connections" (Part Differential geometry) on manifolds and connections in the fiber bundle meaning. The demonstration gives many useful details about the definition of both the bundle of orthonormal frames and of the induced connection, in practical terms.

Theorem 2102 *For any real m dimensional pseudo riemannian manifold (M,g) with the signature (r,s) of g, any principal connection on the principal bundle $P(M, O(\mathbb{R}, r, s), \pi)$ of its orthogonal frames induces a metric, affine connection on (M,g). Conversely any affine connection on (M,g) induces a principal connection on P.*

Proof. i) P is a principal bundle with atlas $(O_a, \varphi_a)_{a \in A}$. The trivialization φ_a is defined by a family of maps : $L_a \in C_0(O_a; GL(\mathbb{R}, m))$ such that : $\varphi_a(x, 1) = e_{ai}(x) = \sum_\alpha [L_a]_i^\alpha \partial \xi_\alpha$ is an orthonormal basis of (M,g) and $\varphi_a(x, S) = \rho(e_{ai}(x), g) = \sum_\alpha [L_a]_j^\alpha [S]_i^j \partial \xi_\alpha$ with $[S] \in O(\mathbb{R}, r, s)$. So in a holonomic basis of TM we have the matrices relation: $[L_a(x)]^t [g(x)] [L_a(x)] = [\eta]$ with $\eta_{ij} = \pm 1$. At the transitions of P : $e_{bi}(x) = \varphi_b(x, 1) = \varphi_a(x, g_{ab}) = \sum_\alpha [L_a]_j^\alpha [g_{ab}]_i^j \partial \xi_\alpha = \sum_\alpha [L_b]_i^\alpha \partial \xi_\alpha$ so : $[L_b] = [L_a][g_{ab}]$ with $[g_{ab}] \in O(\mathbb{R}, r, s)$.

E is the associated vector bundle $P[\mathbb{R}^m, \imath]$ where $(\mathbb{R}^m, \imath)$ is the standard representation of $O(\mathbb{R}, r, s)$. A vector of E is a vector of TM defined in the orthonormal basis $e_{ai}(x)$:$U_p = \sum_i U_a^i e_{ai}(x) = \sum_{i\alpha} U_a^i [L_a]_i^\alpha \partial \xi_\alpha = \sum_\alpha u^\alpha \partial \xi_\alpha$ with $U_a^i = \sum_\alpha [L'_a]_\alpha^i u^\alpha$ and $[L'_a] = [L_a]^{-1}$.

Any principal connexion on P induces the linear connexion on E :
$\Gamma(x) = \sum_{\alpha i j} [\Gamma_{a\alpha}]_i^j d\xi^\alpha \otimes e_a^i(x) \otimes e_{aj}(x)$ where $[\Gamma_\alpha] = \imath'(1)\grave{A}_\alpha$ is a matrix on the standard representation $(\mathbb{R}^m, \imath'(1))$ of $o(\mathbb{R}, r, s)$ so : $[\Gamma_\alpha]^t [\eta] + [\eta][\Gamma_\alpha] = 0$

At the transitions : $\Gamma_b(x) = Ad_{r(g_{ba})}\Gamma_a(x) - r'_g(1)R'_{g_{ba}^{-1}}(g_{ba})g'_{ba}(x)$ which reads :
$[\Gamma_b] = -[\partial_\alpha g_{ba}][g_{ab}] + [g_{ba}][\Gamma_a][g_{ab}] = [g_{ba}][\partial_\alpha g_{ab}] + [g_{ba}][\Gamma_a][g_{ab}]$

iii) The covariant derivative of a vector field $U = \sum_i U^i(x) e_i(x) \in \mathfrak{X}(E)$ is :

$\nabla U = \sum (\partial_\alpha U^i + \Gamma_{a\alpha j}^i U^j) d\xi^\alpha \otimes e_{ai}(x)$
which reads in the holonomic basis of M :
$\nabla U = \left(\sum \partial_\alpha \left([L'_a]_\beta^i u^\beta\right) + \Gamma_{a\alpha j}^i [L'_a]_\beta^j u^\beta\right) d\xi^\alpha \otimes [L_a]_i^\gamma \partial \xi_\gamma$
$\nabla U = \left(\sum \partial_\alpha u^\gamma + \widetilde{\Gamma}_{\alpha\beta}^\gamma u^\beta\right) d\xi^\alpha \otimes \partial \xi_\gamma$
with : $\widetilde{\Gamma}_{a\alpha\beta}^\gamma = [L_a]_i^\gamma \left([\partial_\alpha L'_a]_\beta^i + [\Gamma_{a\alpha}]_j^i [L'_a]_\beta^j\right) \Leftrightarrow \left[\widetilde{\Gamma}_{a\alpha}\right] = [L_a]([\partial_\alpha L'_a] + [\Gamma_{a\alpha}][L'_a])$
As ∇U is intrinsic its defines an affine connection, at least on $\pi_E^{-1}(O_a)$.

iv) At the transitions of P :

$$\left[\widetilde{\Gamma}_{b\alpha}\right] = [L_b]\left([\partial_\alpha L'_b] + [\Gamma_{b\alpha}][L'_b]\right)$$

$$= [L_a][g_{ab}]\left(\partial_\alpha\left([g_{ab}][L'_a]\right) + \left([g_{ba}][\partial_\alpha g_{ab}] + [g_{ba}][\Gamma_a][g_{ab}]\right)\left([g_{ba}][L'_a]\right)\right)$$

$$= [L_a][g_{ab}][\partial_\alpha g_{ba}][L'_a] + [L_a][\partial_\alpha L'_a] + [L_a][\partial_\alpha g_{ab}][g_{ba}][L'_a] + [L_a][\Gamma_a][L'_a]$$

$$= [L_a]\left([\partial_\alpha L'_a] + [\Gamma_a][L'_a]\right) + [L_a]\left([g_{ab}][\partial_\alpha g_{ba}] + [\partial_\alpha g_{ab}][g_{ba}]\right)[L'_a]$$

$$= \left[\widetilde{\Gamma}_{a\alpha}\right] + [L_a]\,\partial_\alpha\left([g_{ab}][g_{ba}]\right)[L'_a] = \left[\widetilde{\Gamma}_{a\alpha}\right]$$

Thus the affine connection is defined over TM.

v) This connexion is metric with respect to g in the usual meaning for affine connection.

The condition is : $\forall \alpha, \beta, \gamma : \partial_\gamma g_{\alpha\beta} = \sum_\eta \left(g_{\alpha\eta}\Gamma^\eta_{\gamma\beta} + g_{\beta\eta}\Gamma^\eta_{\gamma\alpha}\right) \Leftrightarrow [\partial_\alpha g] = $

$[g]\left[\widetilde{\Gamma}_\alpha\right] + \left[\widetilde{\Gamma}_\alpha\right]^t[g]$

with the metric g defined from the orthonormal basis :$[L]^t[g][L] = [\eta] \Leftrightarrow$
$[g] = [L']^t[\eta][L']$

$$[\partial_\alpha g] = [\partial_\alpha L']^t[\eta][L'] + [L']^t[\eta][\partial_\alpha L']$$
$$- [L']^t[\partial_\alpha L]^t[L'][\eta][L'] + [L']^t[\eta][\partial_\alpha L']^t$$

$$[g]\left[\widetilde{\Gamma}_\alpha\right] + \left[\widetilde{\Gamma}_\alpha\right]^t[g] = [L']^t[\eta][L']\left([L]\left([\partial_\alpha L'] + [\Gamma_\alpha][L']\right)\right) + \left([\partial_\alpha L']^t + [L']^t[\Gamma_\alpha]^t\right)[L]^t[L']^t[\eta][L']$$

$$= [L']^t[\eta][L'][L][\partial_\alpha L'] + [L']^t[\eta][L'][L][\Gamma_\alpha][L'] + [\partial_\alpha L']^t[\eta][L'] + [L']^t[\Gamma_\alpha]^t[\eta][L']$$

$$= [L']^t[\eta][\partial_\alpha L'] + [\partial_\alpha L']^t[\eta][L'] + [L']^t\left([\eta][\Gamma_\alpha] + [\Gamma_\alpha]^t[\eta]\right)[L']$$

$$= [L']^t[\eta][\partial_\alpha L'] + [\partial_\alpha L']^t[\eta][L'] \text{ because } [\Gamma_\alpha] \in o(\mathbb{R}, r, s)$$

vi) Conversely an affine connection $\widetilde{\Gamma}^\gamma_{\alpha\beta}$ defines locally the maps:

$\Gamma_a(x) = \sum_{\alpha ij}[\Gamma_{a\alpha}]^j_i\, d\xi^\alpha \otimes e^i_a(x) \otimes e_{aj}(x)$ with $[\Gamma_{a\alpha}] = [L'_a]\left(\left[\widetilde{\Gamma}_\alpha\right][L_a] + [\partial_\alpha L_a]\right)$

At the transitions :

$$[\Gamma_{b\alpha}] = [L'_b]\left(\left[\widetilde{\Gamma}_\alpha\right][L_b] + [\partial_\alpha L_b]\right) = [g_{ba}][L'_a]\left(\left[\widetilde{\Gamma}_\alpha\right][L_a][g_{ab}] + [\partial_\alpha L_a][g_{ab}] + [L_a][\partial_\alpha g_{ab}]\right)$$

$$= [g_{ba}][L'_a]\left(\left[\widetilde{\Gamma}_\alpha\right][L_a] + [\partial_\alpha L_a]\right)[g_{ab}] + [g_{ba}][L'_a][L_a][\partial_\alpha g_{ab}]$$

$$= Ad_{[g_{ba}]}[\Gamma_{a\alpha}] - [\partial_\alpha g_{ba}][g_{ab}] = Ad_{[g_{ba}]}\left([\Gamma_{a\alpha}] - [g_{ab}][\partial_\alpha g_{ba}]\right)$$

$$= Ad_{r'(1)g_{ba}}\left(r'(1)\grave{A}_{a\alpha} - r'(1)L'_{g^{-1}_{ba}}(g_{ba})\partial_\alpha g_{ba}\right)$$

$$r'(1)\grave{A}_{b\alpha} = r'(1)Ad_{g_{ba}}\left(\grave{A}_{a\alpha} - L'_{g^{-1}_{ba}}(g_{ba})\partial_\alpha g_{ba}\right)$$

$$\grave{A}_b = Ad_{g_{ba}}\left(\grave{A}_a - L'_{g^{-1}_{ba}}(g_{ba})g'_{ba}\right)$$

So it defines a principal connection on P. ∎

Remarks :

i) Whenever we have a principal bundle $P(M, O(\mathbb{R}, r, s), \pi)$ of frames on TM, using the standard representation of $O(\mathbb{R}, r, s)$ we can build E, which can be assimilated to TM. E can be endowed with a metric, by importing the metric on $\mathbb{R}^m$ through the standard representation, which is equivalent to define g by : $[g] = [L']^t[\eta][L']$. (M,g) becomes a pseudo-riemannian manifold. The bundle of frames is the geometric definition of a pseudo-riemannian manifold. Not any manifold admit such a metric, so the topological obstructions lie also on the level of P.

ii) With any principal connection on P we have a metric affine connexion on TM. So such connections are not really special : using the geometric definition of (M,g), any affine connection is metric with the induced metric. The true particularity of the Lévy-Civita connection is that it is torsion free.

26.4.4   Connection on a Spin bundle

The following theorem is new. With a principal bundle $Sp(M,Spin\,(\mathbb{R},p,q),\pi_S)$ any representation (V,r) of the Clifford algebra $Cl(\mathbb{R},r,s)$ becomes an associated vector bundle E=P[V,r], and even a fiber bundle. So any principal connection on Sp induces a linear connection on E the usual way. However the existence of the Clifford algebra action permits to define another connection, which has some interseting properties and is the starting point of the Dirac operator seen in Functional Analysis. For the details of the demonstration see Clifford Algebras in the Algebra part.

.

Theorem 2103 *For any representation (V,r) of the Clifford algebra $Cl(\mathbb{R},r,s)$ and principal bundle $Sp(M,Spin\,(\mathbb{R},r,s),\pi_S)$ the associated bundle $E=Sp[V,r]$ is a spin bundle. Any principal connection Ω on Sp with potential $\grave{A}$ induces a linear connection with form $\Gamma = r\left(\sigma\left(\grave{A}\right)\right)$ on E and covariant derivative ∇ . Moreover, the representation $[\mathbb{R}^m,\boldsymbol{Ad}]$ of $Spin\,(\mathbb{R},r,s),\pi_S)$ leads to the associated vector bundle $F=Sp[\mathbb{R}^m,\boldsymbol{Ad}]$ and Ω induces a linear connection on F with covariant derivative $\widehat{\nabla}$. There is the relation :*
$$\forall X \in \mathfrak{X}(F), U \in \mathfrak{X}(E) : \nabla\left(r\left(X\right)U\right) = r\left(\widehat{\nabla}X\right)U + r\left(X\right)\nabla U$$

This property of ∇ makes of the connection on E a "Clifford connection".
Proof. i) The ingredients are the following :

The Lie algebra $o(\mathbb{R},r,s)$ with a basis $(\overrightarrow{\kappa}_\lambda)_{\lambda=1}^q$.

$(\mathbb{R}^m,g)$ endowed with the symmetric bilinear form γ of signature (r,s) on $\mathbb{R}^m$ and its basis $(\varepsilon_i)_{i=1}^m$

The standard representation $(\mathbb{R}^m,\jmath)$ of $SO(\mathbb{R},r,s)$ thus $(\mathbb{R}^m,\jmath'(1))$ is the standard representation of $o(\mathbb{R},r,s)$ and $[J] = \jmath'(1)\overrightarrow{\kappa}$ is the matrix of $\overrightarrow{\kappa} \in o(\mathbb{R},r,s)$

The isomorphism : $\sigma : o(\mathbb{R},r,s) \to T_1 SPin(\mathbb{R},r,s) :: \sigma\left(\overrightarrow{\kappa}\right) = \sum_{ij}[\sigma]_j^i\, \varepsilon_i \cdot \varepsilon_j$ with $[\sigma] = \frac{1}{4}[J][\eta]$

A principal bundle $Sp(M,Spin\,(\mathbb{R},r,s),\pi_S)$ with atlas $(O_a,\varphi_a)_{a\in A}$.transition maps $g_{ba}(x) \in Spin\,(\mathbb{R},r,s)$ and right action ρ endowed with a connection of potential $\grave{A}$, $\grave{A}_b(x) = Ad_{g_{ba}}\left(\grave{A}_a(x) - L'_{g_{ba}^{-1}}(g_{ba})g'_{ba}(x)\right)$. On a Spin group the adjoint map reads :$Ad_s(\overrightarrow{\kappa}) = \boldsymbol{Ad}_s\sigma(\overrightarrow{\kappa})$ so : $\sigma\left(\grave{A}_b(x)\right) = g_{ba}\cdot\sigma\left(\grave{A}_a(x) - g_{ab}(x)\cdot g'_{ba}(x)\right)\cdot g_{ab}$

A representation (V,r) of $Cl\,(\mathbb{R},r,s)$, with a basis $(e_i)_{i=1}^n$ of V, defined by the nxn matrices $\gamma_i = r(\varepsilon_i)$ and : $\gamma_i\gamma_j + \gamma_j\gamma_i = 2\eta_{ij}I, \eta_{ij} = \pm 1$.

An associated vector bundle E=Sp[V,r] with atlas $(O_a,\psi_a)_{a\in A}$, holonomic basis : $e_{ai}(x) = \psi_a(x,e_i)$, $e_{bi}(x) = r(\varphi_{ba}(x))e_{bi}(x)$ and transition maps on E are : $\psi_{ba}(x) = r(g_{ba}(x))$

$\psi_a : O_a \times V \to E :: \psi_a(x, u) = (\varphi_a(x, 1), u)$

For $x \in O_a \cap O_b : (\varphi_a(x, 1), u_a) \sim (\varphi_b(x, 1), r(g_{ba}) u_a)$

Following the diagram :

$T_x M \overset{\grave{A}}{\to} o(\mathbb{R}, r, s) \overset{\sigma}{\to} Cl(\mathbb{R}, r, s) \overset{r}{\to} G\mathcal{L}(V; V)$

We define the connection on E :

$\widetilde{\Gamma}_a(\varphi_a(x, 1), u_a) = \left(\varphi_a(x, 1), r\left(\sigma\left(\grave{A}_a\right)\right) u_a\right)$

The definition is consistent : for $x \in O_a \cap O_b$:

$U_x = (\varphi_a(x, 1), u_a) \sim (\varphi_b(x, 1), r(g_{ba}) u_a)$

$\widetilde{\Gamma}_b(\varphi_b(x, 1), r(g_{ba}) u_a) = \left(\varphi_b(x, 1), r\left(\sigma\left(\grave{A}_b\right)\right) r(g_{ba}) u_a\right)$

$\sim \left(\varphi_a(x, 1), r(g_{ab}) r\left(\sigma\left(\grave{A}_b\right)\right) r(g_{ba}) u_a\right)$

$= \left(\varphi_a(x, 1), r(g_{ab}) r\left(g_{ba} \cdot \sigma\left(\grave{A}_a(x) - g_{ab}(x) \cdot g'_{ba}(x)\right) \cdot g_{ab}\right) r(g_{ba}) u_a\right)$

$= \left(\varphi_a(x, 1), r\left(\sigma\left(\grave{A}_a(x) - g_{ab}(x) \cdot g'_{ba}(x)\right)\right) u_a\right)$

$= \widetilde{\Gamma}_a(\varphi_a(x, 1), u_a) - (\varphi_a(x, 1), r(\sigma(g_{ab}(x) \cdot g'_{ba}(x))) u_a)$

So we have a linear connection:

$\Gamma_a(x) = r\left(\sigma\left(\grave{A}_a\right)\right) = \sum_{\alpha \lambda i j} \grave{A}^\lambda_{a\alpha} [\theta_\lambda]^i_j d\xi^\alpha \otimes e_{ai}(x) \otimes e^j_a(x)$

with $[\theta_\lambda] = \frac{1}{4} \sum_{ij} ([\eta][J])^k_l r(\varepsilon_k) r(\varepsilon_l) = \frac{1}{4} \sum_{kl} ([J_\lambda][\eta])^k_l ([\gamma_k][\gamma_l])$

The covariant derivative associated to the connection is :

$U \in \mathfrak{X}(E) : \nabla U = \sum_{i\alpha} \left(\partial_\alpha U^i + \sum_{\lambda j} \grave{A}^\lambda_{a\alpha} [\theta_\lambda]^i_j U^j\right) d\xi^\alpha \otimes e_{ai}(x)$

ii) With the representation $(\mathbb{R}^m, \mathbf{Ad})$ of $Spin(\mathbb{R}, r, s)$ we can define an associated vector bundle $F = Sp[\mathbb{R}^m, \mathbf{Ad}]$ with atlas $(O_a, \phi_a)_{a \in A}$ and holonomic basis : $\varepsilon_{ai}(x) = \phi_a(x, \varepsilon_i)$ such that : $\varepsilon_{bi}(x) = \mathbf{Ad}_{g_{ba}(x)} \varepsilon_{ai}(x)$

Because $\mathbf{Ad}$ preserves γ the vector bundle F can be endowed with a scalar product g. Each fiber (F(x),g(x)) has a Clifford algebra structure Cl(F(x),g(x)) isomorphic to Cl(TM)(x) and to Cl($\mathbb{R}, r, s$).

The connection on Sp induces a linear connection on F :

$\widehat{\Gamma}(x) = (\mathbf{Ad})'|_{s=1}\left(\sigma\left(\grave{A}(x)\right)\right) = \sum_\lambda \grave{A}^\lambda_\alpha(x) [J_\lambda]^i_j \varepsilon_i(x) \otimes \varepsilon^j(x)$

and the covariant derivative :

$X \in \mathfrak{X}(F) : \widehat{\nabla} X = \sum_{i\alpha} \left(\partial_\alpha X^i + \sum_{\lambda j} \grave{A}^\lambda_\alpha [J_\lambda]^i_j X^j\right) d\xi^\alpha \otimes \varepsilon_i(x)$

X(x) is in F(x) so in Cl(F(x),g(x)) and acts on a section of E :

$r(X) U = \sum_i r(X^i \varepsilon_i(x)) U = \sum_i X^i [\gamma_i]^k_l U^l e_k(x)$

which is still a section of E. Its covariant derivative reads :

$\nabla(r(X) U) = \sum \left((\partial_\alpha X^i) [\gamma_i]^k_l U^l + X^i [\gamma_i]^k_l \partial_\alpha U^l + \grave{A}^\lambda_\alpha [\theta_\lambda]^k_p X^i [\gamma_i]^p_l U^l\right) d\xi^\alpha \otimes$
$e_k(x)$

On the other hand $\widehat{\nabla} X$ is a form valued in F, so we can compute :

$r\left(\widehat{\nabla} X\right) U = \sum \left(\left(\partial_\alpha X^i + \grave{A}^\lambda_\alpha [J_\lambda]^i_j X^j\right) [\gamma_i]^k_l\right) U^l d\xi^\alpha \otimes e_k(x)$

and similarly :

$r(X) \nabla U = \sum X^i [\gamma_i]^k_l \left(\partial_\alpha U^l + \grave{A}^\lambda_\alpha [\theta_\lambda]^l_j U^j\right) d\xi^\alpha \otimes e_k(x)$

$\nabla(r(X) U) - r\left(\widehat{\nabla} X\right) U - r(X) \nabla U$

$= \sum_{i\alpha} \{(\partial_\alpha X^i) [\gamma_i]^k_l U^l + X^i [\gamma_i]^k_l \partial_\alpha U^l - (\partial_\alpha X^i) [\gamma_i]^k_l U^l - X^i [\gamma_i]^k_l \partial_\alpha U^l$

$$+ \sum\nolimits_{\lambda j} \grave{A}^\lambda_\alpha \, [\theta_\lambda]^k_p \, X^i \, [\gamma_i]^p_l \, U^l - \grave{A}^\lambda_\alpha \, [J_\lambda]^i_j \, X^j U^l \, [\gamma_i]^k_l - X^i \, [\gamma_i]^k_l \, \grave{A}^\lambda_\alpha \, [\theta_\lambda]^l_j \, U^j \} d\xi^\alpha \otimes$$

$e_{ak}(x)$

$$= \sum \grave{A}^\lambda_\alpha \, [\theta_\lambda]^k_p \, X^i \, [\gamma_i]^p_l \, U^l - \grave{A}^\lambda_\alpha \, [J_\lambda]^i_j \, X^j U^l \, [\gamma_i]^k_l - X^i \, [\gamma_i]^k_l \, \grave{A}^\lambda_\alpha \, [\theta_\lambda]^l_j \, U^j$$

$$= \sum\nolimits_{\lambda j} \left([\theta_\lambda] \, [\gamma_i] - [\gamma_i] \, [\theta_\lambda] - \sum\nolimits_p [J_\lambda]^p_i \, [\gamma_p] \right)^k_l X^i U^l \grave{A}^\lambda_\alpha$$

$$[\theta_\lambda] \, [\gamma_i] - [\gamma_i] \, [\theta_\lambda] = \tfrac{1}{4} \sum\nolimits_{pq} \left(([J_\lambda] \, [\eta])^p_q \, [\gamma_p] \, [\gamma_q] \, [\gamma_i] - [\gamma_i] \, ([J_\lambda] \, [\eta])^p_q \, ([\gamma_p] \, [\gamma_q]) \right)$$

$$= \tfrac{1}{4} \sum\nolimits_{pq} ([J_\lambda] \, [\eta])^p_q \, ([\gamma_p] \, [\gamma_q] \, [\gamma_i] - [\gamma_i] \, [\gamma_p] \, [\gamma_q])$$

$$= \tfrac{1}{4} \sum\nolimits_{pq} ([J_\lambda] \, [\eta])^p_q \, (r \, (\varepsilon_p \cdot \varepsilon_q \cdot \varepsilon_i - \varepsilon_i \cdot \varepsilon_p \cdot \varepsilon_q))$$

$$= \tfrac{1}{4} \sum\nolimits_{pq} ([J_\lambda] \, [\eta])^p_q \, (2r \, (\eta_{iq}\varepsilon_p - \eta_{ip}\varepsilon_q)) = \tfrac{1}{2} \sum\nolimits_{pq} ([J_\lambda] \, [\eta])^p_q \, (\eta_{iq} \, [\gamma_p] - \eta_{ip} \, [\gamma_q])$$

$$= \tfrac{1}{2} \left(\sum\nolimits_p ([J_\lambda] \, [\eta])^p_i \, \eta_{ii} \, [\gamma_p] - \sum\nolimits_q ([J_\lambda] \, [\eta])^i_q \, \eta_{ii} \, [\gamma_q] \right)$$

$$= \tfrac{1}{2}\eta_{ii} \sum\nolimits_p \left(([J_\lambda] \, [\eta])^p_i - ([J_\lambda] \, [\eta])^i_p \right) [\gamma_p] = \eta_{ii} \sum\nolimits_p ([J_\lambda] \, [\eta])^p_i \, [\gamma_p] = \sum\nolimits_{pj} \eta_{ii} \, [J_\lambda]^p_j \, \eta_{ji} \, [\gamma_p] =$$

$\sum\nolimits_p [J_\lambda]^p_i \, [\gamma_p]$

$[\theta_\lambda] \, [\gamma_i] - [\gamma_i] \, [\theta_\lambda] - \sum\nolimits_p [J_\lambda]^j_i \, [\gamma_j] = 0$ ∎

Notice that the Clifford structure Cl(M) is not defined the usual way, but starting from the principal Spin bundle by taking the Adjoint action.

26.4.5   Chern theory

Given a manifold M and a Lie group G, one can see that the potentials $\sum_{i,\alpha} \grave{A}^i_\alpha(x) \, d\xi^\alpha \otimes \varepsilon_i$ and the forms $\mathcal{F} = \sum_{i,\alpha\beta} \mathcal{F}^i_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta \otimes \varepsilon_i$ resume all the story about the connection. They are forms over M, valued in the Lie algebra, but defined locally, in the domains O_a of the cover of a principal bundle, with transition functions depending on G. So looking at these forms gives some insight about the topological constraints limiting the possible matching of manifolds and Lie groups to build principal bundles.

Chern-Weil theorem

Definition 2104 *For any finite dimensional representation (F,r) on a field K, of a Lie group G $I_s(G, F, r)$ is the set of symmetric s-linear maps $L_s \in \mathcal{L}^s(F; K)$ wich are r invariant.*

$L_s \in I_s(G, F, r)$: $\forall (X_k)^r_{k=1} \in V$, $\forall g \in G$: $L_s(r(g)X_1, ..., r(g)X_s) = L_s(X_1, .., X_s)$

With the product :

$(L_s \times L_t)(X_1, ..X_{t+s}) = \frac{1}{(t+s)!} \sum_{\sigma \in \mathfrak{S}(t+s)} L_r(X_{\sigma(1)}, ...X_{\sigma(s)}) \, L_s(X_{\sigma(s+1)}, ...X_{\sigma(t+s)})$

$I(G, V, r) = \oplus^\infty_{s=0} I_r(G, V, r)$ is a real algebra on the field K.

For any finite dimensional manifold M, any finite dimensional representation (V,r) of a group G, $L_s \in I_s(G, F, r)$ one defines the map : $\widehat{L}_s : \Lambda_p(M; V) \to \Lambda_{sp}(M; \mathbb{R})$ by taking F=$\Lambda_p(M; V)$

$\forall (X_k)^r_{k=1} \in \mathfrak{X}(TM) : \widehat{L}_s(\varpi)(X_1, .., X_{rp})$
$= \frac{1}{(rp)!} \sum_{\sigma \in \mathfrak{S}(rp)} L_r \left(\varpi(X_{\sigma(1)}, ...X_{\sigma(p)}), .., \varpi(X_{\sigma((s-1)p+1)}, ...X_{\sigma(sp)}) \right)$

In particular for any finite dimensional principal bundle P(M,G,π) endowed with a principal connection, the strength form $\mathcal{F} \in \Lambda_2\left(M; T_1G\right)$ and (T_1G, Ad) is a representation of G. So we have for any linear map $L_r \in I_r\left(G, T_1G, Ad\right)$:

$$\forall\left(X_k\right)_{k=1}^r \in T_1G : \widehat{L}_r\left(\mathcal{F}\right)\left(X_1, .., X_{2r}\right)$$
$$= \frac{1}{(2r)!} \sum_{\sigma \in \mathfrak{S}(2r)} L_r\left(\mathcal{F}\left(X_{\sigma(1)}, X_{\sigma(2)}\right), .., \mathcal{F}\left(X_{\sigma(2r-1)}, X_{\sigma(2r)}\right)\right)$$

Theorem 2105 *Chern-Weil theorem (Kobayashi 2 p.293, Nakahara p.422) : For any principal bundle P(M,G,π) and any map $L_r \in I_r\left(G, T_1G, Ad\right)$:*

i) for any connection on P with strength form $\mathcal{F}$: $d\widehat{L}_r\left(\mathcal{F}\right) = 0$

ii) for any two principal connections with strength form $\mathcal{F}_1, \mathcal{F}_2$ there is some form $\lambda \in \Lambda_{2r}\left(M; \mathbb{R}\right)$ such that $\widehat{L}_r\left(\mathcal{F}_1 - \mathcal{F}_2\right) = d\lambda$.

iii) the map : $\chi : L_r \in I_r\left(G, T_1G, Ad\right) \to H^{2r}\left(M\right) :: \chi\left(L_r\right) = [\lambda]$ is linear and when extended to $\chi : I(G) \to H^*\left(M\right) = \oplus_r H^r\left(M\right)$ is a morphism of algebras.

iv) If N is a manifold and f a differentiable map : $f : N \to M$ we have the pull back of P and : $\chi_{f^*} = f^*\chi$

All the forms $\widehat{L}_r\left(\mathcal{F}\right)$ for any principal connection on P(M,G,π) belong to the same class of the cohomology space $H^{2r}\left(M\right)$ of M, which is called the **characteristic class** of P related to the linear map L_r. This class does not depend on the connection, but depends both on P and L_r.

Theorem 2106 *(Nakahara p.426) The characteristic classes of a trivial principal bundle are trivial (meaning $[0]$).*

If we have a representation (V,r) of the group G, the map $r'(1) : T_1G \to \mathcal{L}\left(V; V\right)$ is an isomorphism of Lie algebras. From the identity : $\mathrm{Ad}_{r(g)} r'(1) = r'(1) Ad_g$ where $\mathrm{Ad}_{r(g)} = r(g) \circ \ell \circ r(g)^{-1}$ is the conjugation over $\mathcal{L}\left(V; V\right)$ we can deduce : $Ad_g = r'(1)^{-1} \circ Ad_{r(g)} \circ r'(1)$. If we have a r linear map : $L_r \in \mathcal{L}^r\left(V; K\right)$ which is invariant by $\mathrm{Ad}_{r(g)} : L \circ Ad_{r(g)} = L$ then :

$\widetilde{L}_r = r'(1)^{-1} \circ L \circ r'(1) \in \mathcal{L}^r\left(T_1G; K\right)$ is invariant by Ad_g :

$\widetilde{L}_r = r'(1)^{-1} \circ L \circ r'(1)\left(Ad_g\right) = r'(1)^{-1} \circ L \circ Ad_{r(g)} r'(1) = r'(1)^{-1} \circ L \circ r'(1)$

So with such representations we can deduce other maps in $I_r\left(G, T_1G, Ad\right)$

A special case occurs when we have a real Lie group, and we want to use complex representations.

Chern and Pontryagin classes

Let V be a vector space on the field K. Any symmetric s-linear maps $L_s \in \mathcal{L}^s\left(V; K\right)$ induces a monomial map of degree s : $Q : V \to K :: Q\left(X\right) = L_s\left(X, X, ..., X\right)$. Conversely by polarization a monomial map of degree s induces a symmetric s-linear map. This can be extended to polynomial maps and to sum of symmetric s-linear maps valued in the field K (see Algebra - tensors).

If (V,r) is a representation of the group G, then Q is invariant by the action r of G iff the associated linear map is invariant.

So the set $I_s\left(G, V, r\right)$ is usually defined through invariant polynomials.

Of particular interest are the polynomials $Q:V \to K :: Q(X) = \det(I + kX)$ with $k \in K$ a fixed scalar. If (V,r) is a representation of G, then Q is invariant by $Ad_{r(g)}$:

$$Q\left(Ad_{r(g)}X\right) = \det\left(I + kr(g)Xr(g)^{-1}\right) = \det\left(r(g)(I + kX)r(g)^{-1}\right) = \det(r(g))\det(I + kX)\det r(g)^{-1} = Q(X)$$

The degree of the polynomial is n=dim(V). They define n linear symmetric Ad invariant maps $\widetilde{L}_s \in \mathcal{L}^s(T_1G; K)$

For any principal bundle $P(M,G,\pi)$ the previous polynomials define a sum of maps and for each we have a characteritic class.

If $K=\mathbb{R}, k = \frac{1}{2\pi}, Q(X) = \det(I + \frac{1}{2\pi}X)$ we have the **Pontryagin classes** denoted $p_n(P) \in \Lambda_{2n}(M;\mathbb{C}) \subset H^{2n}(M;\mathbb{C})$

If $K=\mathbb{C}, k = i\frac{1}{2\pi}, Q(X) = \det(I + i\frac{1}{2\pi}X)$ we have the **Chern classes** denoted $c_n(P) \in \Lambda_{2n}(M;\mathbb{R}) \subset H^{2n}(M)$

For 2n > dim(M) then $c_n(P) = p_n(P) = [0]$

27 BUNDLE FUNCTORS

With fiber bundles it is possible to add some mathematical structures above manifolds, and functors can be used for this purpose. However in many cases we have some relations between the base and the fibers, involving differential operators, such as connection or the exterior differential on the tensorial bundle. To deal with them we need to extend the functors to jets. The category theory reveals itself useful, as all the various differential constructs come under a limited number of well known cases.

27.1 Bundle functors

27.1.1 Definitions

(Kolar IV.14)

A functor is a map between categories, a bundle functor is a map between a category of manifolds and the category of fibered manifolds. Here we use fibered manifolds rather than fiber bundles because the morphisms are more easily formalized : this is a pair of maps (F,f) F between the total spaces and f between the bases.

Notation 2107 $\mathfrak{M}$ *is the category of manifolds (with the relevant class of differentiability whenever necessary),*

$\mathfrak{M}_m$ *is the subcategory of m dimensional real manifolds, with local diffeormorphisms,*

$\mathfrak{M}_\infty$ *is the category of smooth manifolds and smooth morphisms*

$\mathfrak{FM}$ *is the category of fibered manifolds with their morphisms,*

$\mathfrak{FM}_m$ *is the category of fibered manifolds with m dimensional base and their local diffeormorphisms,*

The local diffeomorphisms in $\mathfrak{M}_m$ are maps : $f \in \hom_{\mathfrak{M}_m}(M,N) : f \in C_1(M;N) : f'(x) \neq 0$

A functor $F : \mathfrak{M} \mapsto \mathfrak{FM}$ associates :

- to each manifold M a fibered manifold : $M \to F_{o1}(M)(F_{o2}(M), \pi_M)$

- to each local diffeomorphism $f \in C_1(M;N), \phi'(x) \neq 0$ a fibered manifold morphism $F_h(f) \in \hom(F_{o1}(M), F_{o1}(N))$

$F_h(f) = (F_{h1}(f), F_{h2}(f)) : \pi_N \circ F_{h1}(f) = F_{h2}(f) \circ \pi_M$

The base functor $\mathfrak{B} : \mathfrak{FM} \mapsto \mathfrak{M}$ associates its base to a fibered manifold and f to each morphism (F,f).

Definition 2108 *A **bundle functor** (also called natural bundle) is a functor F from the category $\mathfrak{M}_m$ of m dimensional real manifolds and local diffeomorphisms to the category $\mathfrak{FM}$ of fibered manifolds which is :*

i) Base preserving : the composition by the base functor gives the identity :
$\mathfrak{B} \circ F = Id_{\mathfrak{M}}$

ii) Local : if N is a submanifold of M then $F(N)$ is the subfibered manifold of $F(M)$

So if N is an open submanifold of M and $\imath : N \to M$ is the inclusion, then :
$F(N) = \pi_M^{-1}(N)$
$F(\imath) : \pi_M^{-1}(N) \to F(M)$
These two properties are crucial : a bundle functor does not tangle with the base manifolds : they add an object at x in M "over" x.

Theorem 2109 *(Kolar p.138, 204) A bundle functor is :*

*i) **regular** : smoothly parametrized systems of local morphisms are transformed into smoothly parametrized systems of fibered manifold morphisms*

*ii) locally **of finite order** : the operations in the functor do not involve the derivatives of order above r*

i) reads :
If $F : \mathfrak{M}_m \mapsto \mathfrak{FM}$: then if $f : P \times M \to N$ is such that $f(x,.) \in \hom_{\mathfrak{Mm}}(M, N)$, then $\Phi : P \times F(M) \to F(N)$ defined by : $\Phi(x,.) = F f(x,.)$ is smooth.

If $F : \mathfrak{M}_\infty \mapsto \mathfrak{FM}$: then if $f : P \times M \to N$ is smooth , then $\Phi : P \times F(M) \to F(N)$ defined by : $\Phi(x,.) = F f(x,.)$ is smooth.

ii) reads :
$\exists r \in \mathbb{N} : \forall M, N \in \mathfrak{M}_{m,r}, \forall f, g \in \hom_{\mathfrak{M}}(M, N), \forall x \in M : j_x^r f = j_x^r g \Rightarrow F f = F g$

(the equality is checked separately at each point x inside the classes of equivalence).

As the morphisms are local diffeomorphisms $\forall x \in M, j_x^r f \in GJ^r(M, N)$

These two results are quite powerful : they mean that there are not so many ways to add structures above a manifold.

Notation 2110 F^r *is the set of r order bundle functors acting on $\mathfrak{M}$*
F_m^r *is the set of r order bundle functors acting on $\mathfrak{M}_m$*
F_∞^r *is the set of r order bundle functors acting on $\mathfrak{M}_\infty$*

Examples of bundle functors
Take V any vector space, and the functor which associates to each manifold the vector bundle $E(M, \otimes_s^r V, \pi)$

The functor which associates to a manifold the bundle of r-forms : $\Lambda_r : M \mapsto \Lambda_r TM$

The functor $J^r : \mathfrak{M} \mapsto \mathfrak{FM}$ which gives the r jet prolongation of a manifold

The functor $T_m^r : \mathfrak{M}_m \mapsto \mathfrak{FM}$ which gives the principal bundle of r differentiable linear frames of a manifold

The functor $J^r : \mathfrak{FM}_m \mapsto \mathfrak{FM}$ which gives the r jet prolongation of a fibered manifold

The functor which associates the bundle of riemannian metrics (which exists on most manifolds)

27.1.2 Description of r order bundle functors

The strength of the concept of bundle functor is that it applies in the same manner to any manifold, thus to the simple manifold $\mathbb{R}^m$. So when we know what happens in this simple case, we can deduce what happens with any manifold. In many ways this is just the implementation of demonstrations in differential geometry when we take some charts to come back in $\mathbb{R}^m$. Functors enable us to generalize at once all these results.

Fundamental theorem

A bundle functor F acts in particular on the manifold $\mathbb{R}^m$ endowed with the appropriate morphisms.

The set $F_0(\mathbb{R}^m) = \pi_{R^m}^{-1}(0) \in F(\mathbb{R}^m)$ is just the fiber above 0 in the fibered bundle $F(\mathbb{R}^m)$ and is called the standard fiber of the functor

For any two manifolds $M, N \in \mathfrak{M}_{mr}$, we have an action of $GJ^r(M, N)$ (the invertible elements of $J^r(M, N)$) on $F(M)$ defined as follows :

$\Phi_{M,N} : GJ^r(M, N) \times_M F(M) \to F(N) :: \Phi_{M,N}(j_x^r f, y) = (F(f))(y)$

The maps $\Phi_{M,N}$,called the associated maps of the functor,are smooth.

For $\mathbb{R}^m$ at x=0 this action reads :

$\Phi_{R^m,R^m} : GJ^r(\mathbb{R}^m, \mathbb{R}^m) \times_{\mathbb{R}^m} F_0(\mathbb{R}^m) \to F_0(\mathbb{R}^m) :: \Phi_{R^m,R^m}(j_0^r f, 0) = (F(f))(0)$

But $GJ_0^r(\mathbb{R}^m, \mathbb{R}^m)_0 = GL^r(\mathbb{R}, m)$ the r differential group and $F_0(\mathbb{R}^m)$ is the standard fiber of the functor. So we have the action :

$\ell : GL^r(\mathbb{R}, m) \times_{\mathbb{R}^m} F_0(\mathbb{R}^m) \to F_0(\mathbb{R}^m)$

Thus for any manifold $M \in \mathfrak{M}_{mr}$ the bundle of r frames $GT_m^r(M)$ is a principal bundle $GT_m^r(M)(M, GL^r(\mathbb{R}, m), \pi^r)$

And we have the following :

Theorem 2111 *(Kolar p.140) For any bundle functor $F_m^r : \mathfrak{M}_{m,r} \mapsto \mathfrak{FM}$ and manifold $M \in \mathfrak{M}_{mr}$ the fibered manifold $F(M)$ is an associated bundle $GT_m^r(M)[F_0(\mathbb{R}^m), \ell]$ to the principal bundle $GT_m^r(M)$ of r frames of M with the standard left action ℓ of $GL^r(\mathbb{R}, m)$ on the standard fiber $F_0(\mathbb{R}^m)$. A fibered atlas of $F(M)$ is given by the action of F on the charts of an atlas of M. There is a bijective correspondance between the set F_m^r of r order bundle functors acting on m dimensional manifolds and the set of smooth left actions of the jet group $GL^r(\mathbb{R}, m)$ on manifolds.*

That means that *the result of the action of a bundle functor is always some associated bundle*, based on the principal bundle defined on the manifold, and a standard fiber and actions known from $F(\mathbb{R}^m)$. Conversely a left

action $\lambda : GL^r(\mathbb{R}, m) \times V \to V$ on a manifold V defines an associated bundle $GT_m^r(M)[V, \lambda]$ which can be seen as the action of a bundle functor with $F_0(\mathbb{R}^m) = V, \ell = \lambda$.

Theorem 2112 *For any r order bundle functor in F_m^r, its composition with the functor $J^k : \mathfrak{FM} \mapsto \mathfrak{FM}$ gives a bundle functor of order k+r: $F \in F_m^r \Rightarrow J^k \circ F \in F_m^{r+k}$*

Examples

i) the functor which associates to each manifold its tangent bundle, and to each map its derivative, which is pointwise a linear map, and so corresponds to the action of $GL^1(\mathbb{R}, m)$ on $\mathbb{R}^m$.

ii) the functor $T_k^r : T_k^r(M)$ is the principal bundle of r frames $GT_m^r(M)$ itself, the action λ is the action of $GL^r(\mathbb{R}, m)$ on a r frame.

Vector bundle functor

A bundle functor is a vector bundle functor if the result is a vector bundle.

Theorem 2113 *(Kolar p.141) There is a bijective correspondance between the set F_m^r of r order bundle functors acting on m dimensional manifolds and the set of smooth representation $\left(\overrightarrow{V}, \overrightarrow{\ell}\right)$ of $GL^r(\mathbb{R}, m)$. For any manifold M, $F_m^r(M)$ is the associated vector bundle $GT_m^r(M)\left[\overrightarrow{V}, \overrightarrow{\ell}\right]$*

The standard fiber is the vector space $\overrightarrow{V}$ and the left action is given by the representation : $\overrightarrow{\ell} : GL^r(\mathbb{R}, m) \times \overrightarrow{V} \to \overrightarrow{V}$

Example : the tensorial bundle $\otimes TM$ is the associated bundle $GT_m^1(M)\left[\otimes\mathbb{R}^m, \overrightarrow{\ell}\right]$ with the standard action of GL$(\mathbb{R}, m)$ on $\otimes\mathbb{R}^m$

Affine bundle functor

An affine r order bundle functor is such that each $F(M)$ is an affine bundle and each $F(f)$ is an affine morphism. It is deduced from a vector bundle functor $\overrightarrow{F}$ defined by a smooth $\left(\overrightarrow{V}, \overrightarrow{\ell}\right)$ representation of $GL^r(\mathbb{R}, m)$. The action ℓ of $GL^r(\mathbb{R}, m)$ on the standard fiber $V = F_0 R^m$ is given by : $\ell(g)y = \ell(g)x + \overrightarrow{\ell}\left(\overrightarrow{y - x}\right)$.

Example : The bundle of principal connections on a principal fiber bundle QP=J^1E is an affine bundle over E, modelled on the vector bundle $TM^* \otimes VE \to E$

27.2   Natural operators

A common issue in physics is to check gauge invariance of objects defined in a model. When this object, as it is often the case, is some object in a fiber bundle,

the basic requirement is that it does not depend on the choice of a chart. This is for instance the "covariance condition" in General Relativity : the equations should stay valid whatever the maps which are used. When the equations are written in coordinates these requirements impose some conditions, which are expressed in terms of the jacobian J of the transition maps (think to the relations for tensors). When we have several such objects which appear in the same model, there must be some relations between them, which are directly related to the group GL($\mathbb{R}$,n) and its higher derivatives. This is the basic idea behind "natural operators". It appears that almost all such operators are simple combinations of the usual tools, such that connection, curvature, exterior differential,..which legitimates their choice by the physicists in their mathematical models.

27.2.1 Definitions

Natural transformation between bundle fonctors

A **natural transformation** ϕ between two functors $F_1, F_2 : \mathfrak{M} \mapsto \mathfrak{FM}$ is a map denoted $\phi : F \hookrightarrow F_2$ such that the diagram commutes :

$$
\begin{array}{ccccccc}
\mathfrak{M} & & & \mathfrak{FM} & & & \mathfrak{M} \\
\ulcorner \quad \urcorner & & \ulcorner & & & \urcorner & \\
M & \overset{\pi_1}{\leftharpoonup} & F_{1o}(M) & \overset{\phi(M)}{\rightarrow} & F_{2o}(M) & \overset{\pi_2}{\rightharpoonup} & M \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
f \;\downarrow & & \downarrow F_{1m}(f) & & \downarrow F_{2m}(f) & & \downarrow f \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
N & \overset{\pi_1}{\leftharpoonup} & F_{1o}(N) & \overset{\phi(N)}{\rightarrow} & F_{2o}(N) & \overset{\pi_2}{\rightharpoonup} & N \\
\end{array}
$$

So $\phi(M) \in \hom_{\mathfrak{FM}}(F_{1o}(M), F_{2o}(M))$

Any natural transformation is comprised of base preserving morphisms, in the meaning :

$\forall p \in F_{1o}(M) : \pi_2(\phi(M)(p)) = \pi_1(p)$

If F is a vector bundle functor, then the vertical bundle VF is naturally equivalent to $F \oplus F$ (Kolar p.385)

Example : lagrangians.

A r order **lagrangian** on a fiber bundle $E(M, V, \pi)$ is a m-form on M : $L\left(\xi^\alpha, z^i(x), z^i_{\{\alpha_1...\alpha_s\}}(x)\right) d\xi^1 \wedge ... \wedge d\xi^m$ where $\left(\xi^\alpha, z^i(x), z^i_{\{\alpha_1...\alpha_s\}}(x)\right)$ are the coordinates of a section Z of $J^r E$. For a section $S \in \mathfrak{X}(E)$ the lagrangian reads $L \circ j^r S$

We have two functors : $J^r : \mathfrak{M}_m \mapsto \mathfrak{FM}$ and $\Lambda_m : \mathfrak{M}_m \mapsto \mathfrak{FM}$ and $L : J^r \hookrightarrow \Lambda_r$ is a base preserving morphism and a natural transformation.

Theorem 2114 *(Kolar p.142) There is a bijective correspondance between the set of all natural transformations between functors of F^r_m and the set of smooth $GL^r(\mathbb{R}, m)$ equivariant maps between their standard fibers.*

A smooth equivariant map is a map : $\phi : F_{10}R^m \to F_{20}R^m : \forall g \in GL^r(\mathbb{R}, m), p \in F_{10}R^m : \phi(\lambda_1(g, p)) = \lambda_2(g, \phi(p))$

Local operator

Definition 2115 *Let $E_1 (M, \pi_1), E_2 (M, \pi_2)$ be two fibered manifolds with the same base M. A **r order local operator** is a map between their sections $D : \mathfrak{X}_r (E_1) \to \mathfrak{X} (E_2)$ such that $\forall S, T \in \mathfrak{X}_r (E_1), \forall x \in M : l = 0..r : j_x^l S = j_x^l T \Rightarrow D (S) (x) = D (T) (x)$*

So it depends only on the r order derivatives of the sections. It can be seen as the r jet prolongation of a morphism between fibered manifolds.

Naural operator

Definition 2116 *A **natural operator** between two r order bundle functors $F_1, F_2 \in F_m^r$ is :*
i) a set of local operators : $D (F_1 (M) ; F_2 (M)) : \mathfrak{X}_\infty (F_1 (M)) \to \mathfrak{X}_\infty (F_2 (M))$
ii) a map : $\Phi : M \to D (F_1 (M) ; F_2 (M))$ which associates to each manifold an operator between sections of the fibered manifolds, such that :
iii) $\forall S \in \mathfrak{X}_\infty (F_1 (M)), \forall f \in \text{hom} (M, N) : \Phi (N) \left(F_1 (f) \circ S \circ f^{-1} \right) = F_2 (f) \circ \Phi (M) \circ f^{-1}$
iv) for any open submanifold N of M and section S on M : $\Phi (N) (S|_N) = \Phi (M) (S) |_N$

A natural operator is an operator whose local description does not depend on the choice of charts.

Theorem 2117 *There is a bijective correspondance between the set of :*
r order natural operators between functors F_1, F_2 and the set of all natural transformations $J^r \circ F_1 \hookrightarrow F_2$.
k order natural operators between functors $F_1 \in F_m^r, F_2 \in F_m^s$ and the set of all smooth $GL^q (\mathbb{R}, m)$ equivariant maps between $T_m^k (F_1 0 R^m)$ and $F_2 0 R^m$ where $q = \max(r + k, s)$

Examples
i) the commutator between vector fields can be seen as a bilinear natural operator between the functors $T \oplus T$ and T, where T is the 1st order functor : $M \mapsto TM$. The bilinear $GL^2 (\mathbb{R}, m)$ equivariant map is the relation between their coordinates :
$[X, Y]^\alpha = X^\beta \partial_\beta Y^\alpha - Y^\beta X_\beta^\alpha$
And this is the only natural 1st order operator : $T \oplus T \hookrightarrow T$
ii) A r order differential operator is a local operator between two vector bundles : $D : J^r E_1 \to E_2$ (see Functional analysis)

27.2.2 Theorems about natural operators

Theorem 2118 *(Kolar p.222) All natural operators $\wedge_k T^* \hookrightarrow \wedge_{k+1} T^*$ are a multiple of the exterior differential*

Theorem 2119 *(Kolar p.243) The only natural operators $\phi : T_1^r \hookrightarrow T_1^r$ are the maps : $X \to kX, k \in \mathbb{R}$*

Prolongation of a vector field

Theorem 2120 *Any vector field X on a manifold M induces, by a bundle functor F, a projectable vector field, denoted FX and called the **prolongation** of X.*

Proof. The flow $\Phi_X(x,t)$ of X is a local diffeomorphism which has for image by F a fiber bundle morphism :

$\Phi_X(.,t) \in \hom(M,M) \longmapsto F\Phi_X(.,t) \in \hom(FM,FM)$

$F\Phi_X(.,t) = (F(.,t), \Phi_X(.,t))$

Its derivative : $\frac{\partial}{\partial t}F(p.,t)|_{t=0} = W(p)$ defines a vector field on E, which is projectable on M as X. ∎

This operation is a natural operator : $\phi : T \hookrightarrow TF :: \phi(M)X = FX$

Lie derivatives

1. For any bundle functor F, manifold M, vector field X on M, and for any section $S \in \mathfrak{X}(FM)$, the Lie derivative $\pounds_{FX}S$ exists because FX is a projectable vector field on FM.

Example : with $F = $ the tensorial functors we have the Lie derivative of a tensor field over a manifold.

$\pounds_{FX}S \in VFM$ the vertical bundle of FM : this is a section on the vertical bundle with projection S on FM

2. For any two functors $F_1, F_2 : \mathfrak{M}_n \mapsto \mathfrak{FM}$, the Lie derivative of a base preserving morphism $f : F_1M \to F_2M :: \pi_2(f(p)) = \pi_1(p)$, along a vector field X on a manifold M is : $\pounds_X f = \pounds_{(F_1X, F_2X)}f$, which exists (see Fiber bundles) because (F_1X, F_2X) are vector fields projectable on the same manifold.

3. Lie derivatives commute with linear natural operators

Theorem 2121 *(Kolar p.361) Let $F_1, F_2 : \mathfrak{M}_n \mapsto \mathfrak{FM}$ be two functors and $D : F_1 \hookrightarrow F_2$ a natural operator which is linear. Then for any section S on F_1M, vector field X on M : $D(\pounds_X S) = \pounds_X D(S)$*

(Kolar p.362) If D is a natural operator which is not linear we have $VD(\pounds_X S) = \pounds_X DS$ where VD is the vertical prolongation of $D : VD : VF_1M \hookrightarrow VF_2M$

A section Y of the vertical bundle VF_1 can always be defined as : $Y(x) = \frac{d}{dt}U(t,x) = \frac{d}{dt}\varphi(x, u(t,x))|_{t=0}$ where $U \in \mathfrak{X}(F_1M)$ and VD(Y) is defined as : $VD(Y) = \frac{\partial}{\partial t}DU(t,x)|_{t=0} \in VF_2M$

Theorem 2122 *(Kolar p.365) If $E_k, k = 1..n, F$ are vector bundles over the same oriented manifold M, D a linear natural operator $D : \oplus_{k=1}^n E_k \to F$ then :*
$\forall S_k \in \mathfrak{X}_\infty(E_k), X \in \mathfrak{X}_\infty(TM) : \pounds_X D(S_1, .., S_n) = \sum_{k=1}^n D(S_1, ., \pounds_X S_k, ., S_n)$
    Conversely every local linear operator which satisfies this identity is the value of a unique natural operator on $\mathfrak{M}_n$

The bundle of affine connections on a manifold

Any affine connection on a m dimensional manifold (in the usual meaning of Differential Geometry) can be seen as a connection on the principal fiber bundle $P^1(M, GL^1(\mathbb{R}, m), \pi)$ of its first order linear frames. The set of connections on P^1 is a quotient set $QP^1 = J^1 P^1 / GL^1(\mathbb{R}, m)$ and an affine bundle $QP^1 M$ modelled on the vector bundle $TM \otimes TM^* \otimes TM^*$. The functor $QP^1 : \mathfrak{M}_m \mapsto \mathfrak{FM}$ associates to any manifold the bundle of its affine connections.

Theorem 2123 *(Kolar p.220)*
    i) all natural operators : $QP^1 \hookrightarrow QP^1$ are of the kind :
    $\phi(M) = \Gamma + k_1 S + k_2 I \otimes \widehat{S} + k_3 \widehat{S} \otimes I, k_i \in \mathbb{R}$
    with : the torsion $S = \Gamma^\alpha_{\beta\gamma} - \Gamma^\alpha_{\gamma\beta}$ and $\widehat{S} = \Gamma^\alpha_{\beta\gamma} + \Gamma^\alpha_{\gamma\beta}$
    ii) the only natural operator acting on torsion free connections is the identity

Curvature like operators

The curvature Ω of a connection on a fiber bundle $E(M, V, \pi)$ can be seen as a 2-form on M valued in the vertical bundle : $\Omega(p) \in \Lambda_2 T_{\pi(p)} M^* \otimes V_p E$. So, viewing the connection itself as a section of $J^1 E$, the map which associates to a connection its curvature is a natural operator : $J^1 \hookrightarrow \Lambda_2 TB^* \otimes V$ between functors acting on $\mathfrak{FM}$, where B is the base functor and V the vertical functor.

Theorem 2124 *(Kolar p.231) : all natural operators $J^1 \hookrightarrow \Lambda_2 TB^* \otimes V$ are a constant multiple of the curvature operator.*

Operators on pseudo-riemannian manifolds

Theorem 2125 *(Kolar p.244) Any r order natural operator on pseudo-riemannian metrics with values in a first order natural bundle factorzes through the metric itself and the derivative, up to the order r-2, of the covariant derivative of the curvature of the Lévy-Civita connection.*

Theorem 2126 *(Kolar p.274, 276) The Lévy-Civita connection is the only conformal natural connection on pseudo-riemannian manifolds.*

With the precise meaning :
For a pseudo-riemmanian manifold (M,g), a map : $\phi(g, S) = \Gamma$ which defines the Christoffel form of an affine connection on M, with S some section of a functor bundle over M, is conformal if $\forall k \in \mathbb{R} : \phi(k^2 g, S) = \phi(g, S)$. So the only natural operators ϕ which are conformal are the map which defines the Lévy-Civita connection.

27.3 Gauge functors

Gauge functors are basically functors for building associated bundles over a principal bundle with a fixed group. We have definitions very similar to the previous ones.

27.3.1 Definitions

Category of principal bundles

Notation 2127 $\mathfrak{PM}$ *is the category of principal bundles*

It comprises :
Objects : principal bundles $P(M, G, \pi)$
Morphisms : $\hom(P_1, P_2)$ are defined by a couple $F : P_1 \rightarrow P_2$ and $\chi \in \hom(G_1, G_2)$

The principal bundles with the same group G is a subcategory $\mathfrak{PM}(G)$. If the base manifold has dimension m we have the subcategory $\mathfrak{PM}_m(G)$

Gauge functors

Definition 2128 *A* **gauge functor** *(or gauge natural bundle) is a functor* $F : \mathfrak{PM}(G) \mapsto \mathfrak{FM}$ *from the category of principal bundles with group G to the category $\mathfrak{FM}$ of fibered manifolds which is :*

i) Base preserving : the composition by the base functor gives the identity : $\mathfrak{B} \circ F = Id_{\mathfrak{M}}$

ii) Local : if N is a submanifold of M then $F(N)$ is the subfibered manifold of $F(M)$

i) Base preserving means that :
Every principal bundle $P \in \mathfrak{PM}$ is transformed in a fibered manifold with the same base : $BF(P) = B(P)$
The projections : $\pi : P \rightarrow BP = M$ form a natural transformation : $F \hookrightarrow B$
Every morphism $(F : P_1 \rightarrow P_2, \chi = Id_G) \in \hom(P_1, P_2)$ is transformed in a fibered manifold morphism (FF, f) such that $f(\pi_1(p)) = \pi_2(FF(p))$
The morphism in G is the identity
ii) Local means that :
if N is an open submanifold of M and $\imath : N \rightarrow M$ is the inclusion, then :
$F(N) = \pi_M^{-1}(N)$
$F(\imath) : \pi_M^{-1}(N) \rightarrow F(M)$

Theorem 2129 *(Kolar p.397) Any gauge functor is :*
i) regular
ii) of finite order

i) Regular means that :

if $f : P \times M \to N$ is a smooth map such that $f(x,.)$ is a local diffeomorphism, then $\Phi : P \times F(M) \to F(N)$ defined by : $\Phi(x,.) = F f(x,.)$ is smooth.

ii) Of finite order means that :

$\exists r, 0 \leq r \leq \infty$ if $\forall f, g \in \hom_{\mathfrak{M}_n}(M, N), \forall p \in M : j_p^r f = j_p^r g \Rightarrow F f = F g$

Natural gauge operators

Definition 2130 *A **natural gauge operator** between two gauge functors.*
F_1, F_2 *is :*

i) a set of local operators, maps between sections on the fibered manifolds :
$A(F_1(P); F_2(P)) : C_\infty(M; F_1(P)) \to C_\infty(M; F_2(P))$

ii) a map : $\Phi : M \to A(F_1(M); F_2(M))$, *such that :*

iii) $\forall S \in C_\infty(M; F_1(M)), \forall f \in \hom(M, N) : \Phi(N)\left(F_1(f) \circ S \circ f^{-1}\right) = F_2(f) \circ \Phi(M) \circ f^{-1}$

ii) for any open submanifold N of M and section S on M : $\Phi(N)(S|_N) = \Phi(M)(S)|_N$

Every gauge natural operator has a finite order r.

27.3.2 Theorems

$W^r P = GT_m^r(M) \times_M J^r P, W_m^r G = GL^r(\mathbb{R}, m) \rtimes T_m^r(G)$ (see Jets)

Theorem 2131 *(Kolar p.398) There is a bijective correspondance between the set of r order gauge natural operators between gauge functors F_1, F_2 and the set of all natural transformations $J^r \circ F_1 \hookrightarrow F_2$.*

Theorem 2132 *(Kolar p.398) There is a canonical bijection between natural transformations between to r order gauge functors $F_1, F_2 : \mathfrak{PM}_m(G) \mapsto \mathfrak{FM}$ and the $W_m^r G$ equivariant maps :* $\pi_{F_1(R^m \times G)}^{-1}(0) \to \pi_{F_2(R^m \times G)}^{-1}(0)$

Theorem 2133 *(Kolar p.396) For any r order gauge functor F , and any principal bundle $P \in \mathfrak{PM}_m(G)$, $F P$ is an associated bundle to the r jet prolongation of P: $W^r P = GT_m^r(M) \times_M J^r P, W_m^r G = GL^r(\mathbb{R}, m) \rtimes T_m^r(G)$.*

Theorem 2134 *(Kolar p.396) Let G be a Lie group, V a manifold and and λ a left action of $W_m^r G$ on V, then the action factorizes to an action of $W_m^k G$ with $k \leq 2 \dim V + 1$. If m>1 then $k \leq \max\left(\frac{\dim V}{m-1}, \frac{\dim V}{m} + 1\right)$*

27.3.3 The Utiyama theorem

(Kolar p.400)

Let P(M,G,π) be a principal bundle. The bundle of its principal connections is the quotient set : $QP = J^1 P/G$ which can be assimilated to the set of

potentials : $\left\{ \grave{A}\left(x\right) _{\alpha }^{i}\right\}$. The adjoint bundle of P is the associated vector bundle E=$P\left[T_{1}G,Ad\right]$.

QP=$J^{1}E$ is an affine bundle over E, modelled on the vector bundle $TM^{\ast }\otimes VE\rightarrow E$

The strength form $\mathcal{F}$ of the connection can be seen as a 2-form on M valued in the adjoint bundle $\mathcal{F}\in \Lambda _{2}\left(M;E\right) .$

A r order lagrangian on the connection bundle is a map : $\mathcal{L}:J^{r}QP\rightarrow \mathfrak{X}\left(\Lambda _{m}TM^{\ast }\right)$ where m = dim(M). This is a natural gauge operator between the functors : $J^{r}Q\hookrightarrow \Lambda _{m}B$

The Utiyama theorem reads : all first order gauge natural lagrangian on the connection bundle are of the form $A\circ \mathcal{F}$ where A is a zero order gauge lagrangian on the connection bundle and $\mathcal{F}$ is the strength form $\mathcal{F}$ operator.

More simply said : any first order lagrangian on the connection bundle involves only the curvature $\mathcal{F}$ and not the potential $\grave{A}$.

Part VII

PART 7 : FUNCTIONAL ANALYSIS

Functional anaysis studies functions, meaning maps which are value in a field, which is $\mathbb{R}$ or, usually, $\mathbb{C}$, as it must be complete. So the basic property of the spaces of such functions is that they have a natural structure of topological vector space, which can be enriched with different kinds of norms, going from locally convex to Hilbert spaces. Using these structures they can be enlarged to "generalized functions", or distributions.

Functional analysis deals with most of the day to day problems of applied mathematics : differential equations, partial differential equations, optimization and variational calculus. For this endeaour some new tools are defined, such that Fourier transform, Fourier series and the likes. As there are many good books and internet sites on these subjects, we will focus more on the definitions and principles than on practical methods to solve these problems.

28 SPACES OF FUNCTIONS

The basic material of functional analysis is a space of functions, meaning maps from a topological space to a field.The field is usually $\mathbb{C}$ and this is what we assume if not stated otherwise.

Spaces of functions have some basic algebraic properties, which are usefull. But this is their topological properties which are the most relevant. Functions can be considered with respect to their support, boundedness, continuity, integrability. For these their domain of definition does not matter much, because their range in the very simple space $\mathbb{C}$.

When we consider differentiability complications loom. First the domain must be at least a manifold M. Second the partial derivatives are no longer functions : they are maps from M to tensorial bundles over M. It is not simple to define norms over such spaces. The procedures to deal with such maps are basically the same as what is required to deal with sections of a vector bundle (indeed the first derivative f'(p) of a function over M is a vector field in the cotangent bundle TM*). So we will consider both spaces of functions, on any topological space, including manifolds, and spaces of sections of a vector bundle.

In the first section we recall the main results of algebra and analysis which will be useful, in the special view when they are implemented to functions. We define the most classical spaces of functions of functional analysis and we add some new results abut spaces of sections on a vector bundle. A brief recall of functionals lead naturally to the definition and properties of distributions, which can be viewed as "generalised functions", with another new result : the definition of distributions over vector bundles.

28.1 Preliminaries

28.1.1 Algebraic preliminaries

This is a reminder of topics addressed in the previous parts.

Theorem 2135 *The space of functions : $V : E \rightarrow \mathbb{C}$ is a commutative *-algebra*

This is a vector space and a commutattive algebra with pointwise multiplication : $(f \cdot g)(x) = f(x) g(x)$. It is unital with $I : E \rightarrow \mathbb{C} :: I(x) = 1$, and a *-algebra with the involution : $: C(E; \mathbb{C}) \rightarrow C(E; \mathbb{C}) :: \overline{(f)}(x) = \overline{f}(x)$.

With this involution the functions C(E;$\mathbb{R}$) are the subalgebra of hermitian elements in C(E;$\mathbb{C}$).

The usual algebraic definitions of ideal, commutant, self-adjoint,...functions are fully valid. Notice that no algebraic or topological structure on E is necessary for this purpose.

Theorem 2136 *The spectrum of a function in $V : E \rightarrow \mathbb{C}$ is its range :* Sp(f)=f(E).

Proof. The spectrum $Sp(f)$ of f is the subset of the scalars $\lambda \in \mathbb{C}$ such that $(f - \lambda Id_V)$ has no inverse in V.

$\forall y \in f(E) : (f(x) - f(y)) g(x) = x$ has no solution ∎

Of course quite often we will consider only some subset F of this algebra, which may or may not be a subalgebra. Usually when V is a vector subspace it is infinite dimensional, thus it has been useful in the previous parts to extend all definitions and theorems to this case.

We have also maps from a set of functions to another : $L : C(E; \mathbb{C}) \rightarrow C(F; \mathbb{C})$. When the map is linear (on $\mathbb{C}$) it is customary to call it an **operator**. The set of operators between two algebras of functions is itself a *-algebra.

Definition 2137 *The **tensorial product of two functions** $f_1 \in C(E_1; \mathbb{C}), f_2 \in C(E_2; \mathbb{C})$ is the function :*

$f_1 \otimes f_2 \in C(E_1 \times E_2; \mathbb{C}) :: f_1 \otimes f_2(x_1, x_2) = f_1(x_1) f_2(x_2)$

Definition 2138 *The **tensorial product** $V_1 \otimes V_2$ of two vector spaces $V_1 \subset C(E_1; \mathbb{C}), V_2 \subset C(E_2; \mathbb{C})$ is the vector subspace of $C(E_1 \times E_2; \mathbb{C})$ generated by the maps : $E_1 \times E_2 \rightarrow \mathbb{C} :: F(x_1, x_2) = \sum_{i \in I, j \in J} a_{ij} f_i(x_1) f_j(x_2)$ where $(f_i)_{i \in I} \in V_1^I, (f_j)_{j \in J} \in V_2^J, a_{ij} \in \mathbb{C}$ and I,J are finite sets.*

It is customary to call **functional** the maps : $\lambda : C(E; \mathbb{C}) \rightarrow \mathbb{C}$. The space of functionals of a vector space V of functions is the algebraic dual of V, denoted V*. The space of continuous functionals (with the topology on V) is its topological dual denoted V'. It is isomorphic to V iff V is finite dimensional, so this will not be the case usually.

Notice that the convolution * of functions (see below), when defined, brings also a structure of *-algebra, that we will call the convolution algebra.

28.1.2   Topologies on spaces of functions

The set of functions $C(E; \mathbb{C}) = \mathbb{C}^E$ is huge, and the first way to identify interesting subsets is by considering the most basic topological properties, such that continuity. On the $\mathbb{C}$ side everything is excellent : $\mathbb{C}$ is a 1 dimensional Hilbert space. So most will depend on E, and the same set can be endowed with different topologies. It is common to have several non equivalent topologies on a space of functions. The most usual are the following (see Analysis-Normed vector spaces), in the pecking order of interesting properties.

This is mostly a reminder of topics addressed in the previous parts.

Weak topologies

These topologies are usually required when we look for "pointwise convergence".

1. General case:

One can define a topology without any assumption about E.

Definition 2139 *For any vector space of functions $V : E \to \mathbb{C}$, Λ a subspace of V^*, the weak topology on V with respect to Λ is defined by the family of semi-norms : $\lambda \in \Lambda : p_\lambda(f) = |\lambda(f)|$*

This is the initial topology, usually denoted $\sigma(V, \Lambda)$. The open subsets of V are defined by the base : $\{\lambda^{-1}(\varpi), \varpi \in \Omega_{\mathbb{C}}\}$ where $\Omega_{\mathbb{C}}$ are the open subsets of $\mathbb{C}$. With this topology all the functionals in Λ are continuous.

A sequence $(f_n)_{n \in \mathbb{N}} \in V^{\mathbb{N}}$ converges to $f \in V$ iff $\forall \lambda \in \Lambda : \lambda(f_n) \to \lambda(f)$

We have similarly :

Definition 2140 *For any vector space of functions $V : E \to \mathbb{C}$, the *weak topology on V^* with respect to V is defined by the family of semi-norms : $f \in V : p_f(\lambda) = |\lambda(f)|$*

This is the initial topology on V^* using the evaluation maps : $\widehat{f} : V^* \to \mathbb{C} ::$ $\widehat{f}(\lambda) = \lambda(f)$. These maps are continuous with the *weak topology.

The weak topology is Hausdorff iff Λ is separating : $\forall f, g \in V, f \neq g, \exists \lambda \in \Lambda : \lambda(f) \neq \lambda(g)$

2. If V is endowed with a topology:

Usually V has a topology, defined directly, but it is useful to consider weak topologies for some applications.

If V is endowed with some topology and is a vector space, it is a topological vector space. It has a topological dual, then the weak topology on V is the $\sigma(V, V')$ topology. This topology is Hausdorff.

Similarly the *weak topology on V^* is the topology $\sigma(V^*, V)$ and it is also Hausdorff.

The weak and *weak topologies are not equivalent to the initial topology on V, V^*.

If V is endowed with a semi-norm, the weak topology is that of a semi-normed space, with norm : $p(u) = \sup_{\|\lambda\|=1} p_\lambda(u)$ which is not equivalent to the initial norm. This is still true for a norm.

If V is a topological vector space, the product and involution on the algebra C(E;$\mathbb{C}$) are continuous so it is a topological algebra.

Fréchet spaces

A Fréchet space is a Hausdorff, complete, topological vector space, whose metric comes from a countable family of seminorms $(p_i)_{i \in I}$. It is locally convex. As the space is metric we can consider uniform convergence.

But the topological dual of V is not necessarily a Fréchet space.

Compactly supported functions

The functions with compact support are of constant use in functional analysis.

Definition 2141 *A compactly supported function is a function in domain in a topological space E, whose support is enclosed in some compact subset of E.*

the support of f$\in C\left(E;\mathbb{C}\right)$ is the subset of E : Supp(f)=$\overline{\{x \in E : f(x) \neq 0\}}$ or equivalently the complement of the largest open set where f(x) is zero. So the support is a closed subset of E.

Theorem 2142 *If V is a space of functions on a topological space E, then the subspace V_c of compactly supported functions of V is dense in V. So if V is a Fréchet space, then V_c is a Fréchet space.*

Proof. Let V be a space of functions of $C\left(E;\mathbb{C}\right)$, and define V_K the subset of V of functions whose support is the compact K, and V_c the subset of V of functions with compact support : $V_c = \cup_{K \subset E} V_K$. We can define on V the final topology, characterized by the set of opens Ω, with respect to the family of embeddings : $\imath_K : V_K \to V$. Then $\varpi \in \Omega$ is open in V if $\varpi \cap V_K$ is open in each V_K and we can take the sets $\{\varpi \cap V_K, \varpi \in \Omega, K \subset E\}$ as base for the topology of V_c. and V_c is dense in V.

If V is a Fréchet space then each V_K is closed in V so is a Fréchet space. ∎

Theorem 2143 *If V is a space of continuous functions on a topological space E which can be covered by countably many compacts, then the subspace V_c of compactly supported functions of V is a Fréchet space, closed in V.*

(it is sometimes said that E is $\sigma-$compact)

Proof. There is a countable family of compacts $(\kappa_j)_{j \in J}$ such that any compact K is contained in some κ_j . Then V_c is a Fréchet space, closed in V, with the seminorms $q_{i,j}\left(f\right) = \sup_j p_i\left(f|_{\kappa_j}\right)$ where $p_i = \sup \left|f|_{\kappa_j}\right|$. ∎

Banach spaces

A vector space of functions $V \subset C\left(E;\mathbb{C}\right)$ is normed if we have a map : $\|\| : V \to \mathbb{R}_+$ so that :

$\forall f, g \in V, k \in \mathbb{C}$:

$\|f\| \geq 0; \|f\| = 0 \Rightarrow f = 0$

$\|kf\| = |k| \|f\|$

$\|f + g\| \leq \|f\| + \|g\|$

Usually the conditions which are the most difficult to meet are : $\|f\| < \infty$ and $\|f\| = 0 \Rightarrow f = 0$. For the latter it is usually possible to pass to the quotient, meaning that two functions which are equal "almost everywhere" are identical.

The topological dual of V is a normed space with the strong topology : $\|\lambda\| = \sup_{\|f\|=1} |\lambda\left(f\right)|$

If V is a *-algebra such that : $\left\|\overline{f}\right\| = \|f\|$ and $\|f\|^2 = \left\||f|^2\right\|$ then it is a normed *-algebra.

If V is normed and complete then it is a Banach space. Its topological dual is a Banach space.

Theorem 2144 *If a vector space of functions V is a complete normed algebra, it is a Banach algebra. Then the range of any function is a compact of $\mathbb{C}$.*

Proof. the spectrum of any element is just the range of f

the spectrum of any element is a compact subset of $\mathbb{C}$ ∎

If V is also a normed *- algebra, it is a C*-algebra.

Theorem 2145 *The norm on a C*-algebra of functions is necessarily equivalent to :* $\|f\| = \sup_{x \in E} |f(x)|$

Proof. the spectrum of any element is just the range of f

In a normed *-algebra the spectral radius of a normal element f is : $r_\lambda(f) = \|f\|,$

all the functions are normal ff*=f*f

In a C*-algebra : $r_\lambda(f) = \max_{x \in E} |f(x)| = \|f\|$ ∎

Notice that no property required from E.

Hilbert spaces

A Hilbert space H is a Banach space whose norm comes from a positive definite hermitian form, usually denoted $\langle \rangle$. Its dual H' is a also a Hilbert space. In addition to the properties of Banach spaces, Hilbert spaces offer the existence of Hilbertian bases : any function can be written as the series : $f = \sum_{i \in I} f_i e_i$ and of an adjoint operation : $* : H \to H' :: \forall \varphi \in H : \langle f, \varphi \rangle = f^*(\varphi)$

28.1.3 Vector bundles

Manifolds

In this part ("Functional Analysis") we will, if not otherwise stated, assumed that a manifold M is a *finite m dimensional* real Hausdorff class 1 manifold M. Then M has the following properties (see Differential geometry) :

i) it has an equivalent smooth structure, so we can consider only smooth manifolds

ii) it is locally compact, paracompact, second countable

iii) it is metrizable and admits a riemannian metric

iv) it is locally connected and each connected component is separable

v) it admits an atlas with a finite number of charts

vi) every open covering $(O_i)_{i \in I}$ has a refinement $(Q_i)_{i \in I}$ such that $Q_i \sqsubseteq O_i$ and :Q_i has a compact closure, $(Q_i)_{i \in I}$ is locally finite (each points of M meets only a finite number of Q_i), any non empty finite intersection of Q_i is diffeomorphic with an open of $\mathbb{R}^m$

So we will assume that M has a countable atlas $(O_a, \psi_a)_{a \in A}, \overline{O_a}$ compact.

Sometimes we will assume that M is endowed with a volume form ϖ_0. A volume form induces an absolutely continuous Borel measure on M, locally finite (finite on any compact). It can come from any non degenerate metric on M. With the previous properties there is always a riemannian metric so such a volume form always exist.

Vector bundle

In this part ("Functional Analysis") we will, if not otherwise stated, assumed that a vector bundle $E(M, V, \pi)$ with atlas $(O_a, \varphi_a)_{a \in A}$, transition maps $\varphi_{ab}(x) \in \mathcal{L}(V; V)$ and V a Banach vector space, has the following properties :

i) the base manifold M is a smooth finite m dimensional real Hausdorff manifold (as above)

ii) the trivializations $\varphi_a \in C(O_a \times V; E)$ and the transitions maps $\varphi_{ab} \in C(O_a \cap O_b; \mathcal{L}(V; V))$ are smooth.

As a consequence :

i) The fiber E(x) over x is canonically isomorphic to V,. V can be infinite dimensional, but *fiberwise* there is a norm such that E(x) is a Banach vector space.

ii) the trivilialization $(O_a, \varphi_a)_{a \in A}$ is such that $\overline{O_a}$ compact, and for any x in M there is a finite number of a$\in$ A such that $x \in O_a$ and any intersection $O_a \cap O_b$ is relatively compact.

Sometimes we will assume that E is endowed with a scalar product, meaning there is a family of maps $(g_a)_{a \in A}$ with domain O_a such that $g_a(x)$ is a non degenerate, either bilinear symmetric form (real case), or sesquilinear hermitian form (complex case) on each fiber E(x) and on the transition : $g_{bij}(x) = \sum_{kl} \overline{[\varphi_{ab}(x)]}_i^k [\varphi_{ab}(x)]_j^l g_{akl}(x)$. It is called an inner product it is definite positive.

A scalar product g on a vector space V induces a scalar product on a vector bundle $E(M, V, \pi)$ iff the transitions maps preserve g

Notice :

i) a metric on M, riemannian or not, is of course compatible with any tensorial bundle over TM (meaning defined through the tangent or the cotangent bundle), and can induce a scalar product on a tensorial bundle (see Algebra - tensorial product of maps), but the result is a bit complicated, except for the vector bundles TM, TM*, $\Lambda_r TM^*$.

ii) there is no need for a scalar product on V to be induced by anything on the base manifold. In fact, as any vector bundle can be defined as the associated vector bundle $P[V, r]$ to a principal bundle P modelled on a group G, if (V,r) is a unitary representation of V, any metric on V induces a metric on E. The potential topological obstructions lie therefore in the existence of a principal bundle over the manifold M. For instance not any manifold can support a non riemannian metric (but all can support a riemannian one).

Sections of a vector bundle

A section S on a vector bundle $E(M, V, \pi)$ with trivilialization $(O_a, \varphi_a)_{a \in A}$ is defined by a family $(\sigma_a)_{a \in A}$ of maps $\sigma_a : O_a \to V$ such that : $U(x) = \varphi_a(x, u_a(x)), \forall a, b \in A, O_a \cap O_b \neq \varnothing, \forall x \in O_a \cap O_b : u_b(x) = \varphi_{ba}(x) u_a(x)$ where the transition maps $\varphi_{ba}(x) \in \mathcal{L}(V; V)$

So we assume that the sections are defined over the same open cover as E.

The space $\mathfrak{X}(E)$ of sections on E is a complex vector space of infinite dimension.

As a consequence of the previous assumptions there is, *fiberwise*, a norm for the sections of a fiber bundle and, possibly, a scalar product. But it does not provide by itself a norm or a scalar product for $\mathfrak{X}(E)$: as for functions we need some way to agregate the results of each fiber. This is usually done through a volume form on M or by taking the maximum of the norms on the fibers.

To E we can associate the dual bundle $E'(M, V', \pi)$ where V' is the topological dual of V (also a Banach vector space) and there is fiberwise the action of $\mathfrak{X}(E')$ on $\mathfrak{X}(E)$

If M is a real manifold we can consider the space of complex valued functions over M as a complex vector bundle modelled over $\mathbb{C}$. A section is then simply a function. This identification is convenient in the definition of differential operators.

Sections of the r jet prolongation of a vector bundle

1. The r jet prolongation $J^r E$ of a vector bundle is a vector bundle $J^r E\left(E, J_0^r\left(\mathbb{R}^m, V\right), \pi_0^r\right)$. The vector space $J_0^r\left(\mathbb{R}^m, V\right)$ is a set of r components : $J_0^r\left(\mathbb{R}^m, V\right) = \left(z_s\right)_{s=1}^r$ which can be identified either to multilinear symmetric maps $\mathcal{L}_S^s\left(\mathbb{R}^m, V\right)$ or to tensors $\otimes^r T M^* \otimes V$.

So an element of $J^r E(x)$ is identified with $\left(z_s\right)_{s=0}^r$ with $z_0 \in V$

Sections $Z \in \mathfrak{X}\left(J^r E\right)$ of the r jet prolongation $J^r E$ are maps : $Z : M \to \mathfrak{X}\left(J^r E\right) :: Z(x) = \left(z_s(x)\right)_{s=.0}^r$

Notice that a section on $\mathfrak{X}\left(J^r E\right)$ does not need to come from a r differentiable section in $\mathfrak{X}_r(E)$

2. Because V is a Banach space, there is a norm on the space $\mathcal{L}_S^s\left(\mathbb{R}^m, V\right)$: $z_s \in \mathcal{L}_S^s\left(\mathbb{R}^{\dim M}, V\right) : \|z_s\| = \sup \|z_s\left(u_1, ..., u_s\right)\|_V$ for $\|u_k\|_{\mathbb{R}^{\dim M}} = 1, k = 1...s$

So we can define, fiberwise, a norm on $Z \in J^r E(x)$ by :

either : $\|Z\| = \left(\sum_{s=0}^r \|z_s\|^2\right)^{1/2}$

or : $\|Z\| = \max\left(\|z_s\|, s = 0..r\right)$

which are equivalent.

Depending on the subspace of $\mathfrak{X}\left(J^r E\right)$ we can define the norms :

either : $\|Z\| = \sum_{s=} \int_M \|z_s(x)\| \mu$ for integrable sections, if M is endowed with a measure μ

or : $\|Z\| = \max_{x \in M}\left(\|z_s(x)\|, s = 0..r\right)$

28.1.4 Miscellaneous theorems

About functions which can be useful ...

Theorem 2146 *(Lieb p.76) : let $f \in C(\mathbb{R}; \mathbb{R})$ be a measurable function such that f(x+y)=f(x)+f(y) , then f(x)=kx for some constant k*

Gamma function
$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt = (z-1)\Gamma(z-1) = (z-1)!$
$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z}$

$\sqrt{\pi}\Gamma\left(2z\right) = 2^{2z-1}\Gamma\left(z\right)\Gamma\left(z + \frac{1}{2}\right)$

The area $A\left(S^{n-1}\right)$ of the unit sphere S^{n-1} in $\mathbb{R}^n$ is $A\left(S^{n-1}\right) = \frac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)}$

$\int_0^1 \left(1+u\right)^{-x-y} u^{x-1} du = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$

The Lebegue volume of a ball B(0,r) in $\mathbb{R}^n$ is $\frac{1}{n}A\left(S^{n-1}\right)r^n$

28.2 Spaces of bounded or continuous maps

28.2.1 Spaces of bounded or continuous functions

Let E be a topological space, C(E,$\mathbb{C}$) the set of functions $f : E \to \mathbb{C}$

Notation 2147 *$C_c\left(E;\mathbb{C}\right)$ is the set of functions with compact support*
$C_0\left(E;\mathbb{C}\right)$ is the set of continuous functions
$C_{0b}\left(E;\mathbb{C}\right)$ is the set of bounded continuous functions
$C_{0c}\left(E;\mathbb{C}\right)$ is the set of continuous functions with compact support
$C_{0v}\left(E;\mathbb{C}\right)$ is the set of continuous functions vanishing at infinity : $\forall\varepsilon > 0, \exists K$ compact in E: $\forall x \in K^c : |f(x)| < \varepsilon$

Let (E,d) be a metric space, $\gamma \in [0,1]$

Notation 2148 *$C^\gamma\left(E;\mathbb{C}\right)$ is the space of order γ Lipschitz functions :$\exists C > 0 :$*
$\forall x, y \in E : |f\left(x\right) - f\left(y\right)| \leq Cd\left(x,y\right)^\gamma$

Theorem 2149 *Are commutative C*-algebra with pointwise product and the norm : $\|f\| = \max_{x\in E}|f\left(x\right)|$*
i) $C_b\left(E;\mathbb{C}\right)$
ii) $C_{0b}\left(E;\mathbb{C}\right)$ is E is Hausdorff
iii) $C_c\left(E;\mathbb{C}\right), C_{0v}\left(E;\mathbb{C}\right)$ if E is Hausdorff, locally compact
iv) $C_0\left(E;\mathbb{C}\right)$ is E is compact

Theorem 2150 *$C_{0c}\left(E;\mathbb{C}\right)$ is a commutative normed *-algebra (it is not complete) with pointwise product and the norm : $\|f\| = \max_{x\in E}|f\left(x\right)|$ if E is Hausdorff, locally compact. Moreover $C_{0c}\left(E;\mathbb{C}\right)$ is dense in $C_{0\nu}\left(E;\mathbb{C}\right)$*

Theorem 2151 *$C^\gamma\left(E;\mathbb{C}\right)$ is a Banach space with norm : $\|f\| = \sup|f\left(x\right)| + \sup_{x,y\in E}\frac{|f(x)-f(y)|}{d(x,y)^\gamma}$ if E is metric, locally compact*

Theorem 2152 *Tietze extension theorem (Thill p.257) If E is a normal Hausdorff space and X a closed subset of E, then every function in $C_0\left(X;\mathbb{R}\right)$ can be extended to $C_0\left(E;\mathbb{R}\right)$*

Theorem 2153 *Stone-Weierstrass (Thill p.27): Let E be a locally compact Hausdorff space, then a unital subalgebra A of $C_0\left(E;\mathbb{R}\right)$ such that $\forall x \neq y \in E, \exists f \in A : f\left(x\right) \neq f\left(y\right)$ and $\forall x \in E, \exists f \in A : f\left(x\right) \neq 0$ is dense in $C_0\left(E;\mathbb{R}\right)$.*

Theorem 2154 *(Thill p.258) If X is any subset of E, $C_c(X;\mathbb{C})$ is the set of the restrictions to X of the maps in $C_c(E;\mathbb{C})$*

Theorem 2155 *Arzela-Ascoli (Gamelin p.82) A subset F of $C_0(E;\mathbb{R})$ is relatively compact iff it is equicontinuous*

Theorem 2156 *(Gamelin p.82) If E is a compact metric space, the inclusion : $\imath : C^\gamma(E;\mathbb{C}) \to C_0(E;\mathbb{C})$ is a compact map*

28.2.2 Spaces of bounded or continuous sections of a vector bundle

1. According to our definition, a complex vector bundle $E(M,V,\pi)$ V is a Banach vector space and there is always fiberwise a norm for sections on the vector bundle. So most of the previous definitions can be generalized to sections U$\in \mathfrak{X}(E)$ on E by taking the functions : $M \to \mathbb{R} : \|U(x)\|_E$

Definition 2157 *Let $E(M,V,\pi)$ be a complex vector bundle*
  $\mathfrak{X}(E)$ the set of sections $U : M \to E$
  $\mathfrak{X}_c(E)$ the set of sections with compact support
  $\mathfrak{X}_0(E)$ the set of continuous sections
  $\mathfrak{X}_{0b}(E)$ the set of bounded continuous sections
  $\mathfrak{X}_{0c}(E)$ the set of continuous sections with compact support
  $\mathfrak{X}_{0\nu}(E)$ the set of continuous sections vanishing at infinity : $\forall \varepsilon > 0, \exists K$ compact in M: $\forall x \in K^c : \|U(x)\| < \varepsilon$

As E is necessarily Hausdorff:

Theorem 2158 *Are Banach vector space with the norm : $\|U\|_E = \max_{x \in M} \|U(x)\|$*
  i) $\mathfrak{X}_0(E), \mathfrak{X}_{0b}(E)$
  ii) $\mathfrak{X}_c(E), \mathfrak{X}_{0\nu}(E), \mathfrak{X}_{0c}(E)$ if V is finite dimensional, moreover $\mathfrak{X}_{0c}(E)$ is dense in $\mathfrak{X}_{0\nu}(E)$

2. The r jet prolongation $J^r E$ of a vector bundle is a vector bundle, and there is fiberwise the norm (see above) :
  $\|Z\| = \max(\|z_s\|_s, s = 0..r)$
  so we have similar results for sections $\mathfrak{X}(J^r E)$

28.2.3 Rearrangements inequalities

They addres more specifically the functions on $\mathbb{R}^m$ which vanish at infinity. This is an abstract from Lieb p.80.

Definition 2159 *For any Lebesgue measurable set $E \subset \mathbb{R}^m$ the **symmetric rearrangement** $E^\times$ is the ball B(0,r) centered in 0 with a volume equal to the volume of E.*

$\int_E dx = A(S^{m-1})\frac{r^m}{m}$ where $A(S^{m-1})$ is the area of the unit sphere S^{m-1} in $\mathbb{R}^m$ is $A(S^{m-1}) = \frac{2\pi^{m/2}}{\Gamma(\frac{m}{2})}$

Definition 2160 *The symmetric decreasing rearrangement of any characteristic function 1_E of a measurable set E is the characteristic function of its symmetric rearrangement : $1_{E^\times} = 1_E$*

Theorem 2161 *Any measurable function $f \in C(\mathbb{R}^m; \mathbb{C})$ vanishes at infinity iff : $\forall t > 0 : \int_{|f(x)|>t} dx < \infty$*

Theorem 2162 *For a measurable, vanishing at infinity function $f \in C(\mathbb{R}^m; \mathbb{C})$ the function $f^\times(x) = \int_0^\infty 1^\times_{|f(\tau)>t|}(x)\, dt$. has the following properties :*
 i) $f^\times(x) \geq 0$
 ii) it is radially symmetric and non increasing :
 $\|x\| = \|y\| \Rightarrow f^\times(x) = f^\times(y)$
 $\|x\| \geq \|y\| \Rightarrow f^\times(x) \geq f^\times(y)$
 iii) it is lower semicontinuous : the sets $\{x : f^\times(x) > t\}$ are open
 iv) $|f|, |f^\times|$ are equimeasurable :
 $\{x : f^\times(x) > t\} = \{x : |f(x)| > t\}^\times$
 $\int \{x : f^\times(x) > t\}\, dx = \int \{x : |f(x)| > t\}\, dx$
 If $f \in L^p(\mathbb{R}^m, S, dx, \mathbb{C}) : \|f\|_p = \|f^\times\|_p$
 iv) If $f, g \in C_\nu(\mathbb{R}^m; \mathbb{C})$ and $f \leq g$ then $f^\times \leq g^\times$

Theorem 2163 *If f,g are positive, measurable, vanishing at infinity functions $f, g \in C(\mathbb{R}^m; \mathbb{R}_+)$ $f \geq 0, g \geq 0$ then:$\int_M f(x) g(x)\, dx \leq \int_M f^\times(x) g^\times(x)\, dx$ possibly infinite. If $\forall \|x\| > \|y\| \Rightarrow f^\times(x) > f^\times(y)$ then $\int_M f(x) g(x)\, dx = \int_M f^\times(x) g^\times(x)\, dx \Leftrightarrow g = g^\times$*
 if J is a non negative convex function $J : \mathbb{R} \to \mathbb{R}$ such that $J(0)=0$ then
 $\int_{\mathbb{R}^m} J(f^\times(x) - g^\times(x))\, dx \leq \int_{\mathbb{R}^m} J(f(x) - g(x))\, dx$

Theorem 2164 *For any positive, measurable, vanishing at infinity functions $(f_n)_{n=1}^N \in C(\mathbb{R}^m; \mathbb{R}_+)$, kxN matrix $A=[A_{ij}]$ with $k \leq N : \int_{\mathbb{R}^m} \cdots \int_{\mathbb{R}^m} \prod_{n=1}^N f_n\left(\sum_{i=1}^k A_{in} x_i\right) dx \leq$*

$\int_{\mathbb{R}^m} \cdots \int_{\mathbb{R}^m} \prod_{n=1}^N f_n^\times\left(\sum_{i=1}^k A_{in} x_i\right) dx \leq \infty$

28.3 Spaces of integrable maps

28.3.1 L^p spaces of functions

Definition

(Lieb p.41)

Let (E,S,μ) a measured space with μ a positive measure. If the $\sigma-$algebra S is omitted then it is the Borel algebra. The Lebesgue measure is denoted dx as usual.

1. Take p$\in \mathbb{N}, p \neq 0$, consider the sets of measurable functions
$\mathcal{L}^p (E, S, \mu, \mathbb{C}) = \left\{ f : E \to \mathbb{C} : \int_E |f|^p \mu < \infty \right\}$
$N^p (E, S, \mu, \mathbb{C}) = \left\{ f \in \mathcal{L}^p (E, S, \mu) : \int_E |f|^p \mu = 0 \right\}$
$L^p (E, S, \mu, \mathbb{C}) = \mathcal{L}^p (E, S, \mu, \mathbb{C}) / N^p (E, S, \mu, \mathbb{C})$

Theorem 2165 $L^p (E, S, \mu, \mathbb{C})$ *is a complex Banach vector space with the norm:*
$\|f\|_p = \left(\int_E |f|^p \mu \right)^{1/p}$
$L^2 (E, S, \mu, \mathbb{C})$ *is a complex Hilbert vector space with the scalar product :*
$\langle f, g \rangle = \int_E \overline{f} g \mu$

2. Similarly the sets of bounded measurable functions :
$\mathcal{L}^\infty (E, S, \mu, \mathbb{C}) = \{ f : E \to \mathbb{C} : \exists C \in \mathbb{R} : |f(x)| < C \}$
$\|f\|_\infty = \inf (C \in \mathbb{R} : \mu (\{|f(x)| > C\}) = 0)$
$N^\infty (E, S, \mu, \mathbb{C}) = \{ f \in \mathcal{L}^\infty (E, S, \mu, \mathbb{C}) : \|f\|_\infty = 0 \}$
$L^\infty (E, S, \mu, \mathbb{C}) = \mathcal{L}^\infty (E, S, \mu, \mathbb{C}) / N^\infty (E, S, \mu, \mathbb{C})$

Theorem 2166 $L^\infty (E, S, \mu, \mathbb{C})$ *is a C*-algebra (with pointwise multiplication)*

3. Similarly one defines the spaces :

Notation 2167 $L_c^p (E, S, \mu, \mathbb{C})$ *is the subspace of* $L^p (E, S, \mu, \mathbb{C})$ *with compact support*
$L_{loc}^p (E, S, \mu, \mathbb{C})$ *is the space of functions in* $C(E; \mathbb{C})$ *such that* $\int_K |f|^p \mu < \infty$ *for any compact K in E*

Inclusions

Theorem 2168 *(Lieb p.43)* $\forall p, q, r \in \mathbb{N}$:
$1 \leq p \leq \infty : f \in \mathcal{L}^p (E, S, \mu, \mathbb{C}) \cap \mathcal{L}^\infty (E, S, \mu, \mathbb{C}) \Rightarrow f \in \mathcal{L}^q (E, S, \mu, \mathbb{C})$ *for q* $\geq p$
$1 \leq p \leq r \leq q \leq \infty : \mathcal{L}^p (E, S, \mu, \mathbb{C}) \cap \mathcal{L}^q (E, S, \mu, \mathbb{C}) \subset \mathcal{L}^r (E, S, \mu, \mathbb{C})$

Warning ! Notice that we *do not have* $\mathcal{L}^q (E, S, \mu, \mathbb{C}) \subset \mathcal{L}^p (E, S, \mu, \mathbb{C})$ for q $> $ p unless μ is a finite measure

Theorem 2169 $\lim_{p \to \infty} \|f\|_p = \|f\|_\infty$

28.3.2 Spaces of integrable sections of a vector bundle

Definition

According to our definition, in a complex vector bundle $E(M, V, \pi)$ V is a Banach vector space, thus a normed space and there is always pointwise a norm for sections on the vector bundle. If there is a positive measure μ on M we can generalize the previous definitions to sections $U \in \mathfrak{X}(E)$ on E by taking the functions : $M \rightarrow \mathbb{R} : \|U(x)\|_E$

So we define on the space of sections $U \in \mathfrak{X}(E)$

$p \geq 1 :$

$\mathcal{L}^p(M, \mu, E) = \left\{ U \in \mathfrak{X}(E) : \int_M \|U(x)\|^p \mu < \infty \right\}$

$N^p(M, \mu, E) = \left\{ U \in \mathcal{L}^p(M, \mu, E) : \int_M \|U(x)\|^p \mu = 0 \right\}$

$L^p(M, \mu, E) = \mathcal{L}^p(M, \mu, E) / N^p(M, \mu, E)$

and similarly for $p = \infty$

$\mathcal{L}^\infty(M, \mu, E) = \{ U \in \mathfrak{X}(E) : \exists C \in \mathbb{R} : \|U(x)\| < C \}$

$\|U\|_\infty = \inf \left(C \in \mathbb{R} : \int_{U_c} \mu = 0 \text{ with } U_c = \{x \in M : \|U(x)\| > C\} \right)$

$N^\infty(M, \mu, E) = \{ U \in \mathcal{L}^\infty(M, \mu, E) : \|U\|_\infty = 0 \}$

$L^\infty(M, \mu, E) = \mathcal{L}^\infty(M, \mu, E) / N^\infty(M, \mu, E)$

Notice:

i) the measure μ can be any measure. Usually it is the Lebesgue measure induced by a volume form ϖ_0. In this case the measure is absolutely continuous. Any riemannian metric induces a volume form, and with our assumptions the base manifold has always such a metric, thus has a volume form.

ii) this definition for $p = \infty$ is quite permissive (but it would be the same in $\mathbb{R}^m$) with a volume form, as any submanifold with dimension < dim(M) has a null measure. For instance if the norm of a section takes very large values on a 1 dimensional submanifold of M, it will be ignored.

Theorem 2170 $L^p(M, \mu, E)$ *is a complex Banach vector space with the norm:* $\|U\|_p = \left(\int_M \|U\|^p \varpi_0 \right)^{1/p}$

If the vector bundle is endowed with an inner product g which defines the norm on each fiber E(x), meaning that E(x) is a Hilbert space, we can define the scalar product $\langle U, V \rangle$ of two sections U,V over the fiber bundle with a positive measure on M :

$\langle U, V \rangle_E = \int_M g(x)(U(x), V(x)) \mu$

Theorem 2171 *On a vector bundle endowed with an inner product,* $L^2(M, \mu, E)$ *is a complex Hilbert vector space.*

Spaces of integrable sections of the r jet prolongation of a vector bundle

The r jet prolongation $J^r E$ of a vector bundle E is a vector bundle. If the base M is endowed with a positive measure μ we can similarly define spaces of integrable sections of $J^r E$ with a procedure similar to the procedure for E.

A section $Z \in \mathfrak{X}(J^r E)$ can conveniently be defined, coordinate free, as a map :

$M \to V \times \prod_{s=1}^{r} \left\{ \mathcal{L}_S^s \left(\mathbb{R}^{\dim M}, V \right) \right\} :: Z(s) = (z_s(x), s = 0...r)$

We have a norm on V and so on the space of s linear symmetric continuous maps $\mathcal{L}_S^s \left(\mathbb{R}^{\dim M}, V \right)$:

$z_s(x) \in \mathcal{L}_S^s \left(\mathbb{R}^{\dim M}, V \right) : \|Z_s(x)\| = \sup \|Z_s(u_1, ..., u_s)\|_V$ for $\|u_k\|_{\mathbb{R}^{\dim M}} = 1, k = 1...s$

thus, fiberwise, we have norms $\|z_s(x)\|_s$ for each component and we can define, for each s=0...r the set of functions :

$\mathcal{L}^p \left(M, \mu, \mathcal{L}_S^s \left(\mathbb{R}^{\dim M}, V \right) \right) = \left\{ z_s \in C_0 \left(M; \mathcal{L}_S^s \left(\mathbb{R}^{\dim M}, V \right) \right) : \int_M \|z_s(x)\|_s^p \mu < \infty \right\}$

and the norm :

$\|z_s\|_p = \left(\int_M \|z_s(x)\|_s^p \mu \right)^{1/p}$

From there we define :

$\mathcal{L}^p \left(M, \mu, J^r E \right) = \left\{ Z \in \mathfrak{X}(J^r E) : \sum_{s=0}^{r} \int_M \|z_s(x)\|_s^p \mu < \infty \right\}$

Theorem 2172 *The space $L^p \left(M, \mu, J^r E \right)$ is a Banach space with the norm :*
$\|Z\|_p = \left(\sum_{s=0}^{r} \int_M \|z(x)\|_s^p \varpi_0 \right)^{1/p}$

If the vector bundle is endowed with a scalar product g which defines the norm on each fiber E(x), meaning that E(x) is a Hilbert space, we can define the scalar product on the fibers of $J^r E(x)$.

Z(s) reads as a tensor $V(x) \otimes \odot_s \mathbb{R}^m$

If V_1, V_2 are two finite dimensional real vector space endowed with the bilinear symmetric forms g_1, g_2 there is a unique bilinear symmetric form G on $V_1 \otimes V_2$ such that : $\forall u_1, u_1' \in V_1, u_2, u_2' \in V_2$:$G(u_1 \otimes u_1', u_2 \otimes u_2') = g_1(u_1, u_1') g(u_2, u_2')$ and G is denoted $g_1 \otimes g_2$.(see Algebra - tensorial product of maps) . So if we define G on $V \otimes \odot_s \mathbb{R}^m$ such that :

$G_s \left(Z_{\alpha_1...\alpha_m}^i e_i(x) \otimes \varepsilon^{\alpha_1} .. \otimes \varepsilon^{\alpha_m}, T_{\alpha_1...\alpha_m}^i e_i(x) \otimes \varepsilon^{\alpha_1} .. \otimes \varepsilon^{\alpha_m} \right)$
$= \sum_{i,j=1}^{n} \sum_{\alpha_1..\alpha_m} \overline{Z}_{\alpha_1...\alpha_m}^i T_{\alpha_1...\alpha_m}^j g(x) (e_i(x), e_j(x))$

This is a sesquilinear hermitian form on $V(x) \otimes \odot_s \mathbb{R}^m$ and we have : $\|Z_s(x)\|^2 = G_s(x) (Z_s, Z_s)$

The scalar product is extended to M by : $\langle Z, T \rangle = \sum_{s=0}^{r} \int_M G_s(x) (Z_s, T_s) \mu$

and we have :

Theorem 2173 *On a vector bundle endowed with an inner product $L^2(M, \mu, J^r E)$ is a complex Hilbert vector space.*

Remarks : M is not involved because the maps $\mathcal{L}_S^s \left(\mathbb{R}^{\dim M}, V \right)$ are defined over $\mathbb{R}^{\dim M}$ and not M : there is no need for a metric on M.

Most of the following results can be obviously translated from $\mathcal{L}^p(M, \mu, \mathbb{C})$ to $\mathcal{L}^p(M, \mu, E)$ by replacing E by M.

28.3.3 Weak and strong convergence

The strong topology on the L^p spaces is the normed topology. The weak topology is driven by the topological dual L$^{p'}$ (which is L^q for $\frac{1}{p} + \frac{1}{q} = 1$ for p<∞ see below): a sequence f_n converges weakly in L^p if : $\forall \lambda \in L^{p'} : \lambda(f_n) \to \lambda(f)$. We have several results :

Theorem 2174 *(Lieb p.56) If $f \in L^p(E, S, \mu, \mathbb{C}) : \forall \lambda \in L^p(E, S, \mu, \mathbb{C})' : \lambda(f) = 0$ then f=0. The theorem holds for 1$\leq p < \infty$ and for p=∞ if E is $\sigma-$finite.*

(this is an application of the Hahn Banach theorem)

Theorem 2175 *(Lieb p.57) If $f_n \in L^p(E, S, \mu, \mathbb{C})$ converges weakly to f, then $\liminf_{n\to\infty} \|f_n\|_p = \|f\|_p$. The theorem holds for 1$\leq p < \infty$ and for p=∞ if E is $\sigma-$finite.*

Theorem 2176 *(Lieb p.57) If $f_n \in L^p(E, S, \mu, \mathbb{C})$ is such that $\forall \lambda \in L^p(E, S, \mu, \mathbb{C})$, $|\lambda(f_n)|$ is bounded, then $\exists C > 0 : \|f_n\|_p < C$. The theorem holds for 1$\leq p < \infty$ and for p=∞ if E is $\sigma-$finite.*

Theorem 2177 *(Lieb p.57) If $f_n \in L^p(E, S, \mu, \mathbb{C})$ converges weakly to f then there are $c_{nj} \in [0,1], \sum_{j=1}^n c_{nj} = 1$ such that : $\varphi_n = \sum_{j=1}^n c_{nj} f_j$ converges strongly to f.The theorem holds for 1$\leq p < \infty$*

Theorem 2178 *(Lieb p.68) If O is a measurable subset of $\mathbb{R}^m$, $f_n \in L^p(O, S, dx, \mathbb{C})$ a bounded sequence, then there is a subsequence F_k and $f \in L^p(O, S, dx, \mathbb{C})$ such that F_k converges weakly to f.*

28.3.4 Inequalities

$$(\|f\|_r)^{\frac{1}{p} - \frac{1}{q}} \leq \left(\|f\|_p\right)^{\frac{1}{r} - \frac{1}{p}} \left(\|f\|_q\right)^{\frac{1}{q} - \frac{1}{r}}$$

Theorem 2179 *Hölder's inequality (Lieb p.45) For $1 \leq p \leq \infty, \frac{1}{p} + \frac{1}{q} = 1$*
If $f \in \mathcal{L}^p(E, S, \mu, \mathbb{C}), g \in \mathcal{L}^q(E, S, \mu, \mathbb{C})$ then :
i) $f \times g \in \mathcal{L}^1(E, S, \mu, \mathbb{C})$ and $\|fg\|_1 \leq \int_E |f| |g| \mu \leq \|f\|_p \|g\|_q$
ii) $\|fg\|_1 = \int_E |f| |g| \mu$ iff $\exists \theta \in \mathbb{R}, \theta = Ct : f(x) g(x) = e^{i\theta} |f(x)| |g(x)|$ almost everywhere
iii) if $f \neq 0, \int_E |f| |g| \mu = \|f\|_p \|g\|_q$ iff $\exists \theta \in \mathbb{R}, \theta = Ct$ such that :
if 1<p<∞ : $|g(x)| = \lambda |f(x)|^{p-1}$ almost everywhere
if p=1: $|g(x)| \leq \lambda$ almost everywhere and $|g(x)| = \lambda$ when f(x)$\neq$ 0
if p=∞ : $|f(x)| \leq \lambda$ almost everywhere and $|f(x)| = \lambda$ when g(x)$\neq$ 0

Thus : p=q=2: $f, g \in \mathcal{L}^2(E, S, \mu, \mathbb{C})$ $\|fg\|_1 \leq \|f\|_2 \|g\|_2$

Theorem 2180 *The map:* $\lambda : L^\infty (E, S, \mu, \mathbb{C}) \to \mathcal{L} \left(L^2 (E, S, \mu, \mathbb{C}) ; L^2 (E, S, \mu, \mathbb{C}) \right) ::$ $\lambda (f) g = fg$ *is a continuous operator* $\|\lambda (f)\|_2 \leq \|f\|_\infty$ *and an isomorphism of* C^**-algebra. If* μ *is* $\sigma-finite$ *then* $\|\lambda (f)\|_2 = \|f\|_\infty$

Theorem 2181 *Hanner's inequality (Lieb p.40) For* $f, g \in \mathcal{L}^p (E, S, \mu, \mathbb{C})$
$\forall 1 \leq p \leq \infty : \|f + g\|_p \leq \|f\|_p + \|g\|_p$ *and if 1<p<∞ the equality stands iff*
$\exists \lambda \geq 0 : g = \lambda f$
$1 \leq p \leq 2 :$
$$\|f + g\|_p^p + \|f - g\|_p^p \leq \left(\|f\|_p + \|g\|_p \right)^p + \left(\|f\|_p - \|g\|_p \right)^p$$
$$\left(\|f + g\|_p + \|f - g\|_p \right)^p + \left| \|f + g\|_p - \|f - g\|_p \right|^p \leq 2^p \left(\|f\|_p^p + \|g\|_p^p \right)$$
$2 \leq p < \infty :$
$$\|f + g\|_p^p + \|f - g\|_p^p \geq \left(\|f\|_p + \|g\|_p \right)^p + \left(\|f\|_p - \|g\|_p \right)^p$$
$$\left(\|f + g\|_p + \|f - g\|_p \right)^p + \left| \|f + g\|_p - \|f - g\|_p \right|^p \geq 2^p \left(\|f\|_p^p + \|g\|_p^p \right)$$

Theorem 2182 *(Lieb p.51) For* $1<p<\infty, f, g \in \mathcal{L}^p (E, S, \mu, \mathbb{C})$, *the function*
$:\phi : \mathbb{R} \to \mathbb{R} :: \phi (t) = \int_E |f + tg|^p \mu$ *is differentiable and*
$$\frac{d\phi}{dt}|_{t=0} = \frac{p}{2} \int_E |f|^{p-2} \left(\overline{f} g + f \overline{g} \right) \mu$$

Theorem 2183 *(Lieb p.98) There is a fixed countable family* $(\varphi_i)_{i \in I}, \varphi \in$ $C (\mathbb{R}^m ; \mathbb{C})$ *such that for any open subset O in* $\mathbb{R}^m, 1 \leq p \leq \infty, f \in L^p (O, dx, \mathbb{C}), \varepsilon >$ $0 : \exists i \in I : \|f - \varphi_i\|_p < \varepsilon$

Theorem 2184 *Young's inequalities (Lieb p.98): For p,q,r>1 such that* $\frac{1}{p} +$ $\frac{1}{q} + \frac{1}{r} = 2$,
for any $f \in L^p (\mathbb{R}^m, dx, \mathbb{C}), g \in L^q (\mathbb{R}^m, dx, \mathbb{C}), h \in L^r (\mathbb{R}^m, dx, \mathbb{C})$
$$\left| \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} f (x) g (x - y) h (y) dxdy \right| \leq \sqrt{\frac{p^{1/p}}{p'^{1/9'}}} \|f\|_p \|g\|_q \|h\|_r \text{ with } \frac{1}{p} + \frac{i}{p'} = 1$$

The equality occurs iff each function is gaussian :
$f (x) = A \exp \left(-p' (x - a, J (x - a)) + ik.x \right)$
$g (x) = B \exp \left(-q' (x - b, J (x - b)) + ik.x \right)$
$h (x) = C \exp \left(-r' (x - a, J (x - a)) + ik.x \right)$
with A,B,C $\in \mathbb{C}, J$ *a real symmetric positive definite matrix,a=b+c* $\in \mathbb{R}^m$

Theorem 2185 *Hardy-Litllewood-Sobolev inequality (Lieb p.106) For p,r >1,* $0 < \lambda m$ *such that* $\frac{1}{p} + \frac{\lambda}{m} + \frac{1}{r} = 2$,
there is a constant $C (p, r, \lambda)$ *such that :*
$\forall f \in L^p (\mathbb{R}^m, dx, \mathbb{C}), \forall h \in L^r (\mathbb{R}^m, dx, \mathbb{C})$
$$\left| \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} f (x) \|(x - y)\|^\lambda h (y) dxdy \right| \leq C (p, r, \lambda) \|f\|_p \|h\|_r$$

(Riesz' representation theorem) See below linear functionals for the dual $L^p (E, S, \mu, \mathbb{C})'$

28.3.5 Density theorems

Theorem 2186 *(Neeb p.43) If E is a topological Hausdorff locally compact space and μ a Radon measure, then the set $C_{0c}(E;\mathbb{C})$ is dense in $L^2(E,\mu,\mathbb{C})$.*

Theorem 2187 *(Zuily p.14) If O is an open subset of $\mathbb{R}^m$, μ the Lebesgue measure, then*

the subset $L^p_c(O,\mu,\mathbb{C})$ of $L^p(O,\mu,\mathbb{C})$ of functions with compact support is dense in $L^p(O,\mu,\mathbb{C})$ for $1 \leq p < \infty$

the subset $C_{\infty c}(O,\mu,\mathbb{C})$ of smooth functions with compact support is dense in $L^p(O,\mu,\mathbb{C})$ for $1 \leq p < \infty$

Theorem 2188 *If E is a Hausdorff locally compact space with its Borel algebra S, P is a Borel, inner regular, probability, then any function $f \in L^p(E,S,P,\mathbb{C})$, $1 \leq p < \infty$, is almost everywhere equal to a function in $C_b(E;\mathbb{C})$ and there is a sequence (f_n), $f_n \in C_c(E;\mathbb{C})$ which converges to f in $L^p(E,S,P,\mathbb{C})$ almost everywhere.*

28.3.6 Integral operators

Integral operators are linear maps : $K \in \mathcal{L}(L^p(E,S,\mu,\mathbb{C});L^q(E,S,\mu,\mathbb{C}))$

Theorem 2189 *(Minkowski's inequality) (Lieb p.47) Let (E,S,μ), (F,S',ν) $\sigma-$finite measured spaces with positive measure,*

$f : E \times F \to \mathbb{R}_+$ a measurable function,

$1 \leq p < \infty : \int_F \left(\int_E f(x,y)^p \mu(x)\right)^{1/p} \nu(y) \geq \left(\int_E \left(\int_F f(x,y)\nu(y)\right)^p \mu(x)\right)^{1/p}$
and if one term is $<\infty$ so is the other. If the equality stands and p>1 then there are measurable functions $u \in C(E;\mathbb{R}_+)$, $v \in C(F;\mathbb{R}_+)$ such that $f(x,y) = u(x)v(y)$ almost everywhere

Integral kernel

Theorem 2190 *(Taylor 2 p.16) Let (E,S,μ) a measured space with μ a positive measure, $k : E \times E \to \mathbb{C}$ a measurable function on ExE such that : $\exists C_1, C_2 \in \mathbb{R} : \forall x,y \in E : \int_E |k(t,y)|\mu(t) \leq C_1, \int_E |k(x,t)|\mu(t) \leq C_2$*
*then $\forall \infty \geq p \geq 1$, $K : L^p(E,S,\mu,\mathbb{C}) \to L^q(E,S,\mu,\mathbb{C}) :: K(f)(x) = \int_E k(x,y)f(y)\mu(y)$ is a linear continuous operator called the **integral kernel** of K*

$K \in \mathcal{L}(L^p(E,S,\mu,\mathbb{C});L^q(E,S,\mu,\mathbb{C}))$ with : $\frac{1}{p} + \frac{1}{q} = 1$, $\|K\| \leq C_1^{1/p}C_2^{1/q}$
The transpose K^t of K is the operator with integral kernel k^t(x,y)=k(y,x).
If p=2 the adjoint K* of K is the integral operator with kernel k*(x,y)=$\overline{k(y,x)}$

Hilbert-Schmidt operator

(see Hilbert spaces for the definition)

Theorem 2191 *(Taylor 1 p.500) Let $(E_1, S_1, \mu_1), (E_2, S_2, \mu_2)$ be measured spaces with positive measures, and $T \in \mathcal{L}\left(L^2(E_1, S_1, \mu_1, \mathbb{C}); L^2(E_2, S_2, \mu_2, \mathbb{C})\right)$ be a Hilbert-Schmidt operator, then there is a function $K \in L^2\left(E_1 \times E_2, \mu_1 \otimes \mu_2, \mathbb{C}\right)$ such that :*

$\langle Tf, g \rangle = \int \int K\left(x_1, x_2\right) \overline{f\left(x_1\right)} g\left(x_2\right) \mu_1\left(x_1\right) \mu_2\left(x_2\right)$
and we have $\|T\|_{HS} = \|K\|_{L^2}$

Theorem 2192 *(Taylor 1 p.500) Let K_1, K_2 two Hilbert-Schmidt, integral operators on $L^2\left(E, S, \mu, \mathbb{C}\right)$ with kernels k_1, k_2. Then the product $K_1 \circ K_2$ is an Hilbert-Schmidt integral operator with kernel :* $k\left(x, y\right) = \int_E k_1\left(x, t\right) k_2\left(t, y\right) \mu\left(t\right)$

28.3.7 Convolution

Convolution is defined in the Lie group part (see integration). The convolution is a map on functions defined on a locally compact topological unimodular group G :

$*{:}L^1\left(G, S, \mu, \mathbb{C}\right) \times L^1\left(G, S, \mu, \mathbb{C}\right) \to L^1\left(G, S, \mu, \mathbb{C}\right) :: \varphi * \psi\left(g\right) = \int_G \varphi\left(x\right) \psi\left(x^{-1}g\right) \mu\left(x\right) = \int_G \varphi\left(gx\right) \psi\left(x^{-1}\right) \mu\left(x\right)$

The involution is defined as : $\varphi^*\left(g\right) = \overline{\varphi\left(g^{-1}\right)}$

All the results presented in the Lie group part can be extended to functions on the abelian group $\left(\mathbb{R}^m, +\right)$ endowed with the Lebesgue measure.

Definition

Definition 2193 *The **convolution** is the map :*

$*{:}L^1\left(\mathbb{R}^m, dx, \mathbb{C}\right) \times L^1\left(\mathbb{R}^m, dx, \mathbb{C}\right) \to L^1\left(\mathbb{R}^m, dx, \mathbb{C}\right) :: \left(f * g\right)\left(x\right) = \int_{\mathbb{R}^m} f\left(y\right) g\left(x - y\right) dy = \int_{\mathbb{R}^m} f\left(x - y\right) g\left(x\right) dy$

Convolution is well defined for other spaces of functions.

Whenever a function is defined on an open $O \subset \mathbb{R}^m$ it can be extended by taking $\widetilde{f}\left(x\right) = f\left(x\right), x \in O, \widetilde{f}\left(x\right) = 0, x \notin O$ so that many results still stand for functions defined in O.

Theorem 2194 *Hörmander : The convolution f*g exists if f$\in C_c\left(\mathbb{R}^m; \mathbb{C}\right)$ and g$\in L^1_{loc}\left(\mathbb{R}^m, dx, \mathbb{C}\right)$, then f*g is continuous*

Theorem 2195 *The convolution f*g exists if f,g$\in S\left(\mathbb{R}^m\right)$ (Schwartz functions) then f*g$\in S\left(\mathbb{R}^m\right)$*

Theorem 2196 *The convolution f*g exists if $f \in L^p(\mathbb{R}^m, dx, \mathbb{C})$, $g \in L^q(\mathbb{R}^m, dx, \mathbb{C})$, $1 \leq p, q \leq \infty$ then the convolution is a continuous bilinear map :*

$* \in \mathcal{L}^2(L^p(\mathbb{R}^m, dx, \mathbb{C}) \times L^q(\mathbb{R}^m, dx, \mathbb{C}); L^r(\mathbb{R}^m, dx, \mathbb{C}))$ *with* $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$

and $\|f * g\|_r \leq \|f\|_p \|g\|_q$

if $\frac{1}{p} + \frac{1}{q} = 1$ *then* $f * g \in C_{0\nu}(\mathbb{R}^m, \mathbb{C})$ *(Lieb p.70)*

This is a consequence of the Young's inequality

Definition 2197 *On the space of functions $C(\mathbb{R}^m; \mathbb{C})$*

*the **involution** is :* $f^*(x) = \overline{f(-x)}$

*the **translation** is :* $\tau_a f(x) = f(x - a)$ *for* $a \in \mathbb{R}$

So the left action is $\Lambda(a) f(x) = \tau_a f(x) = f(x - a)$ and the right action is : $P(a) f(x) = \tau_{-a} f(x) = f(x + a)$

Properties

Theorem 2198 *Whenever the convolution is defined on a vector space V of functions, it makes V a commutative complex algebra without identity, and convolution commutes with translation*

$f * g = g * f$

$(f * g) * h = f * (g * h)$

$f * (ag + bh) = af * g + bf * h$

$\tau_a(f * g) = f * \tau_a g = \tau_a f * g$ with $\tau_a : V \to V :: (\tau_a f)(x) = f(x - a)$

Theorem 2199 *Whenever the convolution of f,g is defined:*

$Supp(f * g) \subset Supp(f) \cap Supp(g)$

$(f * g)^* = f^* * g^*$

If f,g are integrable, then : $\int_{\mathbb{R}^m} (f * g) \, dx = \left(\int_{\mathbb{R}^m} f(x) \, dy \right) \left(\int_{\mathbb{R}^m} g(x) \, dx \right)$

Theorem 2200 *With convolution as internal operation and involution, $L^1(\mathbb{R}^m, dx, \mathbb{C})$ is a commutative Banach *-algebra and the involution, right and left actions are isometries.*

If $f, g \in L^1(\mathbb{R}^m, dx, \mathbb{C})$ and f or g has a derivative which is in $L^1(\mathbb{R}^m, dx, \mathbb{C})$ then f*g is differentiable and :

$\frac{\partial}{\partial x_\alpha}(f * g) = \left(\frac{\partial}{\partial x_\alpha} f \right) * g = f * \frac{\partial}{\partial x_\alpha} g$

$\forall f, g \in L^1(\mathbb{R}^m, dx, \mathbb{C}) : \|f * g\|_1 \leq \|f\|_1 \|g\|_1, \|f^*\|_1 = \|f\|_1, \|\tau_a f\|_1 = \|f\|_1$

Approximation of the identity

The absence of an identity element is compensated with an approximation of the identity defined as follows :

Definition 2201 *An **approximation of the identity** in the Banach *-algebra* $\left(L^1\left(\mathbb{R}^m, dx, \mathbb{C}\right), *\right)$ *is a family of functions* $(\rho_\varepsilon)_{\varepsilon \in \mathbb{R}}$ *with* $\rho_\varepsilon = \varepsilon^{-m}\rho\left(\frac{x}{\varepsilon}\right) \in C_{\infty c}\left(\mathbb{R}^m; \mathbb{C}\right)$ *where* $\rho \in C_{\infty c}\left(\mathbb{R}^m; \mathbb{C}\right)$ *such that :* $Sup(\rho) \subset B\left(0, 1\right), \rho \geq 0, \int_{\mathbb{R}^m} \rho\left(x\right) dx = 1$

Theorem 2202 *(Lieb p.76) The family* ρ_ε *has the properties :*

$\forall f \in C_{rc}\left(\mathbb{R}^m; \mathbb{C}\right), \forall \alpha = (\alpha_1, ...\alpha_s), s \leq r : D_{(\alpha)}\left(\rho_n * f\right) \to D_{(\alpha)}f$ *uniformly when* $\varepsilon \to 0$

$\forall f \in L_c^p\left(\mathbb{R}^m, dx, \mathbb{C}\right), 1 \leq p < \infty, \rho_\varepsilon * f \to f$ *when* $\varepsilon \to 0$

These functions are used to approximate any function f$\in L^1\left(\mathbb{R}^m, dx, \mathbb{C}\right)$ by a sequence of functions $\in C_{\infty c}\left(\mathbb{R}^m; \mathbb{C}\right)$

28.4 Spaces of differentiable maps

The big difference is that these spaces are usually not normable. This can be easily understood : the more regular the functions, the more difficult it is to get the same properties for limits of sequences of these functions. Differentiability is only defined for functions over manifolds, which can be open subsets of $\mathbb{R}^m$.Notice that we address only differentiability with respect to real variables, over real manifolds. Indeed complex differentiable functions are holomorphic, and thus C-analytic, so for most of the purposes of functional analysis, it is this feature which is the most useful.

We start by the spaces of differentiable sections of a vector bundle, as the case of functions follows.

Notation 2203 $D_{(\alpha)} = D_{\alpha_1...\alpha_s} = \dfrac{\partial}{\partial\xi^{\alpha_1}}\dfrac{\partial}{\partial\xi^{\alpha_2}}...\dfrac{\partial}{\partial\xi^{\alpha_s}}$ *where the* $\alpha_k = 1...m$ *can be identical*

28.4.1 Spaces of differentiable sections of a vector bundle

Theorem 2204 *The spaces* $\mathfrak{X}_r\left(E\right)$ *of r (1$\leq r \leq \infty$) continuously differentiable sections on a vector bundle E are Fréchet spaces.*

Proof. Let $E\left(M, V, \pi\right)$ be a complex vector bundle with a finite atlas $(O_a, \varphi_a)_{a=1}^N$, such that $(F, O_a, \psi_a)_{a=1}^N$ is an atlas of the m dimensional manifold M.

$S \in \mathfrak{X}_r\left(E\right)$ is defined by a family of r differentiable maps : $u_a \in C_r\left(O_a; V\right) :: S\left(x\right) = \varphi_a\left(x, u_a\left(x\right)\right)$. Let be $\Omega_a = \psi_a\left(O_a\right) \subset \mathbb{R}^m$. $F_a = u_a \circ \psi_a^{-1} \in C_r\left(\Omega_a; V\right)$ and $D_{\alpha_1...\alpha_s}F_a(\xi)$ is a continuous multilinear map $\mathcal{L}^s\left(\mathbb{R}^m; V\right)$ with a finite norm $\|D_{\alpha_1...\alpha_s}F_a(\xi)\|$

For each set Ω_a we define the sequence of sets :

$K_p = \{\xi \in \Omega_a, \|\xi\| \leq p\} \cap \{\xi \in \Omega_a, d\left(\xi, \Omega_a^c\right) \geq 1/p\}$

It is easily shown (Zuily p.2) that : each K_p is compact, $K_p \subset \overset{\circ}{K}_{p+1}, \Omega_a = \cup_{p=1}^\infty K_p = \cup_{p=2}^\infty \overset{\circ}{K}_p$, for each compact K$\subset \Omega_a$ there is some p such that K$\subset \Omega_p$

The maps p_n :

644

for $1 \leq r < \infty : p_n(S) = \max_{a=..n} \sum_{s=1}^{n} \sum_{\alpha_1...\alpha_s} \sup_{\xi \in K_n} \|D_{\alpha_1...\alpha_s} F_a(\xi)\|$

for $r = \infty : p_n(S) = \max_{a=..n} \sum_{s=1}^{n} \sum_{\alpha_1...\alpha_n} \sup_{\xi \in K_n} \|D_{\alpha_1...\alpha_s} F_a(\xi)\|$

define a countable family of semi-norms on $\mathfrak{X}_r(E)$.

With the topology induced by these semi norms a sequence $S_n \in \mathfrak{X}_r(E)$ converges to $S \in \mathfrak{X}_r(E)$ iff S converges uniformly on any compact. The space of continuous, compactly supported sections is a Banach space, so any Cauchy sequence on K_p converges, and converges on any compact. Thus $\mathfrak{X}_r(E)$ is complete with these semi norms. ∎

A subset A of $\mathfrak{X}_r(E)$ is bounded if : $\forall S \in A, \forall n > 1, \exists C_n : p_n(S) \leq C_n$

Theorem 2205 *The space $\mathfrak{X}_{rc}(E)$ of r differentiable, compactly supported, sections of a vector bundle is a Fréchet space*

This is a closed subspace of $\mathfrak{X}_r(E)$. See general theorem above.

28.4.2 Space of sections of the r jet prolongation of a vector bundle

Theorem 2206 *The space $\mathfrak{X}(J^r E)$ of sections of the r jet prolongation $J^r E$ of a vector bundle is a Fréchet space.*

Proof. If E is a vector bundle on a m dimensional real manifold M and dimV=n then $J^r E$ is a vector bundle $J^r E\left(E, J_0^r(\mathbb{R}^m, V), \pi_0^r\right)$ endowed with norms fiberwise (see above). A section $j^r z$ over $J^r E$ is a map $M \to J^r E$ with coordinates :

$\left(\xi^\alpha, \eta^i(\xi), \eta^i_{\alpha_1..\alpha_s}(\xi), s = 1...r, 1 \leq \alpha_k \leq \alpha_{k+1} \leq m, i = 1..n\right)$

The previous semi-norms read :

for $1 \leq s \leq r : p_n(j^r z) = \max_{a=..n} \sum_{s=1}^{n} \sum_{\alpha_1...\alpha_s} \sup_{\xi \in K_n} \|\eta^i_{\alpha_1..\alpha_s}(\xi)\|$ ∎

A section U on E gives rise to a section on $J^r E : j^r U(x) = j^r_x U$ and with this definition : $p_n(j^r U)_{\mathfrak{X}(J^r E)} = p_n(U)_{\mathfrak{X}_r(E)}$

28.4.3 Spaces of differentiable functions on a manifold

The spaces of differentiable functions on a manifold are a special case of the previous ones.

Definition 2207 $C_r(M;\mathbb{C})$ *the space of r continuously differentiable functions* $f : M \to \mathbb{C}$

$C_{rc}(M;\mathbb{C})$ *the space of r continuously differentiable functions* $f : M \to \mathbb{C}$ *with compact support*

if r=∞ the functions are smooth

Theorem 2208 *The spaces $C_r(M;\mathbb{C}), C_{rc}(M;\mathbb{C})$ are Fréchet spaces*

If M is a *compact* finite dimensional Hausdorff smooth manifold we have simpler semi-norms:

Theorem 2209 *If M is a compact manifold, then :*

i) the space $C_\infty(M; \mathbb{C})$ of smooth functions on M is a Fréchet space with the family of seminorms :

for $1 \leq r < \infty : p_r(f) = \sum_{p=0}^{r} \left| f^{(p)}(p)(u_1, ..., u_p) \right|_{\|u_k\| \leq 1}$

ii) the space $C_r(M; \mathbb{C})$ of r differentiable functions on M is a Banach space with the norm :

$\|f\| = \sum_{p=0}^{r} \left| f^{(p)}(p)(u_1, ..., u_p) \right|_{\|u_k\| \leq 1}$

where the u_k are vectors fields whose norm is measured with respect to a riemannian metric

(which always exist with our assumptions).

Notice that we have no Banach spaces structures for differentiable functions over a manifold if it is not compact.

Theorem 2210 *(Zuily p.2) A subset A of $C_\infty(O; \mathbb{C})$ is compact iff it is closed and bounded with the semi-norms*

this is untrue if r<∞.

28.4.4 Spaces of compactly supported, differentiable functions on an open of $\mathbb{R}^m$

There is an increasing sequence of relatively compacts open O_n which covers any open O. Thus with the notation above $C_{rc}(O; \mathbb{C}) = \sqcup_{i=1}^{\infty} C_r(O_n; \mathbb{C})$. Each of the $C_r(O_n; \mathbb{C})$ endowed with the seminorms is a Fréchet space (and a Banach), $C_r(O_n; \mathbb{C}) \subset C_r(O_{n+1}; \mathbb{C})$ and the latter induces in the former the same topology.

Which entails :

Theorem 2211 *(Zuily p.10) For the space $C_{rc}(O; \mathbb{C})$ of compactly supported, r continuously differentiable functions on an open subset of $\mathbb{R}^m$*

i) there is a unique topology on $C_{rc}(O; \mathbb{C})$ which induces on each $C_r(O_n; \mathbb{C})$ the topology given by the seminorms

ii) A sequence (f_n) in $C_{rc}(O; \mathbb{C})$ converges iff : $\exists N : \forall n : Supp(f_n) \subset O_N$, and (f_n) converges in $C_{rc}(O_N; \mathbb{C})$

iii) A linear functional on $C_{rc}(O; \mathbb{C})$ is continuous iff it is continuous on each $C_{rc}(O_n; \mathbb{C})$

iv) A subset A of $C_{rc}(O; \mathbb{C})$ is bounded iff there is N such that $A \subset C_{rc}(O_N; \mathbb{C})$ and A is bounded on this subspace.

v) For $0 \leq r \leq \infty$ $C_{rc}(O; \mathbb{C})$ is dense in $C_r(O; \mathbb{C})$

So for most of the purposes it is equivalent to consider $C_{rc}(O; \mathbb{C})$ or the families $C_{rc}(K; \mathbb{C})$ where K is a relatively compact open in O.

Theorem 2212 *(Zuily p.18) Any function of $C_{rc}(O_1 \times O_2; \mathbb{C})$ is the limit of a sequence in $C_{rc}(O_1; \mathbb{C}) \otimes C_{rc}(O_2; \mathbb{C})$.*

Functions $f \in C_{\infty c}(\mathbb{R}^m; \mathbb{C})$ can be deduced from holomorphic functions.

Theorem 2213 *Paley-Wiener-Schwartz : For any function $f \in C_{\infty c}(\mathbb{R}^m; \mathbb{C})$ with support in the ball B(0,r), there is a holomorphic function $F \in H(\mathbb{C}^m; \mathbb{C})$ such that :*

$\forall x \in \mathbb{R}^m : f(x) = F(x)$

$\forall p \in \mathbb{N}, \exists c_p > 0, \forall z \in \mathbb{C}^m : |F(z)| \leq c_p(1 + |z|)^{-p} e^{r|\text{Im } z|}$

Conversely if F is a holomorphic function $F \in H(\mathbb{C}^m; \mathbb{C})$ meeting the property above, then there is a function $f \in C_{\infty,c}(\mathbb{R}^m; \mathbb{C})$ with support in the ball B(0,r) such that $\forall x \in \mathbb{R}^m : f(x) = F(x)$

28.4.5 Space of Schwartz functions

This is a convenient space for Fourier transform. It is defined only for functions over the whole of $\mathbb{R}^m$.

Definition 2214 *The **Schwartz space** $S(\mathbb{R}^m)$ of rapidly decreasing smooth functions is the subspace of $C_\infty(\mathbb{R}^m; \mathbb{C})$:*

$f \in C_\infty(\mathbb{R}^m; \mathbb{C}) : \forall n \in \mathbb{N}, \forall \alpha = (\alpha_1, .., \alpha_n), \exists C_{n,\alpha} : \|D_{\alpha_1...\alpha_n} f(x)\| \leq C_{N,\alpha} \|x\|^{-n}$

Theorem 2215 *(Zuily p.108) The space $S(\mathbb{R}^m)$ is a Fréchet space with the seminorms: $p_n(f) = \sup_{x \in \mathbb{R}^m, \alpha, k \leq n} \|x\|^k \|D_{\alpha_1...\alpha_n} f(x)\|$ and a complex commutative *-algebra with pointwise multiplication*

Theorem 2216 *The product of $f \in S(\mathbb{R}^m)$ by any polynomial, and any partial derivative of f still belongs to $S(\mathbb{R}^m)$*

Theorem 2217 $C_{\infty c}(\mathbb{R}^m; \mathbb{C}) \subset S(\mathbb{R}^m) \subset C_\infty(\mathbb{R}^m; \mathbb{C})$

$C_{\infty c}(\mathbb{R}^m; \mathbb{C})$ *is dense in* $S(\mathbb{R}^n)$

$\forall p : 1 \leq p \leq \infty : S(\mathbb{R}^n) \subset L^p(\mathbb{R}^m, dx, \mathbb{C})$

28.4.6 Sobolev spaces

Sobolev spaces consider functions which are both integrable and differentiable : there are a combination of the previous cases. To be differentiable they must be defined at least over a manifold M, and to be integrable we shall have some measure on M. So the basic combination is a finite m dimensional smooth manifold endowed with a volume form ϖ_0.

Usually the domain of Sobololev spaces of functions are limited to an open of $\mathbb{R}^m$ but with the previous result we can give a more extensive definition which is useful for differential operators.

Sobolev spaces are extended (see Distributions and Fourier transform).

Sobolev space of sections of a vector bundle

Definition 2218 *On a vector bundle $E(M, V, \pi)$,with M endowed with a positive measure μ, the Sobolev space denoted $W^{r,p}(E)$ is the subspace of r differentiable sections $S \in \mathfrak{X}_r(E)$ such that their r jet prolongation: $J^r S \in L^p(M, \mu, J^r E)$*

Theorem 2219 *$W^{r,p}(E)$ is a Banach space with the norm :*

$\|Z\|_{p,r} = \left(\sum_{s=1}^{r} \sum_{\alpha_1...\alpha_s} \int_M \|D_{(\alpha)} u(x)\|^p \mu \right)^{1/p}$. *Moreover the map : J^r : $W^{r,p}(E) \to L^p(M, \mu, J^r E)$ is an isometry.*

Proof. $W^{r,p}(E) \subset L^p(M, \varpi_0, E)$ which is a Banach space.

Each section $S \in \mathfrak{X}_r(E)$ $S(x) = \varphi(x, u(x))$ gives a section in $\mathfrak{X}(J^r S)$ which reads :

$M \to V \times \prod_{s=1}^{r} \left\{ \mathcal{L}_S^s \left(\mathbb{R}^{\dim M}, V \right) \right\} :: Z(s) = (z_s(x), s = 0...r)$ with $z_s(x) = D_{(\alpha)} u(x)$ with, fiberwise, norms $\|z_s(x)\|_s$ for each component and $\mathcal{L}^p(M, \mu, J^r E) = \left\{ Z \in \mathfrak{X}(J^r E) : \sum_{s=0}^{r} \int_M \|z_s(x)\|_s^p \mu < \infty \right\}$

Take a sequence $S_n \in W^{r,p}(E)$ which converges to S$\in \mathfrak{X}_r(E)$ in $\mathfrak{X}_r(E)$, then the support of each S_n is surely ϖ_0 integrable and by the Lebesgue theorem $\int_M \|D_{(\alpha)} \sigma_n(x)\|_s^p \mu < \infty$. Thus $W^{r,p}(E)$ is closed in $L^p(M, \mu, J^r E)$ and is a Banach space. ∎

Theorem 2220 *On a vector bundle endowed with an inner product $W^{r,2}(E)$ is a Hilbert space, with the scalar product of $L^2(M, \mu, E)$*

Proof. $W^{r,2}(E)$ is a vector subspace of $L^2(M, \varpi_0, E)$ and the map J^r : $W^{r,2}(E) \to L^2(M, \varpi_0, J^r E)$ is continuous. So $W^{r,2}(E)$ is closed in $L^2(M, \varpi_0, E)$ which is a Hilbert space ∎

As usual :

Notation 2221 *$W_c^{r,p}(E)$ is the subspace of $W^{r,p}(E)$ of sections with compact support.*

$W_{loc}^{r,p}(E)$ is the subspace of $L_{loc}^p(M, \varpi_0, E)$ comprised of r differentiable sections such that : $\forall \alpha, \|\alpha\| \leq r, D_\alpha \sigma \in L_{loc}^p(M, \varpi_0, E)$

Sobolev spaces of functions over a manifold

This is a special case of the previous one.

Definition 2222 *The **Sobolev space,** denoted $W^{r,p}(M)$, of functions over a manifold endowed with a positive measure μ is the subspace of $L^p(M, \mu, \mathbb{C})$ comprised of r differentiable functions f over M such that : $\forall \alpha_1, ...\alpha_s, s \leq r :$ $D_{(\alpha)} f \in L^p(M, \mu, \mathbb{C}), 1 \leq p \leq \infty$*

Theorem 2223 $W^{r,p}(M)$ *It is a Banach vector space with the norm* $:\|f\|_{W^{r,p}} = \left(\sum_{s=1}^{r}\sum_{\alpha_1..\alpha_s}\left\|D_{(\alpha)}f\right\|_{L^p}\right)^{1/p}$.

Theorem 2224 *If there is an inner product in E,* $W^{r,2}(M) = H^r(M)$ *is a Hilbert space with the scalar product :* $\langle\varphi,\psi\rangle = \sum_{s=1}^{r}\sum_{\alpha_1..\alpha_s}\left\langle D_{(\alpha)}\varphi,D_\alpha\psi\right\rangle_{L^2}$

Notation 2225 $H^r(M)$ *is the usual notation for* $W^{r,2}(M)$

Remark : this is the usual definition of Sobolev spaces. As we can see we need a norm for the derivatives, which are multilinear maps : $\mathcal{L}^s(T_xM;\mathbb{C})$ whose norm is not obviously defined from a metric on M, but more obviously by using the fact that $\mathcal{L}^s(T_xM;\mathbb{C})\simeq\mathcal{L}^s(\mathbb{R}^m;\mathbb{C})$.

Sobolev spaces of functions over $\mathbb{R}^m$

Theorem 2226 *For any open subset O of $\mathbb{R}^m$:*
i) $\forall r > r'$ *:* $H^r(O) \subset H^{r'}(O)$ *and if O is bounded then the injection* $\imath : H^r_c(O) \to H^{r'}_c(O)$ *is compact*
ii) $C_\infty(O;\mathbb{C})$ *is dense in* $W^{1,1}_{loc}(O)$ *and* $H^1(O)$

Theorem 2227 $C_{\infty c}(\mathbb{R}^m;\mathbb{C})$ *is dense in* $H^r(\mathbb{R}^m)$ *(but this is no true for* $O\subset\mathbb{R}^m$*)*

Theorem 2228 *(Lieb p.177) For any functions* $u,v\in H^1(\mathbb{R}^m)$ *:* $\forall k = 1..m$ *:* $\int_{\mathbb{R}^m}u\left(\partial_k v\right)dx = -\int_{\mathbb{R}^m}v\left(\partial_k u\right)dx$

If v is real and $\Delta v \in L^1_{loc}(\mathbb{R}^m,dx,\mathbb{R})$ then $\int_{\mathbb{R}^m}u\left(\Delta v\right)dx = -\int_{\mathbb{R}^m}\sum_{k=1}^{m}\left(\partial_k v\right)\left(\partial_k u\right)dx$

Theorem 2229 *(Lieb p.179) For any real valued functions* $f,g\in H^1(\mathbb{R}^m)$ *:*
$\int_{\mathbb{R}^m}\sum_\alpha\left(\partial_\alpha\sqrt{f^2+g^2}\right)^2 dx \leq \int_{\mathbb{R}^m}\sum_\alpha\left(\left(\partial_\alpha f\right)^2+\left(\partial_\alpha g\right)^2\right)dx$
and if $g\geq 0$ *then the equality holds iff* $f = cg$ *almost everywhere for a constant* c

Theorem 2230 *(Zuilly p.148) : For any class r manifold with boundary M in* $\mathbb{R}^m$ *:*
i) $C_{\infty c}(M;\mathbb{C})$ *is dense in* $H^r\left(\overset{\circ}{M}\right)$

ii) There is a map : $P \in \mathcal{L}\left(H^r\left(\overset{\circ}{M}\right);H^r(\mathbb{R}^m)\right)$ *such* $P(f(x)) = f(x)$ *on* $\overset{\circ}{M}$

iii) If $k > \frac{m}{2} + l$ *:* $H^k\left(\overset{\circ}{M}\right) \subset C_l(M;\mathbb{C})$ *so* $\cup_{k\in\mathbb{N}}H^k\left(\overset{\circ}{M}\right) \subset C_\infty(M;\mathbb{C})$

iv) $\forall r > r'$ *:* $H^r\left(\overset{\circ}{M}\right) \subset H^{r'}\left(\overset{\circ}{M}\right)$ *and if M is bounded then the injection* $\imath : H^r\left(\overset{\circ}{M}\right) \to H^{r'}\left(\overset{\circ}{M}\right)$ *is compact*

28.5   DISTRIBUTIONS

28.6   Spaces of functionals

28.6.1   Reminder of elements of normed algebras

see the section Normed Algebras in the Analysis part.

Let V be a topological *-subalgebra of $C\left(E;\mathbb{C}\right),$ with the involution $f \to \overline{f}$ and pointwise multiplication.

The subset of functions of V valued in $\mathbb{R}$ is a subalgebra of hermitian elements, denoted V_R, which is also the real subspace of V with the natural real structure on V : $f = \operatorname{Re} f + i \operatorname{Im} f$. The space V_+ of positive elements of V is the subspace of V_R of maps in $C\left(E : \mathbb{R}_+\right).$ If V is in $C\left(E;\mathbb{R}\right)$ then $V = V_R$ and is hermitian.

A linear functional is just an element of the algebraic dual $V^* = L(V;\mathbb{C})$ so this is a complex linear map. With the real structures on V and $\mathbb{C}$, any functional reads : $\lambda\left(\operatorname{Re} f + i \operatorname{Im} f\right) = \lambda_R\left(\operatorname{Re} f\right) + \lambda_I\left(\operatorname{Im} f\right) + i\left(-\lambda_I\left(\operatorname{Re} f\right) + \lambda_R\left(\operatorname{Im} f\right)\right)$ with two real functionals $\lambda_R, \lambda_I \in C\left(V_R;\mathbb{R}\right).$ In the language of algebras, λ is hermitian if $\lambda\left(\overline{f}\right) = \overline{\lambda\left(f\right)}$ that is $\lambda_I = 0$. Then $\forall f \in V_R : \lambda\left(f\right) \in \mathbb{R}$.

In the language of normed algebras a linear functional λ is weakly continuous if $\forall f \in V_R$ the map $g \in V \to \lambda\left(\left|g\right|^2 f\right)$ is continuous. As the map : $f, g \in V \to \left|g\right|^2 f$ is continuous on V (which is a *-topological algebra) then here weakly continuous = continuous on V_R

28.6.2   Positive linear functionals

Theorem 2231 *A linear functional* $\lambda : V \to \mathbb{C}$ *on a space of functions V is positive iff* $\lambda\left(\overline{f}\right) = \overline{\lambda\left(f\right)}$ *and* $\lambda\left(f\right) \geq 0$ *when* $f \geq 0$

Proof. Indeed a linear functional λ is positive (in the meaning viewed in Algebras) if $\lambda\left(\left|f\right|^2\right) \geq 0$ and any positive function has a square root. ∎

The variation of a positive linear functional is : $v\left(\lambda\right) = \inf_{f \in V}\left\{\gamma : \left|\lambda\left(f\right)\right|^2 \leq \gamma\lambda\left(\left|f\right|^2\right)\right\}.$ If it is finite then $\left|\lambda\left(f\right)\right|^2 \leq v\left(\lambda\right)\lambda\left(\left|f\right|^2\right)$

A positive linear functional λ, continuous on V_R, is a state if $v\left(\lambda\right) = 1$, a quasi-state if $v\left(\lambda\right) \leq 1$. The set of states and of quasi-states are convex.

If V is a normed *-algebra :

i) a quasi-state is continuous on V with norm $\|\lambda\| \leq 1$

Proof. It is $\sigma-$contractive, so $\left|\lambda\left(f\right)\right| \leq r_\lambda\left(f\right) = \|f\|$ because f is normal ∎

ii) the variation of a positive linear functional is $v\left(\lambda\right) = \lambda\left(I\right)$ where I is the identity element if V is unital.

iii) if V has a state the set of states has an extreme point (a pure state).

If V is a Banach *-algebra :

i) a positive linear functional λ is continuous on V_R. If $v\left(\lambda\right) < \infty$ it is continuous on V, if $v\left(\lambda\right) = 1$ it is a state.

ii) a state (resp. a pure state) λ on a closed *-subalgebra can be extended to a state (resp. a pure state) on V

f V is a C*-algebra : a positive linear functional is continuous and $v(\lambda) = \|\lambda\|$, it is a state iff $\|\lambda\| = \lambda(I) = 1$

As a consequence:

Theorem 2232 *A positive functional is continuous on the following spaces :*
i) $C_b(E;\mathbb{C})$ of bounded functions if E is topological
ii) $C_{0b}(E;\mathbb{C})$ of bounded continuous functions if E Hausdorff
iii) $C_c(E;\mathbb{C})$ of functions with compact support, $C_{0v}(E;\mathbb{C})$ of continuous functions vanishing at infinity, if E is Hausdorff, locally compact
iv) $C_0(E;\mathbb{C})$ of continuous functions if E is compact

Theorem 2233 *If E is a locally compact, separable, metric space, then a positive functional $\lambda \in L(C_{0c}(E;\mathbb{R});\mathbb{R})$ can be uniquely extended to a functional in $\mathcal{L}(C_{0\nu}(E;\mathbb{R});\mathbb{R})$*

28.6.3 Functional defined as integral

Functionals, and mainly positive functional, can be defined as integral. The results vary according to the spaces of functions.

Function defined on a compact space

Theorem 2234 *(Taylor 1 p.484) If E is a compact metric space, $C(E;\mathbb{C})'$ is isometrically isomorphic to the space of complex measures on E endowed with the total variation norm.*

$\mathbf{L}^p$ spaces

Theorem 2235 *(Lieb p.61) For any measured space with a positive measure (E,S,μ) , $1 \le p \le \infty, \frac{1}{p} + \frac{1}{q} = 1$*
The map : $L_p : L^q(E,S,\mu,\mathbb{C}) \to L^p(E,S,\mu,\mathbb{C})^* :: L(f)(\varphi) = \int_E f\varphi\mu$ is continuous and an isometry : $\|L_p(f)\|_p = \|f\|_q$ so $L_p(f) \in L^p(E,S,\mu,\mathbb{C})'$
i) If $1 < p < \infty$, or if p=1 and the measure μ is σ-finite (meaning E is the countable union of subsets of finite measure), then L_p is bijective in the topological dual $L^p(E,S,\mu,\mathbb{C})'$ of $L^p(E,S,\mu,\mathbb{C})$, which is isomorphic to $L^q(E,S,\mu,\mathbb{C})$
ii) If $p=\infty$: elements of the dual $L^\infty(E,S,\mu,\mathbb{C})'$ can be identified with bounded signed finitely additive measures on S that are absolutely continuous with respect to μ.

So :
if f is real then $L_p(f)$ is hermitian, and positive if $f \ge 0$ almost everywhere for $1 < p < \infty$
Conversely:

Theorem 2236 *(Doob p.153) For any measured space with a $\sigma-$finite measure (E,S,μ), and any continuous linear functional on $L^1(E,S,\mu,\mathbb{C})$ there is a function f, unique up to a set of null measure, bounded and integrable such that : $\ell(\varphi) = \int_E f\varphi\mu$. Moreover : $\|\ell\| = \|f\|_\infty$ and ℓ is positive iff $f\geq 0$ almost everywhere.*

Radon measures

see Analysis-Measure,

A Radon measure μ is a Borel (defined on the Borel $\sigma-$algebra S of open subsets), locally finite, regular, signed measure on a topological Hausdorff locally compact space (E,Ω). So :

$\forall X \in S : \mu(X) = \inf(\mu(Y), X \sqsubseteq Y, Y \in \Omega)$

$\forall X \in S, \mu(X) < \infty : \mu(X) = \sup(\mu(K), K \sqsubseteq X, K \text{ compact})$

A measure is locally compact if it is finite on any compact

Theorem 2237 *(Doob p.127-135 and Neeb p.42) Let E be a Hausdorff, locally compact, topological space, and S its Borel $\sigma-$algebra.*

i) For any positive, locally finite Borel measure μ the map : $\lambda_\mu(f) = \int_E f\mu$ is a positive linear functional on $C_{0c}(E;\mathbb{C})$ (this is a Radon integral). Moreover it is continuous on $C_0(K;\mathbb{C})$ for any compact K of E.

ii) Conversely for any positive linear functional λ on $C_{0c}(E;\mathbb{C})$ there is a unique Radon measure, called the Radon measure associated to λ ,such that $\lambda = \lambda_\mu$

iii) For any linear functional on $C_{0v}(E;\mathbb{C})$ there is a unique complex measure μ on (E,S) such that $|\mu|$ is regular and $\forall f \in C_{0v}(E;\mathbb{C}) : \lambda(f) = \int_E f\mu$

iv) For any continuous positive linear functional λ on $C_c(E;\mathbb{C})$ with norm 1 (a state) there is a Borel, inner regular, probability P such that : $\forall f \in C_c(E;\mathbb{C})$: $\lambda(f) = \int_E fP$. So we have P(E)=1,$P(\varpi) \geq 0$

Theorem 2238 *Riesz Representation theorem (Thill p.254): For every positive linear functional ℓ on the space $C_c(E;\mathbb{C})$ where E is a locally compact Hausdorff space, bounded with norm 1, there exists a unique inner regular Borel measure μ such that :*

$\forall \varphi \in C_c(E;\mathbb{C}) : \ell(\varphi) = \int_E \varphi\mu$

On a compact K of E, $\mu(K) = \inf\{\ell(\varphi) : \varphi \in C_c(E;\mathbb{C}), 1_K \leq \varphi \leq 1_E\}$

28.6.4 Multiplicative linear functionals

(see Analysis - Normed algebras)

For an algebra A of functions a multiplicative linear functional is an element λ of the algebraic dual A* such that $\lambda(fg) = \lambda(f)\lambda(g)$ and $\lambda \neq 0 \Rightarrow \lambda(I) = 1$. It is necessarily continuous with norm $\|\lambda\| \leq 1$ if A is a Banach *-algebra.

If E is a locally compact Hausdorff space, then the set of multiplicative linear functionals $\Delta(C_{0v}(E;\mathbb{C}))$ is homeomorphic to E :

For $x \in E$ fixed : $\delta_x : C_0(E;\mathbb{C}) \to \mathbb{C}$ with norm $\|\lambda\| \leq 1 :: \delta_x(f) = f(x)$

So the only multiplicative linear functionals are the Dirac distributions.

28.7 Distributions on functions

Distributions, also called generalized functions, are a bright example of the implementation of duality. The idea is to associate to a given space of functions its topological dual, meaning the space of linear continuous functionals. The smaller the space of functions, the larger the space of functionals. We can extend to the functionals many of the operations on functions, such as derivative, and thus enlarge the scope of these operations, which is convenient in many calculii, but also give a more unified understanding of important topics in differential equations. But the facility which is offered by the use of distributions is misleading. Everything goes fairly well when the functions are defined over $\mathbb{R}^m$, but this is another story when they are defined over manifolds.

28.7.1 Definition

Definition 2239 *A **distribution** is a continuous linear functional on a Fréchet space V of functions, called the space of **test functions**.*

Notation 2240 *V' is the space of distributions over the space of test functions V*

There are common notations for the most used spaces of distributions but, in order to avoid the introduction of another symbol, I find it simpler to keep this standard and easily understood notation, which underlined the true nature of the set.

Of course if V is a Banach space the definition applies, because a Banach vector space is a Fréchet space. But when V is a Hilbert space, then its topological dual is a Hilbert space, and when V is a Banach space, then V' is a Banach space, and in both cases we have powerful tools to deal with most of the problems. But the spaces of differentiable maps are only Fréchet spaces and it is not surprising that the most usual spaces of tests functions are space of differentiable functions.

Usual spaces of distributions on $\mathbb{R}^m$

1. The most usual Fréchet spaces of functions on $\mathbb{R}^m$ are the spaces of differentiable functions.

Let O be an open subset of $\mathbb{R}^m$. We have the following spaces of distributions :

$C_{\infty c}(O;\mathbb{C})'$: usually denoted $\mathfrak{D}(O)$
$C_{\infty}(O;\mathbb{C})'$: usually denoted $\mathfrak{E}'(O)$
$C_{rc}(O;\mathbb{C})'$
$S(\mathbb{R}^m)'$ called the space of **tempered distributions.**
$S \in S(\mathbb{R}^m)' \Leftrightarrow S \in L(S(\mathbb{R}^m);\mathbb{C}), \exists p,q \in \mathbb{N}, \exists C \in \mathbb{R} : \forall \varphi \in S(\mathbb{R}^m) :$
$|S(\varphi)| \leq C \sum_{k \leq p,, l \leq q} \sum_{(\alpha_1..\alpha_k)(\beta_1...\beta_l)} \sup_{x \in \mathbb{R}^m} |x_{\alpha_1}...x_{\alpha_k} D_{\beta_1...\beta_l} \varphi(x)|$

2. We have the following inclusions : the larger the space of tests functions, the smaller the space of distributions.

Theorem 2241 $\forall r \geq 1 \in \mathbb{N} : C_\infty (O; \mathbb{C})' \subset C_{rc} (O; \mathbb{C})' \subset C_{\infty c} (O; \mathbb{C})'$
$C_\infty (O; \mathbb{C})' \subset S (\mathbb{R}^m)' \subset C_{\infty c} (\mathbb{R}^m; \mathbb{C})'$

Usual spaces of distributions on a manifold

For any real finite dimensional manifold M, we have the following spaces of distributions :

$C_{\infty c} (M; \mathbb{C})'$: usually denoted $\mathfrak{D} (M)$
$C_\infty (M; \mathbb{C})'$: usually denoted $\mathfrak{E}' (M)$
$C_{rc} (M; \mathbb{C})'$

Linear dependance of distributions

Theorem 2242 *(Lieb p.150) Let O be an open in $\mathbb{R}^m$, if $(S_k)_{k=1}^n, S_k \in C_{\infty c} (O; \mathbb{C})'$ and $S \in C_{\infty c} (O; \mathbb{C})'$ such that : $\forall \varphi \in \cap_{k=1}^n \ker S_k : S (\varphi) = 0$ then there are $c_k \in \mathbb{C} : S = \sum_{k=1}^n c_k S_k$*

28.7.2   Topology

As a Fréchet space V is endowed with a countable family $(p_i)_{i \in I}$ of semi-norms, which induces a metric for which it is a complete locally convex Hausdorff space. The strong topology on V implies that a functional $S : V \rightarrow \mathbb{C}$ is continuous iff for any bounded subset W of V, that is a subset such that : $\forall i \in I, \exists D_{W_i} \in \mathbb{R}, \forall f \in W : p_i (f) \leq D_{Wi}$, we have :

$\exists C_W \in \mathbb{R} : \forall f \in W, \forall i \in I : |S (f)| \leq C_W p_i (f)$

Equivalently a linear functional (it belongs to the algebraic dual V*) S is continuous if for any sequence $(\varphi_n)_{n \in \mathbb{N}} \in V^{\mathbb{N}}$ converging to 0, which reads : $\forall i \in \mathbb{N} : p_i (\varphi_n) \rightarrow 0$ we have : $S(\varphi_n) \rightarrow 0$.

The most common cases are when $V = \sqcup_{K \subset E} V_K$ where V_K are functions with compact support in $K \subset E$ and E is a topological space which is the countable union of compacts subsets. Then a continuous functional can be equivalently defined as a linear functional whose restriction on V_K where K is any compact, is continuous.

The topological dual of a Fréchet space V is usually not a Fréchet space. So the natural topology on V' is the *weak topology : a sequence $(S_n), S_n \in V'$ converges to S in V' if : $\forall \varphi \in V : S_n (\varphi) \rightarrow S (\varphi)$

Notice that S must be defined and belong to V' prior to checking the convergence, the simple convergence of $S_n (\varphi)$ is usually not sufficient to guarantee that there is S in V' such that $\lim S_n (\varphi) = S (\varphi)$. However :

Theorem 2243 *(Zuily p.57) If a sequence $(S_n)_{n \in \mathbb{N}} \in \left(C_{\infty c} (O; \mathbb{C})' \right)^{\mathbb{N}}$ is such that $\forall \varphi \in C_{\infty c} (O; \mathbb{C}) : S_n (\varphi)$ converges, then there is $S \in C_{\infty c} (O; \mathbb{C})'$ such that $S_n (\varphi) \rightarrow S (\varphi)$.*

28.7.3 Identification of functions with distributions

One of the most important feature of distributions is that functions can be
"assimilated" to distributions, meaning that there is a map T between some
space of functions W and the space of distributions V', W being larger than V.

There are many theorems which show that, in most of the cases, a distribution is necessarily an integral for some measure. This question is usually
treated rather lightly. Indeed it is quite simple when the functions are defined
over $\mathbb{R}^m$ but more complicated when they are defined over a manifold. So it
needs attention.

Warning ! there is not always a function associated to a distribution.

General case

As a direct consequences of the theorems on functionals and integrals :

Theorem 2244 *For any measured space (E,S,μ) with a positive measure μ and
Fréchet vector subspace $V \subset L^p(E, S, \mu, \mathbb{C})$ with $1 \leq p \leq \infty$ and μ is $\sigma-$finite
if $p=1, \frac{1}{p} + \frac{1}{q} = 1$, the map : $T : L^q(E, S, \mu, \mathbb{C}) \to V' :: T(f)(\varphi) = \int_E f\varphi\mu$ is
linear, continuous, injective and an isometry.*

So T(f) is a distribution in V'

i) T is continuous : if the sequence $(f_n)_{n\in\mathbb{N}}$ converges to f in $L^p(E, S, \mu, \mathbb{C})$
then $T(f_n) \to T(f)$ in V'

ii) it is an isometry : $\|T(f)\|_{L^p} = \|f\|_{L^q}$

iii) Two functions f,g give the same distribution iff they are equal almost
everywhere on E.

However : T is surjective for $V = L^p(E, S, \mu, \mathbb{C})$: to any distribution $S \in L^p(E, S, \mu, \mathbb{C})'$ one can associate a function $f \in L^q(E, S, \mu, \mathbb{C})$ such that $T(f) : L^p(E, S, \mu, \mathbb{C})'$ but, because V' is larger than $L^p(E, S, \mu, \mathbb{C})'$, it is not surjective
on V'.

While this theorem is general, when the tests functions are defined on a
manifold we have two very different cases.

Functions defined in $\mathbb{R}^m$

Theorem 2245 *For any Fréchet vector subspace $V \subset L^1(O, dx, \mathbb{C})$ where O
is an open subset of $\mathbb{R}^m$, for any $f \in L^\infty(O, \mu, \mathbb{C})$ the map : $T(f) : V \to \mathbb{C} :: T(f)(\varphi) = \int_O f\varphi dx^1 \wedge ... \wedge dx^m$ is a distribution in V', and the map :
$T : L^\infty(O, dx, \mathbb{C}) \to V'$ is linear, continuous and injective*

This is just the application of the previous theorem : $f\varphi dx^1 \wedge ... \wedge dx^m \simeq f\varphi dx^1 \otimes ... \otimes dx^m$ defines a measure which is locally finite.

More precisely we have the following associations (but they are not bijective
!)

$C_{\infty c}(O; \mathbb{C})' \leftrightarrow W = L^p_{loc}(O, dx, \mathbb{C}), 1 \leq p \leq \infty$

$C_\infty (O; \mathbb{C})' \leftrightarrow W = L_c^p (O, dx, \mathbb{C}), 1 \leq p \leq \infty$ if the function is compactly supported so is the distribution

$S (\mathbb{R}^m)' \leftrightarrow W = L^p (\mathbb{R}^m, dx, \mathbb{C}), 1 \leq p \leq \infty$ (Zuily p.113)

$S (\mathbb{R}^m)' \leftrightarrow$ W={ f$\in C (\mathbb{R}^m; \mathbb{C})$ measurable, $|f(x)| \leq P (x)$ where P is a polynomial } (Zuily p.113)

In an all cases : $f \in W : T (f) (\varphi) = \int f \varphi dx$ and the map : $T : W \to V'$ is injective and continuous (but not surjective). So :

$T (f) = T (g) \Leftrightarrow f = g$ almost everywhere

Moreover :

$T (C_{\infty c} (O; \mathbb{C}))$ is dense in $C_{\infty c} (O; \mathbb{C})'$ and $(C_\infty (O; \mathbb{C}))'$

$T (S (\mathbb{R}^m))$ is dense in $S (\mathbb{R}^m)'$

In all the cases above the measure is absolutely continuous, because it is a multiple of the Lebesgue measure. However for $C_{\infty c} (O; \mathbb{C})$ we have more:

Theorem 2246 *(Lieb p.161) Let O be an open in $\mathbb{R}^m$ and $S \in C_{\infty c} (O; \mathbb{C})'$ such that : $\forall \varphi \geq 0 : S (\varphi) \geq 0$ then there is a unique positive Radon measure μ on O such that : $S (\varphi) = \mu (\varphi) = \int_O \varphi \mu$. Conversely any Radon measure defines a positive distribution on O.*

So on $C_{\infty c} (O; \mathbb{C})$ distributions are essentially Radon measures (which are not necessarily absolutely continuous).

Functions over a manifold

If $V \subset C (M; \mathbb{C})$ and M is some m dimensional manifold, other than an open in $\mathbb{R}^m$ the quantity f$d\xi^1 \wedge ... \wedge d\xi^m$ is not a m form and does not define a measure on M. There are two ways to define distributions.

Theorem 2247 *For an oriented Hausdorff class 1 real manifold M, any continuous m form ϖ on M and its induced Lebesgue measure μ, any Fréchet vector space of functions $V \subset L^p (M, \mu, \mathbb{C})$ with $1 \leq p \leq \infty$, any function $f \in L^q (O, \mu, \mathbb{C})$ with $\frac{1}{p} + \frac{1}{q} = 1$, the map : $T (f) : V \to \mathbb{C} :: T (f) (\varphi) = \int_M f \varpi$ is a distribution in V', and the map : $T : L^\infty (O, dx, \mathbb{C}) \to V'$ is linear, continuous and injective.*

Proof. any continuous m form defines a Radon measure μ, and M is $\sigma-$finite

this measure μ can be decomposed into two positive measure μ_+, μ_- which are still Radon measure (they are locally compact because $\mu (K) < \infty \Rightarrow \mu_+ (K), \mu_- (K) < \infty$)

From there it suffices to apply the theorem above. ∎

As μ is locally finite the theorem holds for any space V of bounded functions with compact support. In particular $C_{0c} (M; \mathbb{C})$ is dense in $L^p (M, \mu, \mathbb{C})$ for each $\infty \geq$ p≥ 1.

The other association is more peculiar.

Theorem 2248 *For an oriented Hausdorff class 1 real manifold M, any Fréchet vector space of functions $V \subset C_{0c} (M, \mathbb{C})$ the map : $T (\varpi) : \Lambda_m (M; \mathbb{C}) \to \mathbb{C} :: T (\varpi) (\varphi) = \int_M \varphi \varpi$ is a distribution in V'.*

Proof. The Lebesgue measure μ induced by ϖ is locally finite, so φ is integrable if it is bounded with compact support.

μ can be decomposed in 4 real positive Radon measures : $\mu = (\mu_{r+} - \mu_{r-}) + i(\mu_{I+} - \mu_{I-})$ and for each T is continuous on $C_0(K;\mathbb{C})$ for any compact K of M so it is continuous on $C_{0c}(M,\mathbb{C})$ ∎

Notation : Warning ! It is clear that $f \neq T(f)$, even If is common to use the same symbol for the function and the associated distribution. By experience this is more confusing than helpful. So we will stick to :

Notation 2249 *T(f) is the distribution associated to a function f or a form by one of the maps T above, usually obvious in the context.*

And conversely, if for a distribution S there is a function f such that S=T(f), we say that S is induced by f.

28.7.4 Support of a distribution

Definition 2250 *Let V a Fréchet space in C(E;$\mathbb{C}$). The **support** of a distribution S in V' is defined as the subset of E, complementary of the largest open O in E such as : $\forall \varphi \in V, Supp(\varphi) \subset O \Rightarrow S(\varphi) = 0$.*

Notation 2251 *(V')$_c$ is the set of distributions of V' with compact support*

Definition 2252 *For a Fréchet space V$\subset C(E;\mathbb{C})$ of tests functions, and a subset Ω of E, the **restriction** $S|_\Omega$ of a distribution S$\in V'$ is the restriction of S to the subspace of functions : $V \cap C(\Omega;\mathbb{C})$*

Notice that, contrary to the usual rule for functions, the map : $V' \to (V \cap C(\Omega;\mathbb{C}))'$ is neither surjective or injective. This can be understood by the fact that $V \cap C(\Omega;\mathbb{C})$ is a smaller space, so the space of its functionals is larger.

Definition 2253 *If for a Fréchet space V$\subset C(E;\mathbb{C})$ of tests functions there is a map : $T:W\to V'$ for some subspace of functions on E, the **singular support** of a distribution S is the subset SSup(S) of E, complementary of the largest open O in E such as : $\exists f \in W \cap C(O;\mathbb{C}) : T(f) = S|_O$.*

So S cannot be represented by a function which has its support in SSup(S).
Then :
SSup(S)=$\varnothing \Rightarrow \exists f \in W : T(f) = S$
SSup(S)$\subset Sup(S)$
If S$\in (C_{c\infty}(O;\mathbb{C}))'$, $f \in C_\infty(O;\mathbb{C})$: $SSup(fS) = SSup(S) \cap Supp(f)$, $SSup(S + T(f)) = SSup(S)$

Theorem 2254 *(Zuily p.120) The set $(C_{\infty c}(O;\mathbb{C}))'_c$ of distributions with compact support can be identified with the set of distributions on the smooth functions : $(C_\infty(O;\mathbb{C}))' \equiv (C_{\infty c}(O;\mathbb{C}))'_c$ and $(C_{\infty c}(\mathbb{R}^m;\mathbb{C}))'_c$ is dense in $S(\mathbb{R}^m)'$*

Theorem 2255 *For any family of distributions* $(S_i)_{i \in I}$ *,* $S_i \in C_{\infty c}(O_i; \mathbb{C})'$ *where* $(O_i)_{i \in I}$ *is an open cover of O in* $\mathbb{R}^m$ *, such that :* $\forall i, j \in I, S_i|_{O_i \cap O_j} = S_j|_{O_i \cap O_j}$ *there is a unique distribution* $S \in C_{\infty c}(O; \mathbb{C})'$ *such that* $S|_{O_i} = S_i$

28.7.5 Product of a function and a distribution

Definition 2256 *The product of a distribution* $S \in V'$ *where* $V \subset C(E; \mathbb{C})$ *and a function* $f \in C(E; \mathbb{C})$ *is the distribution :* $\forall \varphi \in V : (fS)(\varphi) = S(f\varphi),$ *defined whenever* $f\varphi \in V$.

The operation $f\varphi$ is the pointwise multiplication : $(f\varphi)(x) = f(x)\varphi(x)$
The product is well defined for :
$S \in C_{rc}(O; \mathbb{C})', 1 \leq r \leq \infty$ and $f \in C_\infty(O; \mathbb{C})$
$S \in S(\mathbb{R}^m)'$ and f any polynomial in $\mathbb{R}^m$
When the product is well defined : $Supp(fS) \subset Supp(f) \cap Supp(S)$

28.7.6 Derivative of a distribution

This is the other most important property of distributions, and the main reason for their use : distributions are smooth. So, using the identification of functions to distributions, it leads to the concept of derivative "in the meaning of distributions". However caution is required on two points. First the "distributional derivative", when really useful (that is when the function is not itself differentiable) is not a function, and the habit of using the same symbol for the function and its associated distribution leads quickly to confusion. Second, the distributional derivative is simple only when the function is defined over $\mathbb{R}^m$. Over a manifold this is a bit more complicated and the issue is linked to the concept of distribution on a vector bundle seen in the next subsection.

Definition

Definition 2257 *The r derivative* $D_{\alpha_1..\alpha_r}S$ *of a distribution* $S \in V'$ *on a Fréchet space of r differentiable functions on a class r m dimensional manifold M, is the distribution :* $\forall \varphi \in V : (D_{\alpha_1..\alpha_r}S)(\varphi) = (-1)^r S(D_{\alpha_1..\alpha_r}\varphi)$

Notice that the test functions φ must be differentiable.
By construction, the derivative of a distribution is well defined whenever : $\forall \varphi \in V : D_{\alpha_1..\alpha_r}\varphi \in V$ which is the case for all the common spaces of tests functions. In particular :
If $S \in C_{rc}(O; \mathbb{C})'$: then $\forall s \leq r : \exists D_{\alpha_1..\alpha_s}S \in C_{rc}(O; \mathbb{C})'$ and $D_{\alpha_1..\alpha_r}S \in C_{r+1,c}(O; \mathbb{C})'$
if $S \in S(\mathbb{R}^m)'$: then $\forall (\alpha_1...\alpha_r) : \exists D_{\alpha_1..\alpha_r}S \in S(\mathbb{R}^m)'$
As a consequence if V is a space of r differentiable functions, any distribution is r differentiable.

Fundamental theorems

Theorem 2258 *Let V be a Fréchet space of r differentiable functions $V \subset C_{rc}(O, \mathbb{C})$ on an open subset O of $\mathbb{R}^m$. If the distribution S is induced by the integral of a r differentiable function $f \in C_{rc}(O; \mathbb{C})$ then we have : $D_{\alpha_1..\alpha_r}(T(f)) = T(D_{\alpha_1..\alpha_r} f)$*

meaning that the derivative of the distribution is the distribution induced by the derivative of the function.

Proof. The map T reads : $T : W \to V' :: T(f)(\varphi) = \int_O f\varphi d\xi^1 \wedge ... \wedge d\xi^m$

We start with r=1 and $D_\alpha = \partial_\alpha$ with $\alpha \in 1...m$ and denote: $d\xi = d\xi^1 \wedge ... \wedge d\xi^m$

$i_{\partial_\alpha} d\xi = (-1)^{\alpha-1} d\xi^1 \wedge ... \wedge \left(\widehat{d\xi^\alpha}\right) \wedge .. \wedge d\xi^m$

$d(i_{\partial_\alpha} \varpi) = 0$

$(d\varphi) \wedge i_{\partial_\alpha} d\xi = \sum_\beta (-1)^{\alpha-1} (\partial_\beta \varphi) d\xi^\beta \wedge d\xi^1 \wedge ... \wedge \left(\widehat{d\xi^\alpha}\right) \wedge .. \wedge d\xi^m = (\partial_\beta \varphi) dx$

$d(f\varphi i_{\partial_\alpha} d\xi) = d(f\varphi) \wedge i_{\partial_\alpha} d\xi - (f\varphi) \wedge d(i_{\partial_\alpha} d\xi) = \varphi(df) \wedge i_{\partial_\alpha} d\xi + f d(\varphi) \wedge i_{\partial_\alpha} d\xi = (f\partial_\alpha \varphi) d\xi + (\varphi \partial_\alpha f) d\xi$

Let N be a manifold with boundary in O. The Stockes theorem gives :

$\int_N d(f\varphi i_{\partial_\alpha} d\xi) = \int_{\partial N} f\varphi i_{\partial_\alpha} d\xi = \int_N (f\partial_\alpha \varphi) d\xi + \int_N (\varphi \partial_\alpha f) d\xi$

N can always be taken such that $Supp(\varphi) \subset \overset{\circ}{N}$ and then :

$\int_N (f\partial_\alpha \varphi) d\xi + \int_N (\varphi \partial_\alpha f) d\xi = \int_O (f\partial_\alpha \varphi) d\xi + \int_O (\varphi \partial_\alpha f) d\xi$

$\int_{\partial N} f\varphi i_{\partial_\alpha} d\xi = 0$ because $Supp(\varphi) \subset \overset{\circ}{N}$

So : $\int_O (f\partial_\alpha \varphi) d\xi = -\int_O (\varphi \partial_\alpha f) d\xi \Leftrightarrow T(f)(\partial_\alpha \varphi) = -T(\partial_\alpha f)(\varphi) = -\partial_\alpha T(f)(\varphi)$

$T(\partial_\alpha f)(\varphi) = \partial_\alpha T(f)(\varphi)$

By recursion over r we get the result ∎

The result still holds if $O = \mathbb{R}^m$

For a manifold the result is different with a similar (but not identical) proof:

Theorem 2259 *Let $V \subset C_{rc}(M, \mathbb{C})$ be a Fréchet space of r differentiable functions on an oriented Hausdorff class r real manifold M, If the distribution S is induced by a r differentiable m form $\varpi = \varpi_0 d\xi^1 \wedge ... \wedge d\xi^m$ then we have : $(D_{\alpha_1...\alpha_r} T(\varpi))(\varphi) = (-1)^r T\left((D_{\alpha_1...\alpha_r} \varpi_0) d\xi^1 \wedge ... \wedge d\xi^m\right)(\varphi)$*

By r differentiable m form we mean that the function ϖ_0 is r differentiable on each of its domain (ϖ_0 changes according to the rule : $\varpi_{b0} = \det[J_{ba}] \varpi_{a0}$)

Proof. The map T reads : $T : W \to V' :: T(\varpi)(\varphi) = \int_M \varphi \varpi$

We start with r=1 and $D_\alpha = \partial_\alpha$ with $\alpha \in 1...m$ and denote: $d\xi^1 \wedge ... \wedge d\xi^m = d\xi, \varpi = \varpi_0 d\xi$

$i_{\partial_\alpha} d\varpi = (-1)^{\alpha-1} \varpi_0 d\xi^1 \wedge ... \wedge \left(\widehat{d\xi^\alpha}\right) \wedge .. \wedge d\xi^m$

$d(i_{\partial_\alpha} \varpi) = \sum_\beta (-1)^{\alpha-1} (\partial_\beta \varpi_0) d\xi^\beta \wedge d\xi^1 \wedge ... \wedge \left(\widehat{d\xi^\alpha}\right) \wedge .. \wedge d\xi^m = (\partial_\alpha \varpi_0) d\xi$

$(d\varphi) \wedge i_{\partial_\alpha} \varpi = \sum_\beta (-1)^{\alpha-1} (\partial_\beta \varphi) \varpi_0 d\xi^\beta \wedge d\xi^1 \wedge ... \wedge \left(\widehat{d\xi^\alpha}\right) \wedge .. \wedge d\xi^m = (\partial_\beta \varphi) \varpi_0 d\xi = (\partial_\beta \varphi) \varpi$

$d\left(\varphi i_{\partial_\alpha}\varpi\right) = d\varphi \wedge i_{\partial_\alpha}\varpi - \varphi d\left(i_{\partial_\alpha}\varpi\right) = \left(\partial_\alpha\varphi\right)\varpi - \varphi\left(\partial_\alpha\varpi_0\right)d\xi$

Let N be a manifold with boundary in M. The Stockes theorem gives :

$\int_N d\left(\varphi i_{\partial_\alpha}\varpi\right) = \int_{\partial N}\varphi i_{\partial_\alpha}\varpi = \int_N \left(\partial_\alpha\varphi\right)\varpi - \int_N \varphi\left(\partial_\alpha\varpi_0\right)d\xi$

N can always be taken such that $Supp\left(\varphi\right) \subset \overset{\circ}{N}$ and then : $\int_{\partial N}\varphi i_{\partial_\alpha}\varpi = 0$

$\int_M \left(\partial_\alpha\varphi\right)\varpi = \int_M \varphi\left(\partial_\alpha\varpi_0\right)d\xi$

$\left(\partial_\alpha T\left(\varpi\right)\right)\left(\varphi\right) = -T\left(\partial_\alpha\varpi_0 d\xi\right)\left(\varphi\right)$

By recursion over r we get :

$\left(D_{\alpha_1...\alpha_r}T\left(\varpi\right)\right)\left(\varphi\right) = \left(-1\right)^r T\left(D_{\alpha_1...\alpha_r}\varpi_0 d\xi\right)\left(\varphi\right)$ ∎

Notice that if T is defined by : $T\left(f\right)\left(\varphi\right) = \int_M f\varphi\varpi$ where $\varpi = \varpi_0 d\xi$ the formula does not hold any more. This is why the introduction of m form is useful. However the factor $\left(-1\right)^r$ is not satisfactory. It comes from the very specific case of functions over $\mathbb{R}^m$, and in the next subsections we have a better solution.

Derivative "in the meaning of distributions"

For a function f defined in $\mathbb{R}^m$, which is not differentiable, but is such that there is some distribution T(f) which is differentiable, its derivative "in the sense of distributions" (or distributional derivative) is the derivative of the distribution $D_\alpha T\left(f\right)$ sometimes denoted $\{D_\alpha f\}$ and more often simply $D_\alpha f$. However the distributional derivative of f is represented by a function iff f itself is differentiable:

Theorem 2260 *(Zuily p.53) Let* $S\in C_{\infty c}\left(O;\mathbb{C}\right)'$ *with O an open in* $\mathbb{R}^m$ *. The following are equivalent :*

i) $\exists f \in C_r\left(O;\mathbb{C}\right) : S = T(f)$

ii) $\forall\alpha_1..\alpha_s = 1...m, s = 0...r : \exists g \in C_0\left(O;\mathbb{C}\right) : \partial_{\alpha_1...\alpha_s}S = T\left(g\right)$

So, if we can extend extend "differentiability" to many functions which otherwise are not differentiable, we must keep in mind that usually $\{D_\alpha f\}$ is not a function, and is defined with respect to some map T.

One finds quite often expressions like "a function f such that its distributional derivatives belong to L^p" . They must be interpreted as "f such that $D_\alpha\left(T\left(f\right)\right) = T\left(g\right)$ with g$\in L^p$" . But if the distributional derivative of f is represented by a function, it means that f is differentiable, and so it would be simpler and clearer to say "f such that $D_\alpha f \in L^p$" .

Because of all these problems we will always stick to the notation T(f) to denote the distribution induced by a function.

In the most usual cases we have the following theorems :

Theorem 2261 *If f$\in L^1_{loc}\left(\mathbb{R}^m, dx, \mathbb{C}\right)$ is locally integrable :* $\forall\left(\alpha_1...\alpha_r\right) : \exists D_{\alpha_1..\alpha_r}\left(T\left(f\right)\right) \in C_{\infty c}\left(\mathbb{R}^m;\mathbb{C}\right)'$

Theorem 2262 *(Lieb p.175) Let O be an open in* $\mathbb{R}^m$. *If f$\in H^1\left(O\right), g \in C_\infty\left(O;\mathbb{C}\right)$ and has bounded derivatives, then :* $f \times g \in H^1\left(O\right)$ *and* $\partial_\alpha T\left(fg\right) \in C_{\infty c}\left(O;\mathbb{C}\right)'$

Theorem 2263 *(Zuily p.39)* $S\in C_{\infty c}\left(\mathbb{R};\mathbb{C}\right)', \frac{dS}{dx}+aS = T\left(f\right), a \in C_\infty\left(\mathbb{R};\mathbb{C}\right), f \in C_0\left(\mathbb{R};\mathbb{C}\right) \Leftrightarrow \exists g \in C_1\left(\mathbb{R};\mathbb{C}\right) : S = T\left(g\right)$

Properties of the derivative of a distribution

Theorem 2264 *Support : For a distribution S, whenever the derivative $D_{\alpha_1..\alpha_r}S$ exist : $Supp\,(D_{\alpha_1..\alpha_r}S) \subset Supp\,(S)$*

Theorem 2265 *Leibnitz rule: For a distribution S and a function f, whenever the product and the derivative exist :*
$$\partial_\alpha\,(fS) = f\partial_\alpha S + (\partial_\alpha f)\,S$$

Notice that for f this is the usual derivative.

Theorem 2266 *Chain rule (Lieb p.152): Let O open in $\mathbb{R}^m, y = (y_k)_{k=1}^n, y_k \in W_{loc}^{1,p}\,(O)\,, F \in C_1\,(\mathbb{R}^n;\mathbb{C})$ with bounded derivative, then :*
$$\frac{\partial}{\partial x_j}T\,(F \circ y) = \sum_k \frac{\partial F}{\partial y_k}\frac{\partial}{\partial x_j}T\,(y_k)$$

Theorem 2267 *Convergence : If the sequence of distributions $(S_n)_{n\in\mathbb{N}} \in V^{\mathbb{N}}$ converges to S in V', and have derivatives in V', then $D_\alpha S_n \to D_\alpha S$*

Theorem 2268 *Local structure of distributions (Zuily p.76): For any distribution S in $C_{\infty c}\,(O;\mathbb{C})'$ with O an open in $\mathbb{R}^m$, and any compact K in O, there is a finite family $(f_i)_{i\in I} \in C_0\,(O;\mathbb{C})^I : S|_K = \sum_{i\in I,\alpha_1...\alpha_r} D_{\alpha_1...\alpha_r}T\,(f_i)$*

As a consequence a distribution whose derivatives are null is constant :
$$S\in C_{\infty c}\,(\mathbb{R}^m;\mathbb{C})'\,, \alpha = 1..m : \frac{\partial}{\partial x^\alpha}S = 0 \Leftrightarrow S = Cst$$

Theorem 2269 *(Lieb p.145) Let O be an open in $\mathbb{R}^m$, $S\in C_{\infty c}\,(O;\mathbb{C})'$, $\varphi \in C_{\infty c}\,(O;\mathbb{C})\,, y \in \mathbb{R}^m$ such that $\forall t \in [0,1] : ty \in O$ then :*
$$S\,(\psi) - S\,(\varphi) = \int_0^1 \sum_{k=1}^m y_k \partial_k S\,(\varphi\,(ty))\,dt \text{ with } \psi : \{ty, t \in [0,1]\} \to \mathbb{C} ::$$
$\psi\,(x) = \varphi\,(ty)$
If $f\in W_{loc}^{1,1}\,(\mathbb{R}^m)$ then : $\forall\varphi \in C_{\infty c}\,(O;\mathbb{C})$:
$$\int_O (f\,(x+y) - f\,(x))\,\varphi\,(x)\,dx = \int_O \left(\int_0^1 \sum_{k=1}^m y_k \partial_k f\,(x+ty)\,dt\right)\varphi\,(x)\,dx$$

Theorem 2270 *Jumps formula (Zuily p.40) Let $f \in C\,(\mathbb{R};\mathbb{R})$ be a function continuous except at the isolated points a_i where it is semi-continuous : $\exists \lim_{\epsilon\to 0} f\,(a_i \pm \epsilon)$. Then the derivative of f reads : $\frac{d}{dx}T\,(f) = f' + \sum_i \sigma_i \delta_{a_i}$ with $\sigma_i = \lim_{\epsilon\to 0} f\,(a_i + \epsilon) - \lim_{\epsilon\to 0} f\,(a_i - \epsilon)$ and f' the usual derivative where it is well defined.*

28.7.7 Heaviside and Dirac functions

Definitions

Definition 2271 *For any Fréchet space V in C(E;$\mathbb{C}$) the **Dirac's distribution** is : $\delta_a : V \to \mathbb{C} :: \delta_a\,(\varphi) = \varphi\,(a)$ where $a\in E$. Its derivatives are : $D_{\alpha_1...\alpha_r}\delta_a\,(\varphi) = (-1)^r D_{\alpha_1...\alpha_r}\varphi|_{x=a}$*

Warning ! If $E\neq \mathbb{R}^m$ usually δ_0 is meaningless

Definition 2272 *The **Heaviside function** on $\mathbb{R}$ is the function given by $H\,(x) = 0, x \leq 0, H(x) = 1, x > 0$*

General properties

The Heaviside function is locally integrable and the distribution T(H) is given in $C_{\infty c}(\mathbb{R};\mathbb{C})$ by : $T(H)(\varphi) = \int_0^\infty \varphi dx$. It is easy to see that : $\frac{d}{dx}T(H) = \delta_0$. And we can define the Dirac *function* : $\delta : \mathbb{R} \to \mathbb{R} :: \delta(0) = 1, x \neq 0 : \delta(x) = 0$. So $\left\{\frac{d}{dx}H\right\} = T(\delta) = \delta_0$. But there is no function such that T(F)=$\frac{d^2}{dx^2}$T(H)

Theorem 2273 *Any distribution S in $C_\infty(O;\mathbb{C})'$ which has a support limited to a unique point has the form :*

$S(\varphi) = c_0 \delta_a(\varphi) + \sum_\alpha c_\alpha D_\alpha \varphi|_a$ *with $c_0, c_\alpha \in \mathbb{C}$*

If $S \in C_{\infty c}(O;\mathbb{C})', xS = 0 \Rightarrow S = \delta_0$
If $f \in L^1(\mathbb{R}^m, dx, \mathbb{C}), \varepsilon > 0 : \lim_{\varepsilon \to 0} \varepsilon^{-m} f(\varepsilon^{-1}x) = \delta_0 \int_{\mathbb{R}^m} f(x) dx$

Laplacian

On $C_\infty(\mathbb{R}^m;\mathbb{R})$ the laplacian is the differential operator : $\Delta = \sum_{\alpha=1}^m \frac{\partial^2}{(\partial x^\alpha)^2}$

The function : $f(x) = \left(\sum_i (x^i - a^i)^2\right)^{1-\frac{m}{2}} = \frac{1}{r^{m-2}}$ is such that for m$\geq$ 3 : $\Delta(T(f)) = (2-m)A(S_{m-1})\delta_a$ where $A(S_{m-1})$ is the Lebesgue surface of the unit sphere in $\mathbb{R}^m$. For m=3 :$\Delta T\left(\left(\sum_i (x^i - a^i)^2\right)^{-1/2}\right) = -4\pi\delta_a$

For m=2 : $\Delta T\left(\ln\left(\sum_i (x^i - a^i)^2\right)^{1/2}\right) = 2\pi\delta_a$

Theorem 2274 *(Taylor 1 p.210) If $S \in S(\mathbb{R}^m)'$ is such that $\Delta S = 0$ then S=T(f) with f a polynomial in $\mathbb{R}^m$*

28.7.8   Tensorial product of distributions

Theorem 2275 *(Zuily p.18,64) For any open subsets $O_1, O_2 \subset \mathbb{R}^m$:*
i) any function in $C_{\infty c}(O_1 \times O_2;\mathbb{C})$ is the limit of a sequence of functions of $C_{\infty c}(O_1;\mathbb{C}) \otimes C_{\infty c}(O_2;\mathbb{C})$
ii) for any distributions $S_1 \in C_{\infty c}(O_1;\mathbb{C})', S_2 \in C_{\infty c}(O_2;\mathbb{C})'$ there is a unique distribution $S_1 \otimes S_2 \in C_{\infty c}(O_1 \times O_2;\mathbb{C})'$ such that : $\forall \varphi_1 \in C_{\infty c}(O_1;\mathbb{C}), \varphi_2 \in C_{\infty c}(O_2;\mathbb{C}) : S_1 \otimes S_2(\varphi_1 \otimes \varphi_2) = S_1(\varphi_1)S_2(\varphi_2)$
iii) $\forall \varphi \in C_{\infty c}(O_1 \times O_2;\mathbb{C}) : S_1 \otimes S_2(\varphi) = S_1(S_2(\varphi(x_1,.))) = S_2(S_1(\varphi(.,x_2)))$
iv) If $S_1 = T(f_1), S_2 = T(f_2)$ then $S_1 \otimes S_2 = T(f_1) \otimes T(f_2) = T(f_1 \otimes f_2)$
v) $\frac{\partial}{\partial x_1^\alpha}(S_1 \otimes S_2) = \left(\frac{\partial}{\partial x_1^\alpha}S_1\right) \otimes S_2$

Theorem 2276 *Schwartz kernel theorem (Taylor 1 p.296) : let M,N be compact finite dimensional real manifolds. $L : C_\infty(M;\mathbb{C}) \to C_{\infty c}(N;\mathbb{C})'$ a continuous linear map, B the bilinear map : $B : C_\infty(M;\mathbb{C}) \times C_\infty(N;\mathbb{C}) \to \mathbb{C} :: B(\varphi,\psi) = L(\varphi)(\psi)$ separately continuous in each factor. Then there is a distribution : $S \in C_{\infty c}(M \times N;\mathbb{C})'$ such that : $\forall \varphi \in C_\infty(M;\mathbb{C}), \psi \in C_\infty(N;\mathbb{C}) : S(\varphi \otimes \psi) = B(\varphi,\psi)$*

28.7.9 Convolution of distributions

Definition

Convolution of distributions are defined so as we get back the convolution of functions when a distribution is induced by a function : $T(f) * T(g) = T(f * g)$. It is defined only for functions on $\mathbb{R}^m$.

Definition 2277 *The **convolution of the distributions** $S_1, S_2 \in C_{rc}(\mathbb{R}^m; \mathbb{C})'$ is the distribution $S_1 * S_2 \in C_{rc}(\mathbb{R}^m; \mathbb{C})'$:*

$\forall \varphi \in C_{\infty c}(\mathbb{R}^m; \mathbb{C}) :: (S_1 * S_2)(\varphi) = (S_1(x_1) \otimes S_2(x_2))(\varphi(x_1 + x_2)) = S_1(S_2(\varphi(x_1 + x_2))) = S_2(S_1(\varphi(x_1 + x_2)))$

It is well defined when at least one of the distributions has a compact support. If both have compact support then $S_1 * S_2$ has a compact support.

The condition still holds for the product of more than two distributions : all but at most one must have compact support.

If is possible to enlarge the spaces of distribution on which convolution is well defined.

If $S_1, S_2 \in C_{\infty c}(\mathbb{R}^m; \mathbb{C})'$ and their support meets the condition: $\forall R > 0, \exists \rho : x_1 \in SupS_1, x_2 \in SupS_2, \|x_1 + x_2\| \leq R \Rightarrow \|x_1\| \leq \rho, \|x_2\| \leq \rho$ then the convolution $S_1 * S_2$ is well defined, even if the distributions are not compacly supported.

If the domain of the functions is some relatively compact open O in $\mathbb{R}^m$ we can always consider them as compactly supported functions defined in the whole of $\mathbb{R}^m$ by : $x \notin O : \varphi(x) = 0$

Properties

Theorem 2278 *Convolution of distributions, when defined, is associative, commutative.*

With convolution $C_{rc}(\mathbb{R}^m; \mathbb{C})'_c$ is an unital commutative algebra with unit element the Dirac distribution δ_0

Theorem 2279 *Over $C_{rc}(\mathbb{R}^m; \mathbb{C})'_c$ the derivative of the convolution product is : $D_{\alpha_1 \dots \alpha_r}(S \otimes U) = (D_{\alpha_1 \dots \alpha_r} S) * U = S * D_{\alpha_1 \dots \alpha_r} U$*

Theorem 2280 *If the sequence $(S_n)_{\in \mathbb{N}} \in \left(C_{rc}(\mathbb{R}^m; \mathbb{C})'_c\right)^{\mathbb{N}}$ converges to S then $\forall U \in C_{rc}(\mathbb{R}^m; \mathbb{C})' : S_n * U \to S * U$ in $C_{rc}(\mathbb{R}^m; \mathbb{C})'_c$*

Convolution of distributions induced by a function

We have the general result : $T(f) * T(g) = T(f * g)$ but more generally :

Theorem 2281 *(Zuily p.69) The convolution of a distribution and a distribution induced by a function gives a distribution induced by a function. If both distribution and function are compactly supported then the product is a compactly supported function.*

$S * T (f) = T (S_y (f (x - y)))$ and $S_y (f (x - y)) \in C_r (\mathbb{R}^m; \mathbb{C})$

With : $S \in C_{rc} (\mathbb{R}^m; \mathbb{C})', f \in C_{rc} (\mathbb{R}^m; \mathbb{C})$ or $S \in C_{rc} (\mathbb{R}^m; \mathbb{C})'_c, f \in C_r (\mathbb{R}^m; \mathbb{C})$

In particular : $\delta_a * T (f) = T (\delta_a (t) f (x - t)) = T (f (x - a))$

Supp(S*T(f))⊂Supp(S)+Supp(T(f))

SSup(S*U)⊂SSup(S)+SSup(U)

28.7.10 Pull back of a distribution

Definition

Theorem 2282 *(Zuily p.82) If O_1 is an open in $\mathbb{R}^m$, O_2 an open in $\mathbb{R}^n$, F:$O_1 \to O_2$ a smooth submersion, there is a map $F^* : C_{\infty c} (O_2; \mathbb{C})' \to C_{\infty c} (O_1; \mathbb{C})'$ such that : $\forall f_2 \in C_0 (O_2; \mathbb{C})$, $F^*T (f_2)$ is the unique functional $F^*T (f_2) \in C_{\infty c} (O_1; \mathbb{C})'$ such that $F^*T (f_2) (\varphi_1) = T (f_2 \circ F) (\varphi_1) = T (F^* f_2) (\varphi_1)$*

F is a submersion = F is differentiable and rank F'(p)= $n \leq m$.

The definition is chosen so that the pull back of the distribution given by f_2 is the distribution given by the pull back of f_2. So this is the assimilation with functions which leads the way. But the map F* is valid for any distribution.

Properties

Theorem 2283 *(Zuily p.82) The pull back of distributions has the following properties :*

i) $Supp(F^*S_2) \subset F^{-1} (Supp (S_2))$

ii) F^*S_2 is a positive distribution if S_2 is positive

iii) $\frac{\partial}{\partial x_1^\alpha} F^*S_2 = \sum_{\beta=1}^n \frac{\partial F_\beta}{\partial x_1^\alpha} F^* \left(\frac{\partial S_2}{\partial x_2^\beta} \right)$

iv) $g_2 \in C_\infty (O_2; \mathbb{C}) : F^* (g_2 S_2) = (g_2 \circ F) F^*S_2 = (F^*g_2) \times F^*S_2$

v) if $G : O_2 \to O_3$ is a submersion, then $(G \circ F)^* S_3 = F^* (G^*S_3)$

vi) If the sequence $(S_n)_{\in \mathbb{N}} \in \left(C_{\infty c} (O_2; \mathbb{C})' \right)^{\mathbb{N}}$ converges to S then $F^*S_n \to F^*S$

Pull back by a diffeomorphism

The application of the previous theorem when F is a smooth diffeomorphism leads to the definition :

Theorem 2284 *If O_1, O_2 are open subsets in $\mathbb{R}^m$, F:$O_1 \to O_2$ a smooth diffeomorphism, $V_1 \subset C (O_1; \mathbb{C})$, $V_2 \subset C (O_2; \mathbb{C})$ two Fréchet spaces, the pull back on V_1' of a distribution $S_2 \in V_2'$ by a smooth diffeomorphism is the map : $F^*S_2 (\varphi_1) = |\det F'|^{-1} S_2 (\varphi_1 \circ F^{-1})$*

$\forall \varphi_1 \in C_{\infty c} (O_1; \mathbb{C})$, $f_2 \in C_0 (O_2; \mathbb{C})$:

$F^*T (f_2) (\varphi_1) = \int_{O_1} f_2 (F (x_1)) \varphi_1 (x_1) dx_1 = \int_{O_2} f_2 (x_2) \varphi_1 \left(F^{-1} (x_2) \right) |\det F' (x_2)|^{-1} dx_2$

$$\Leftrightarrow F^*T\left(f_2\right)\left(\varphi_1\right) = T\left(f_2\right)\left(\varphi_1\left(F^{-1}\left(x_2\right)\right)\left|\det F'\left(x_2\right)\right|^{-1}\right)$$

Notice that we have the absolute value of $\det F'\left(x_2\right)^{-1}$: this is the same formula as for the change of variable in the Lebesgue integral.

We have all the properties listed above. Moreover, if $S \in S\left(\mathbb{R}^m\right)'$ then $F^*S \in S\left(\mathbb{R}^m\right)'$

Applications

For distributions $S \in C_{\infty c}\left(\mathbb{R}^m; \mathbb{C}\right)'$,

1. Translation :
$\tau_a(x) = x - a$
$\tau_a^* S\left(\varphi\right) = S_y\left(\varphi\left(y + a\right)\right) = \int_{O_2} S_y(y)\varphi\left(y + a\right) dy$
$\delta_a = \tau_a^* \delta_0$

2. Similitude : ,
$k \neq 0 \in R : \lambda_k(x) = kx$
$\lambda_k^* S\left(\varphi\right) = \frac{1}{|k|} S_y\left(\varphi\left(\frac{y}{k}\right)\right) = \frac{1}{|k|}\int_{O_2} S_y(y)\varphi\left(\frac{y}{k}\right) dy$

3. Reflexion :
$R(x) = -x$
$R^* S\left(\varphi\right) = S_y\left(\varphi\left(-y\right)\right) = \int_{O_2} S_y(y)\varphi\left(-y\right) dy$

28.7.11 Miscellaneous operations with distributions

Homogeneous distribution

Definition 2285 $S \in C_{\infty c}\left(\mathbb{R}^m; \mathbb{C}\right)'$ *is* **homogeneous** *of degree a if* $\forall k > 0$:
$S\left(\varphi\left(kx\right)\right) = k^{-m-a} S\left(\varphi(x)\right)$

The definition coincides with the usual is $S = T\left(f\right)$

S is homogeneous of degree a iff $\sum_{\alpha=1}^m x^\alpha \dfrac{\partial S}{\partial x^\alpha} = aS$

Distribution independant with respect to a variable

Let us define the translation along the canonical vector e_i of $\mathbb{R}^m$ as :
$h \in \mathbb{R}, \tau_i\left(h\right) : \mathbb{R}^m \to \mathbb{R}^m :: \tau_i\left(h\right)\left(\sum_{\alpha=1}^m x^\alpha e_\alpha\right) = \sum_{\alpha=1}^m x^\alpha e_\alpha - he_i$

A distribution $S \in C_{\infty c}\left(\mathbb{R}^m; \mathbb{C}\right)$ is said to be independant with respect to the variable i if $\forall h : \tau_i\left(h\right)^* S = S \Leftrightarrow \frac{\partial S}{\partial x^i} = 0$

Principal value

1. Principal value of a function :
let $f : \mathbb{R} \to \mathbb{C}$ be some function such that :
either $\int_{-\infty}^{+\infty} f\left(x\right) dx$ is not defined, but $\exists \lim_{X \to \infty}\int_{-X}^{+X} f(x)dx$
or $\int_a^b f\left(x\right) dx$ is not defined, but $\exists \lim_{\varepsilon \to 0}\left(\int_a^{c-\varepsilon} f(x)dx + \int_{c+\varepsilon}^b f(x)dx\right)$
then we define the principal value integral : $\mathrm{pv}\int f(x)dx$ as the limit

2. The **principal value** distribution is defined as $pv\left(\frac{1}{x}\right) \in C_{\infty c}\left(\mathbb{R};\mathbb{C}\right)'$ ::
$pv\left(\frac{1}{x}\right)(\varphi) = \lim_{\varepsilon \to 0}\left(\int_{-\infty}^{-\varepsilon}\frac{f(x)}{x}dx + \int_{+\varepsilon}^{\infty}\frac{f(x)}{x}dx\right)$

If $S \in C_{\infty c}\left(\mathbb{R};\mathbb{C}\right)' : xS = 1$ then $S = pv\left(\frac{1}{x}\right) + C\delta_0$ and we have : $\frac{d}{dx}T\left(\ln|x|\right) = pv\left(\frac{1}{x}\right)$

3. It can be generalized as follows

(Taylor 1 p.241): take $f \in C_{\infty}\left(S^{m-1};\mathbb{C}\right)$ where S^{m-1} is the unit sphere of $\mathbb{R}^m$ such that : $\int_{S^{m-1}} f\varpi_0 = 0$ with the Lebesgue volume form on S^{m-1}. Define : $f_n\left(x\right) = \|x\|^{-n} f\left(\frac{x}{\|x\|}\right), x \neq 0, n \geq -m$

then : $T\left(f_n\right) : T\left(f_n\right)(\varphi) = \int_{S^{m-1}} f_n\varphi\varpi_0$ is a homogeneous distribution in S'($\mathbb{R}^m$) called the principal value of $f_n : T\left(f_n\right) = pv\left(f_n\right)$

28.7.12   Distributions depending on a parameter

Definition

1. Let V be Fréchet space of functions on some manifold M, J an open interval in $\mathbb{R}$ and a map : $S : J \to F'$ :: $S\left(t\right)$ is a family of distributions acting on V.

2. The result of the action on V is a function $S\left(t\right)(\varphi)$ on J. So we can consider the map : $\widetilde{S} : J \times V \to \mathbb{C}$ :: $S\left(t\right)(\varphi)$ and the existence of a partial derivative $\frac{\partial}{\partial t}$ ot this map with respect to the parameter t. We say that S is of class r if : $\forall \varphi \in V : \widetilde{S}\left(.\right)(\varphi) \in C_r\left(J;\mathbb{C}\right)$ that we denote : $S \in C_r\left(J;V'\right)$

3. We can go further and consider families of functions depending on a parameter : $u \in C\left(J;V\right)$ so that $u\left(t\right) \in V$.

If we have on one hand a family of functions $u \in C\left(J;V\right)$ and on the other hand a family of distributions $S \in C\left(J;V'\right)$ we can consider quantities such that : $S\left(t\right)u(t) \in \mathbb{C}$ and the derivative with respect to t of the scalar map : $J \to \mathbb{C}$ :: $S\left(t\right)u(t)$. For this we need a partial derivative of u with respect to t, which assumes that the space V is a normed vector space.

If $V \subset C\left(E;\mathbb{C}\right)$ we can also consider $u \in C\left(E \times J;\mathbb{C}\right)$ and distributions $S \in C\left(E \times J;\mathbb{C}\right)'$ if we have some way to agregate the quantities $S\left(t\right)u(t) \in \mathbb{C}$.

Family of distributions in $C_{\infty c}\left(O;\mathbb{C}\right)'$ If there is a partial derivative $\frac{\partial \widetilde{S}}{\partial t}$ we can hope that it is also a distribution.

Theorem 2286 *(Zuily p.61) Let O be an open of* $\mathbb{R}^m$, *J be an open interval in* $\mathbb{R}$, *if* $S \in C_r\left(J;C_{\infty c}\left(O;\mathbb{C}\right)'\right)$ *then :*
$\forall s : 0 \leq s \leq r : \exists S_s \in C_{r-s}\left(J;C_{\infty c}\left(O;\mathbb{C}\right)'\right) : \forall \varphi \in C_{\infty c}\left(O;\mathbb{C}\right) : \left(\frac{d}{dt}\right)^s \left(S\left(t\right)(\varphi)\right) = S_s\left(t\right)(\varphi)$

Notice that the distribution S_s is not the derivative of a distribution on $J \times \mathbb{R}^m$, even if is common to denote $S_s\left(t\right) = \left(\frac{\partial}{\partial t}\right)^s S$

The theorem still holds if : $S \in C\left(J;C_{\infty c}\left(O;\mathbb{C}\right)'_c\right)$ and $u \in C\left(J;C_{\infty}\left(O;\mathbb{C}\right)\right)$

We have, at least in the most simple cases, the chain rule :

Theorem 2287 *(Zuily p.61) Let O be an open of* $\mathbb{R}^m$, *J be an open interval in* $\mathbb{R}$, *$S \in C_1\left(J;C_{\infty c}\left(O;\mathbb{C}\right)'\right), u \in C_1\left(J;C_{\infty c}\left(O;\mathbb{C}\right)\right)$ then :*

$$\frac{d}{dt}\left(S\left(t\right)\left(u\left(t\right)\right)\right) = \frac{\partial S}{\partial t}\left(u\left(t\right)\right) + S\left(t\right)\left(\frac{\partial u}{\partial t}\right)$$

$\frac{\partial S}{\partial t}$ is the usual derivative of the map $S : J \to C_{\infty c}\left(O;\mathbb{C}\right)'$

The theorem still holds if : $S \in C\left(J; C_{\infty c}\left(O;\mathbb{C}\right)'_c\right)$ and $u \in C\left(J; C_{\infty}\left(O;\mathbb{C}\right)\right)$

Conversely we can consider the integral of the scalar map : $J \to \mathbb{C} ::$ $S\left(t\right) u(t)$ with respect to t

Theorem 2288 *(Zuily p.62) Let O be an open of $\mathbb{R}^m$, J be an open interval in $\mathbb{R}$, $S \in C_0\left(J; C_{\infty c}\left(O;\mathbb{C}\right)'\right), u \in C_{c\infty}\left(J \times O;\mathbb{C}\right)$ then the map : $\widehat{S}$: $C_{\infty c}\left(J \times O;\mathbb{C}\right) \to \mathbb{C} :: \widehat{S}\left(u\right) = \int_J \left(S\left(t\right) u\left(t, .\right)\right) dt$ is a distribution $\widehat{S} : C_{\infty c}\left(J \times O;\mathbb{C}\right)'$*

The theorem still holds if : $S \in C_0\left(J; C_{\infty c}\left(O;\mathbb{C}\right)'_c\right)$ and $u \in C_{\infty}\left(J \times O;\mathbb{C}\right)$

Theorem 2289 *(Zuily p.77) Let J be an open interval in $\mathbb{R}$, $S \in C_r\left(J; C_{\infty c}\left(\mathbb{R}^m;\mathbb{C}\right)'_c\right), U \in C_{\infty c}\left(\mathbb{R}^m;\mathbb{C}\right)'$ then the convolution : $S\left(t\right){*}U \in C_r\left(J; C_{\infty c}\left(\mathbb{R}^m;\mathbb{C}\right)'\right)$ and $\forall s, 0 \leq s \leq r : \left(\frac{\partial}{\partial t}\right)^s\left(S\left(t\right) * U\right) = \left(\left(\frac{\partial}{\partial t}\right)^s S\right) * U$*

Theorem 2290 *(Zuily p.77) Let J be an open interval in $\mathbb{R}$, $S \in C_r\left(J; C_{\infty c}\left(\mathbb{R}^m;\mathbb{C}\right)'_c\right), \varphi \in C_{\infty}\left(\mathbb{R}^m;\mathbb{C}\right)$ then :*
$$S\left(t\right){*}T\left(\varphi\right) = T\left(S\left(t\right)_y\left(\varphi\left(x - y\right)\right)\right) \text{ with } S\left(t\right)_y\left(\varphi\left(x - y\right)\right) \in C_r\left(J \times \mathbb{R}^m;\mathbb{C}\right)$$

28.7.13 Distributions and Sobolev spaces

This is the first extension of Sobolev spaces, to distributions.

Dual of Sobolev's spaces on $\mathbb{R}^m$

Definition 2291 *The Sobolev space, denoted $H^{-r}\left(O\right), r \geq 1$ is the topological dual of the closure $\overline{C_{\infty c}\left(O;\mathbb{C}\right)}$ of $C_{\infty c}\left(O;\mathbb{C}\right)$ in $H^r_c\left(O\right)$ where O is an open subset of $\mathbb{R}^m$*

Theorem 2292 *(Zuily p.87) $H^{-r}\left(O\right)$ is a vector subspace of $C_{\infty c}\left(O;\mathbb{C}\right)'$ which can be identified with :*
the topological dual $\left(H^r_c\left(O\right)\right)'$ of $H^r_c\left(O\right)$
the space of distributions : $\left\{S \in C_{\infty c}\left(M;\mathbb{C}\right)'; S = \sum_{s=0}^r \sum_{\alpha_1...\alpha_s} D_{\alpha_1...\alpha_s} T\left(f_{\alpha_1...\alpha_s}\right), f_{\alpha_1...\alpha_s} \in L^2\left(O, dx, \mathbb{C}\right.\right.$
This is a Hilbert space with the norm : $\|S\|_{H^{-r}} = \inf_{f_{(\alpha)}}\left(\sum_{s=0}^r \sum_{(\alpha)} \|f_{(\alpha)}\|_{L^2}\right)^{1/2}$

Theorem 2293 *(Zuily p.88) $\forall \varphi \in L^2\left(O, dx, \mathbb{C}\right) : \psi \in H^r_c\left(O\right) : \left|\int_M \overline{\varphi}\psi\mu\right| \leq \|\varphi\|_{H^{-r}}\|\psi\|_{H^r}$*

The generalization is stated for compact manifolds in (Taylor 1 p.282) .

Sobolev's inequalities

This is a collection of inequalities which show that many properties of functions are dictated by their first order derivative.

Theorem 2294 *(Zuily p.89) Poincaré inequality : For any subset O open in $\mathbb{R}^m$ with diameter $d<\infty$: $\forall \varphi \in H_c^1(O) : \|\varphi\|_{L^2} \le 2d \sum_{\alpha=1}^m \left\| \frac{\partial \varphi}{\partial x^\alpha} \right\|_{L^2}$*

Theorem 2295 *(Zuily p.89) If O is bounded in $\mathbb{R}^m$ the quantity $\sum_{\alpha_1 \dots \alpha_r} \|D_{\alpha_1 \dots \alpha_r}\varphi\|_{L^2}$ is a norm on $H_c^r(O)$ equivalent to $\|\|_{H^r(O)}$*

For the following theorems we define the spaces of functions on $\mathbb{R}^m$:
$D^1(\mathbb{R}^m) = \left\{ f \in L_{loc}^1(\mathbb{R}^m, dx, \mathbb{C}) \cap C_\nu(\mathbb{R}^m; \mathbb{C}), \partial_\alpha T(f) \in T\left(L^2(\mathbb{R}^m, dx, \mathbb{C})\right) \right\}$
$D^{1/2}(\mathbb{R}^m) = \left\{ f \in L_{loc}^1(\mathbb{R}^m, dx, \mathbb{C}) \cap C_\nu(\mathbb{R}^m; \mathbb{C}), \int_{\mathbb{R}^{2m}} \frac{|f(x)-f(y)|^2}{\|x-y\|^{m+1}} dxdy < \infty \right\}$

Theorem 2296 *(Lieb p.204) For $m>2$ If $f \in D^1(\mathbb{R}^m)$ then $f \in L^q(\mathbb{R}^m, dx, \mathbb{C})$ with $q = \frac{2m}{m-2}$ and :*
$\|f'\|_2^2 \ge C_m \|f\|_q^2$ *with $C_m = \frac{m(m-2)}{4} A(S^m)^{2/m} = \frac{m(m-2)}{4} 2^{2/m} m^{1+\frac{1}{m}} \Gamma\left(\frac{m+1}{2}\right)^{-2/m}$*
The equality holds iff f is a multiple of $\left(\mu^2 + \|x-a\|^2\right)^{-\frac{m-2}{2}}$,$\mu > 0, a \in \mathbb{R}^m$

Theorem 2297 *(Lieb p.206) For $m>1$ If $f \in D^{1/2}(\mathbb{R}^m)$ then $f \in L^q(\mathbb{R}^m, dx, \mathbb{C})$ with $q = \frac{2m}{m-2}$ and :*
$\frac{\Gamma\left(\frac{m+1}{2}\right)}{2\pi^{\frac{m+1}{2}}} \int_{\mathbb{R}^{2m}} \frac{|f(x)-f(y)|^2}{\|x-y\|^{m+1}} dxdy \ge \frac{m-1}{2}(C_m)^{1/m} \|f\|_q^2$ *with $C_m = \frac{m-1}{2} 2^{1/m} \pi^{\frac{m+1}{2m}} \Gamma\left(\frac{m+1}{2}\right)^{-1/m}$*
The equality holds iff f is a multiple of $\left(\mu^2 + \|x-a\|^2\right)^{-\frac{m-1}{2}}$,$\mu > 0, a \in \mathbb{R}^m$

The inequality reads :
$\int_{\mathbb{R}^{2m}} \frac{|f(x)-f(y)|^2}{\|x-y\|^{m+1}} dxdy \ge C_m' \|f\|_q^2$ with $C_m' = \left(\frac{m-1}{2}\right)^{1+\frac{1}{m}} \left(2\pi^{\frac{m+1}{2}} \Gamma\left(\frac{m+1}{2}\right)^{-1}\right)^{1+\frac{1}{m^2}}$

Theorem 2298 *(Lieb p.210) Let (f_n) a sequence of functions $f_n \in D^1(\mathbb{R}^m), m > 2$ such that $f_n' \to g$ weakly in $L^2(\mathbb{R}^m, dx, \mathbb{C})$ then :*
i) $g = f'$ for some unique function $f \in D^1(\mathbb{R}^m)$.
ii) there is a subsequence which converges to f for almost every $x \in \mathbb{R}^m$
iii) If A is a set of finite measure in $\mathbb{R}^m$ then $1_A f_n \to 1_A f$ strongly in $L^p(\mathbb{R}^m, dx, \mathbb{C})$ for $p < \frac{2m}{m-2}$

Theorem 2299 *Nash's inequality (Lieb p.222) For every function $f \in H^1(\mathbb{R}^m) \cap L^1(\mathbb{R}^m, dx, \mathbb{C})$:*
$\|f\|_2^{1+\frac{2}{m}} \le C_m \|f'\|_2 \|f\|_1^{2/m}$ *where $C_m^2 = 2m^{-1+\frac{2}{m}} \left(1+\frac{m}{2}\right)^{1+\frac{2}{m}} \lambda(m)^{-1} A(S^{m-1})^{-2/m}$ and $\lambda(m)$ is a constant which depends only on m*

Theorem 2300 *Logarithmic Sobolev inequality (Lieb p.225) Let $f \in H^1\left(\mathbb{R}^m\right), a > 0$ then :*

$\frac{a^2}{m}\int_{\mathbb{R}^m}\|f'\|^2\,dx \geq \int_{\mathbb{R}^m}|f|^2\ln\left(\frac{|f|^2}{\|f\|_2^2}\right)dx + m\left(1 + \ln a\right)\|f\|_2^2$

The equality holds iff f is, up to translation, a multiple of $\exp\left(-\pi\frac{\|x\|^2}{2a^2}\right)$

Theorem 2301 *(Lieb p.229) Let (E,S,μ) be a $\sigma-$finite measured space with μ a positive measure, $U(t)$ a one parameter semi group which is also symmetric and a contraction on both $L^p\left(E, S, \mu, \mathbb{C}\right), p = 1, 2,$ with infinitesimal generator S whose domain is $D(S), \gamma \in [0, 1]$ a fixed scalar. Then the following are equivalent:*

i) $\exists C_1\left(\gamma\right) > 0 : \forall f \in L^1\left(E, S, \mu, \mathbb{C}\right) : \|U\left(t\right)f\|_\infty \leq C_1 t^{-\frac{\gamma}{1-\gamma}}\|f\|_1$

ii) $\exists C_2\left(\gamma\right) > 0 : \forall f \in L^1\left(E, S, \mu, \mathbb{C}\right)\cap D\left(S\right) : \|f\|_2^2 \leq C_2\left(\langle f, Sf\rangle\right)^\gamma\|f\|_1^{2(1-\gamma)}$

U(t) must satisfy : $, U(0)f = f, U(t+s) = U(t)\circ U(s), \|U\left(t\right)f - U\left(s\right)f\|_p \to_{t\to s} 0$

$\|U(t)f\|_p \leq \|f\|_p, \langle f, U\left(t\right)g\rangle = \langle U(t)f, g\rangle$

28.8 Extension of distributions

28.8.1 Colombeau algebras

Colombeau algebras

It would be nice if we had distributions $S\left(\varphi \times \psi\right) = S\left(\varphi\right)S\left(\psi\right)$ meaning S is a multiplicative linear functional in V', so $S \in \Delta\left(V\right)$. But if E is a locally compact Hausdorff space, then $\Delta\left(C_{0v}\left(E; \mathbb{C}\right)\right)$ is homeomorphic to E and the only multiplicative functionals are the Dirac distributions : $\delta_a\left(\varphi\psi\right) = \delta_a\left(\varphi\right)\delta_a\left(\psi\right)$. This is the basis of a celebrated "no go" theorem by L.Schwartz stating that there is no possibility to define the product of distributions. However there are ways around this issue, by defining quotient spaces of functions and to build "Colombeau algebras" of distributions. The solutions are very technical. See Nigsch on this topic.They usually involved distributions acting on m forms, of which we give an overview below as it is a kind of introduction to the following subsection.

Distributions acting on m forms

1.The general idea is to define distributions acting, not on functions, but on m forms on a m dimensional manifold. To do this the procedure is quite similar to the one used to define Lebesgue integral of forms. It is used, and well defined, with $V = C_{\infty c}\left(M; \mathbb{C}\right)$.

2. Let M be a m dimensional, Hausdorff, smooth, real, orientable manifold, with atlas $\left(O_i, \varphi_i\right)_{i\in I}, \varphi_i\left(O_i\right) = U_i \in \mathbb{R}^m$

A m form ϖ on M has for component in the holonomic basis $\varpi_i \in C\left(O_i; \mathbb{R}\right)$ with the rule in a transition between charts : $\varphi_{ij} : O_i \to O_j :: \varpi_j = \det\left[\varphi'_{ij}\right]^{-1}\varpi_i$

The push forward of ϖ by each chart gives a m form on $\mathbb{R}^m$ whose components in each domain U_i is equal to ϖ_i :

$\varphi_{i*}\varpi\left(\xi\right) = \varpi_i\left(\varphi_i^{-1}\left(\xi\right)\right)e^1 \wedge ... \wedge e^m$ in the canonical basis $\left(e^k\right)_{k=1}^m$ of $\mathbb{R}^{m*}$

If each of the functions ϖ_i is smooth and compactly supported : $\varpi_i \in C_{\infty c}(O_i; \mathbb{R})$ we will say that ϖ is smooth and compactly supported and denote the space of such forms : $\mathfrak{X}_{\infty c}(\wedge_m TM^*)$. Obviously the definition does not depend of the choice of an atlas.

Let $S_i \in C_{\infty c}(U_i; \mathbb{R})'$ then $S_i(\varpi_i)$ is well defined.

Given the manifold M, we can associate to $\mathfrak{X}_{\infty c}(\wedge_m TM^*)$ and any atlas, families of functions $(\varpi_i)_{i \in I}, \varpi_i \in C_{\infty c}(U_i; \mathbb{R})$ meeting the transition properties. As M is orientable, there is an atlas such that for the transition maps $\det\left[\varphi'_{ij}\right] > 0$

Given a family $(S_i)_{i \in I} \in \left(C_{\infty c}(U_i; \mathbb{R})'\right)^I$ when it acts on a family $(\varpi_i)_{i \in I}$, representative of a m form, we have at the intersection $S_i|_{U_i \cap U_j} = S_j|_{U_i \cap U_j}$, so there is a unique distribution $\widehat{S} \in C_{\infty c}(U; \mathbb{R})', U = \cup_{i \in I} U_i$ such that $\widehat{S}|_{U_i} = S_i$.

We define the pull back of $\widehat{S}$ on M by the charts and we have a distribution S on M with the property :

$\forall \psi \in C_{\infty c}(O_i; \mathbb{R}) : \left(\varphi_i^{-1}\right)^* S(\psi) = \widehat{S}|_{U_i}(\psi)$

and we say that S$\in \mathfrak{X}_{\infty c}(\wedge_m TM^*)'$ and denote $S(\varpi) = \widehat{S}(\varphi_* \varpi)$

On the practical side the distribution of a m form is just the sum of the value of distributions on each domain, applied to the component of the form, with an adequate choice of the domains.

3. The pointwise product of a smooth function f and a m form $\varpi \in (\wedge_m TM^*)_{\infty c}$ is still a form f$\varpi \in \mathfrak{X}_{\infty c}(\wedge_m TM^*)$ so we can define the product of such a function with a distribution. Or if $\varpi \in \mathfrak{X}_\infty(\wedge_m TM^*)$ is a smooth m form, and $\varphi \in C_{\infty c}(M; \mathbb{R})$ then $\varphi\varpi \in \mathfrak{X}_{\infty c}(\wedge_m TM^*)$.

4. The derivative takes a different approach. If X a is smooth vector field on M, then the Lie derivative of a distribution S$\in (\mathfrak{X}_{\infty c}(\wedge_m TM^*))'$ is defined as :

$\forall X \in \mathfrak{X}_\infty(TM), \varpi \in \mathfrak{X}_{\infty c}(\wedge_m TM^*) : (\pounds_X S)(\varpi) = -S(\pounds_X \varpi) = -S(d(i_X \varpi))$

If there is a smooth volume form ϖ_0 on M, and $\varphi \in C_{\infty c}(M; \mathbb{R})$ then

$(\pounds_X S)(\varphi\varpi_0) = -S((div X)\varphi\varpi_0) = -(div X) S(\varphi\varpi_0)$

If M is manifold with boundary we have some distribution :

$S(d(i_X \varpi_0)) = S_{\partial M}(i_X \varpi_0) = -(\pounds_X S)(\varpi_0)$

28.8.2 Distributions on vector bundles

In fact the construct above can be generalized to any vector bundle in a more convenient way, but we take the converse of the previous idea : distributions act on sections of vector bundle, and they can be assimilated to m forms valued on the dual vector bundle.

Definition

If $E(M, V, \pi)$ is a complex vector bundle, then the spaces of sections $\mathfrak{X}_r(E)$, $\mathfrak{X}(J^r E), \mathfrak{X}_{rc}(E), \mathfrak{X}_c(J^r E)$ are Fréchet spaces. Thus we can consider continuous linear functionals on them and extend the scope of distributions.

The space of test functions is here $\mathfrak{X}_{\infty, c}(E)$. We assume that the base manifold M is m dimensional and V is n dimensional.

Definition 2302 *A distribution on the vector bundle E is a linear, continuous functional on $\mathfrak{X}_{\infty,c}(E)$*

Notation 2303 $\mathfrak{X}_{\infty,c}(E)'$ *is the space of distributions on E*

Several tools for scalar distributions can be extended to vector bundle distributions.

Product of a distribution and a section

Definition 2304 *The product of distribution $S \in \mathfrak{X}_{\infty,c}(E)'$ by a function $f \in C_{\infty}(M;\mathbb{C})$ is the distribution $fS \in \mathfrak{X}_{\infty,c}(E)' :: (fS)(X) = S(fX)$*

The product of a section $X \in \mathfrak{X}_{\infty,c}(E)$ by a function $f \in C_{\infty}(M;\mathbb{C})$ still belongs to $\mathfrak{X}_{\infty,c}(E)$

Assimilation to forms on the dual bundle

Theorem 2305 *To each continuous m form $\lambda \in \Lambda_m(M;E')$ valued in the topological bundle E' can be associated a distribution in $\mathfrak{X}_{\infty,c}(E)'$*

Proof. The dual E' of $E(M,V,\pi)$ is the vector bundle $E'(M,V',\pi)$ where V' is the topological dual of V (it is a Banach space). A m form $\lambda \in \Lambda_m(M;E')$ reads in a holonomic basis $\left(e_a^i(x)\right)_{i \in I}$ of E' and $(d\xi^{\alpha})_{\alpha=1}^m$ of TM* : $\lambda = \sum \lambda_i(x) e_a^i(x) \otimes d\xi^1 \wedge ... \wedge d\xi^m$

It acts fiberwise on a section $X \in \mathfrak{X}_{\infty,c}(E)$ by : $\sum_{i \in I} X^i(x) \lambda_i(x) d\xi^1 \wedge ... \wedge d\xi^m$

With an atlas $(O_a, \psi_a)_{a \in A}$ of M and $(O_a, \varphi_a)_{a \in A}$ of E, at the transitions (both on M and E) (see Vector bundles) :

$\lambda_b^i = \sum_j \det[\psi'_{ba}] [\varphi_{ab}]_i^j \lambda_a^j$

$X_b^i = \sum_{j \in I} [\varphi_{ba}]_j^i X_a^j$

$\mu_b = \sum_{i \in I} X_b^i(x) \lambda_{bi}(x) = \det[\psi'_{ba}] \sum_{i \in I} [\varphi_{ba}]_i^j \lambda_a^j [\varphi_{ab}]_k^i X_a^k = \det[\psi'_{ba}] \sum_{i \in I} \lambda_a^i X_a^i$

So : $\mu = \sum_{i \in I} X^i(x) \lambda_i(x) d\xi^1 \wedge ... \wedge d\xi^m$ is a m form on M and the actions reads :

$\mathfrak{X}_0(E') \times \mathfrak{X}_{\infty,c}(E) \to \Lambda_m(M;\mathbb{C}) :: \mu = \sum_{i \in I} X^i(x) \lambda_i(x) d\xi^1 \wedge ... \wedge d\xi^m$

μ defines a Radon measure on M, locally finite. And because X is compactly supported and bounded and λ continuous the integral of the form on M is finite. $\blacksquare$

We will denote T the map : $T:\Lambda_{0,m}(M;E') \to \mathfrak{X}_{\infty,c}(E)'::T(\lambda)(X) = \int_M \lambda(X)$.

Notice that the map is not surjective. Indeed the interest of the concept of distribution on vector bundle is that this not a local operator but a global operator, which acts on sections. The m forms of $\Lambda_m(M;E')$ can be seen as the local distributions.

Pull back, push forward of a distribution

Definition 2306 *For two smooth complex finite dimensional vector bundles* $E_1\left(M,V_1,\pi_1\right), E_2\left(M,V_2,\pi_2\right)$ *on the same real manifold and a base preserving morphism* $L:E_1\rightarrow E_2$ *the pull back of a distribution is the map :* $L^*:$ $\mathfrak{X}_{\infty,c}\left(E_2\right)'\rightarrow \mathfrak{X}_{\infty,c}\left(E_1\right)'::L^*S_2\left(X_1\right)=S_2\left(LX_1\right)$

Theorem 2307 *In the same conditions, there is a map :* $L^{t*}:\Lambda_m\left(M;E_2'\right)\rightarrow$ $\Lambda_m\left(M;E_1'\right)$ *such that :* $L^*T\left(\lambda_2\right)=T\left(L^{t*}\left(\lambda_2\right)\right)$

Proof. The transpose of $L:E_1\rightarrow E_2$ is $L^t:E_2'\rightarrow E_1'$. This is a base preserving morphism such that : $\forall X_1\in\mathfrak{X}\left(E_1\right), Y_2\in\mathfrak{X}\left(E_1'\right):L^t\left(Y_2\right)\left(X_1\right)=Y_2\left(LX_1\right)$

It acts on $\Lambda_m\left(M;E_2'\right):L^{t*}:\Lambda_m\left(M;E_2'\right)\rightarrow\Lambda_m\left(M;E_1'\right)::L^{t*}\left(\lambda_2\right)\left(x\right)=$ $L^t\left(x\right)\left(\lambda_2\right)\otimes d\xi^1\wedge...\wedge d\xi^m$

If $S_2=T\left(\lambda_2\right)$ then $L^*S_2\left(X_1\right)=S_2\left(LX_1\right)=T\left(\lambda_2\right)\left(LX_1\right)=\int_M\lambda_2\left(LX_1\right)=$ $\int_M L^t\left(\lambda_2\right)\left(X_1\right)=T\left(L^{t*}\left(\lambda_2\right)\right)\left(X_1\right)$ ■

Similarly :

Definition 2308 *For two smooth complex finite dimensional vector bundles* $E_1\left(M,V_1,\pi_1\right), E_2\left(M,V_2,\pi_2\right)$ *on the same real manifold and a base preserving morphism* $L:E_2\rightarrow E_1$ *the push forward of a distribution is the map :* $L_*:$ $\mathfrak{X}_{\infty,c}\left(E_1\right)'\rightarrow\mathfrak{X}_{\infty,c}\left(E_2\right)'::L_*S_1\left(X_2\right)=S_1\left(LX_2\right)$

Support of a distribution

The support of a section X is defined as the complementary of the largest open O in M such as : $\forall x\in O:X\left(x\right)=0.$

Definition 2309 *The support of a distribution* S *in* $\mathfrak{X}_{\infty,c}\left(E\right)'$ *is the subset of M, complementary of the largest open O in M such as :* $\forall X\in\mathfrak{X}_{\infty,c}\left(E\right), Supp(X)\subset$ $O\Rightarrow S\left(X\right)=0.$

r jet extension of a vectorial distribution

Rather than the derivative of a distribution it is more logical to look for its r jet prolongation. But this needs some adjustments.

Definition 2310 *The action of a distribution* $S\in\mathfrak{X}_{\infty,c}\left(E\right)'$ *on a section* $Z\in$ $\mathfrak{X}_{\infty,c}\left(J^rE\right)$ *is defined as the map :*

$S:\mathfrak{X}_{\infty,c}\left(J^rE\right)\rightarrow\left(\mathbb{C},\mathbb{C}^{sm},s=1..r\right)::$
$S\left(Z_{\alpha_1...\alpha_s}\left(x\right), s=0..r, \alpha_j=1...m\right)=\left(S\left(Z_{\alpha_1...\alpha_s}\right), s=0..r, \alpha_j=1...m\right)$

The r jet prolongation of the vector bundle $E\left(M,V,\pi\right)$ is the vector bundle $J^rE\left(M,J_0^r\left(\mathbb{R}^m,V\right)_0,\pi_0^r\right).$A section $Z\in\mathfrak{X}_{\infty,c}\left(J^rE\right)$ can be seen as the set $\left(Z_{\alpha_1...\alpha_s}\left(x\right), s=0..r, \alpha_j=1...m\right)$ where $Z_{\alpha_1...\alpha_s}\left(x\right)\in E\left(x\right)$

The result is not a scalar, but a set of scalars : $\left(k_{\alpha_1...\alpha_s}, s=0..r, \alpha_j=1...m\right)$

Definition 2311 *The r jet prolongation of a distribution $S \in \mathfrak{X}_{\infty,c}(E)'$ is the map denoted $J^r S$ such that $\forall X \in \mathfrak{X}_{\infty,c}(E) : J^r S(X) = S(J^r X)$*

$J^r E$ is a vector bundle, and the space of its sections, smooth and compactly supported $\mathfrak{X}_{\infty,c}(J^r E)$, is well defined, as the space of its distributions. If $X \in \mathfrak{X}_{\infty,c}(E)$ then $J^r X \in \mathfrak{X}_{\infty,c}(J^r E)$:

$J^r X = \left(X, \sum_{i=1}^n \left(D_{\alpha_1..\alpha_s} X^i\right) e_i(x), i=1..n, s=1..r, \alpha_j = 1...m\right)$

So the action of $S \in \mathfrak{X}_{\infty,c}(E)'$ is :

$S(J^r X) = \left(S\left(D_{\alpha_1..\alpha_s} X\right), s = 0...r, \alpha_j = 1...m\right)$

and this leads to the definition of $J^r S$:

$J^r S : \mathfrak{X}_{\infty,c}(E) \to \left(\mathbb{C}, \mathbb{C}^{sm}, s = 1..r\right) :: J^r S(X) = \left(S\left(D_{\alpha_1..\alpha_s} X\right), s = 0...r, \alpha_j = 1...m\right)$

Definition 2312 *The derivative of a distribution $S \in \mathfrak{X}_{\infty,c}(E)'$ with respect to $(\xi^{\alpha_1}, ..., \xi^{\alpha_s})$ on M is the distribution : $\left(D_{\alpha_1..\alpha_s} S\right) \in \mathfrak{X}_{\infty,c}(E)'$: $\forall X \in \mathfrak{X}_{\infty,c}(E)$: $\left(D_{\alpha_1..\alpha_s} S\right)(X) = S\left(D_{\alpha_1..\alpha_s} X\right)$*

The map : $\left(D_{\alpha_1..\alpha_s} S\right) : \mathfrak{X}_{\infty,c}(E) \to \mathbb{C} :: \left(D_{\alpha_1..\alpha_s} S\right)(X) = S\left(D_{\alpha_1..\alpha_s} X\right)$ is valued in $\mathbb{C}$, is linear and continuous, so it is a distribution.

Theorem 2313 *The r jet prolongation of a distribution $S \in \mathfrak{X}_{\infty,c}(E)'$ is the set of distributions :$J^r S = \left(\left(D_{\alpha_1..\alpha_s} S\right), s = 0..r, \alpha_j = 1...m\right)$*

The space $\left\{J^r S, S \in \mathfrak{X}_{\infty,c}(E)'\right\}$ is a vector subspace of $\left(\mathfrak{X}_{\infty,c}(E)'\right)^N$ where N$=\sum_{s=0}^r \frac{m!}{s!(m-s)!}$

Theorem 2314 *If the distribution $S \in \mathfrak{X}_{\infty,c}(E)'$ is induced by the m form $\lambda \in \Lambda_m(M; E')$ then its r jet prolongation : $J^r(T(\lambda)) = T(J^r \lambda)$ with $J^r \lambda \in \Lambda_m(M; J^r E')$*

Proof. The space $\Lambda_m(M; E') \simeq E' \otimes (TM^*)^m$ and its r jet prolongation is : $J^r E' \otimes (TM^*)^m \simeq \Lambda_m(M; J^r E')$

$\left(D_{\alpha_1..\alpha_s} T(\lambda)\right)(X) = T(\lambda)\left(D_{\alpha_1..\alpha_s} X\right) = \int_M \lambda\left(D_{\alpha_1..\alpha_s} X\right)$

The demonstration is similar to the one for derivatives of distributions on a manifold.

We start with r=1 and $D_\alpha = \partial_\alpha$ with $\alpha \in 1...m$ and denote: $d\xi = d\xi^1 \wedge ... \wedge d\xi^m$

$i_{\partial_\alpha}(\lambda_i d\xi) = (-1)^{\alpha-1} \lambda_i d\xi^1 \wedge ... \wedge \left(\widehat{d\xi^\alpha}\right) \wedge .. \wedge d\xi^m$

$d\left(i_{\partial_\alpha}(\lambda_i d\xi)\right) = \sum_\beta (-1)^{\alpha-1}(\partial_\beta \lambda_i) d\xi^\beta \wedge d\xi^1 \wedge ... \wedge \left(\widehat{d\xi^\alpha}\right) \wedge .. \wedge d\xi^m = (\partial_\alpha \lambda_i) d\xi$

$\left(dX^i\right) \wedge i_{\partial_\alpha}(\lambda_i d\xi) = \sum_\beta (-1)^{\alpha-1}\left(\partial_\beta X^i\right) \lambda_i d\xi^\beta \wedge d\xi^1 \wedge ... \wedge \left(\widehat{d\xi^\alpha}\right) \wedge .. \wedge d\xi^m = \left(\partial_\beta X^i\right) \lambda_i d\xi$

$d\left(X^i i_{\partial_\alpha}(\lambda_i d\xi)\right) = dX^i \wedge i_{\partial_\alpha}(\lambda_i d\xi) - X^i d\left(i_{\partial_\alpha}(\lambda_i d\xi)\right) = \left(\partial_\alpha X^i\right)(\lambda_i d\xi) - X^i(\partial_\alpha \lambda_i) d\xi$

$d \left(\sum_i X^i i_{\partial_\alpha} \lambda_i \right) = \sum_i \left(\left(\partial_\alpha X^i \right) \left(\lambda_i d\xi \right) - X^i \left(\partial_\alpha \lambda_i \right) d\xi \right)$

Let N be a compact manifold with boundary in M. The Stockes theorem gives :

$\int_N d \left(\sum_i X^i i_{\partial_\alpha} \lambda_i \right) = \int_{\partial N} \sum_i X^i i_{\partial_\alpha} \lambda_i = \int_N \sum_i \left(\left(\partial_\alpha X^i \right) \left(\lambda_i d\xi \right) - X^i \left(\partial_\alpha \lambda_i \right) d\xi \right)$

Because X has a compact support we can always choose N such that $Supp\left(X \right) \subset \overset{\circ}{N}$ and then :

$\int_{\partial N} \sum_i X^i i_{\partial_\alpha} \lambda_i = 0$

$\int_N \sum_i \left(\left(\partial_\alpha X^i \right) \left(\lambda_i d\xi \right) - X^i \left(\partial_\alpha \lambda_i \right) d\xi \right) = \int_M \sum_i \left(\left(\partial_\alpha X^i \right) \left(\lambda_i d\xi \right) - X^i \left(\partial_\alpha \lambda_i \right) d\xi \right) = 0$

$T \left(\partial_\alpha \lambda d\xi \right) \left(X \right) = T \left(\lambda \right) \left(\partial_\alpha X \right) = \partial_\alpha \left(T \left(\lambda \right) \right) \left(X \right)$

By recursion over r we get :

$T \left(D_{\alpha_1 \dots \alpha_r} \lambda \right) \left(X \right) = \left(D_{\alpha_1 \dots \alpha_r} T \left(\lambda \right) \right) \left(X \right)$

$J^r T \left(\lambda \right) = \left(\left(D_{\alpha_1 .. \alpha_s} T \left(\lambda \right) \right) \right), s = 0..r, \alpha_j = 1...m)$

$= \left(T \left(D_{\alpha_1 \dots \alpha_r} \lambda \right), s = 0..r, \alpha_j = 1...m \right) = T \left(J^r \lambda \right)$ ∎

Notice that we have gotten rid off the $(-1)^r$ in the process, which was motivated only for the consistency for functions in $\mathbb{R}^m$.

The space of complex functions over a manifold M can be seen as a vector bundle $(M, \mathbb{C}, \pi)$. So, in order to be avoid confusion:

- whenever we deal with functions defined in $\mathbb{R}^m$ we will refer to the usual definition of derivative of distribution and associated function;

- whenever we deal with complex functions defined over a manifold we will refer to the vector bundle concept, with associated m form and r jet extension.

29   THE FOURIER TRANSFORM

Fourier transforms, which work under the form of series or integrals, are a poweful tool in mathematics in that they convert differentiation or convolution into multiplications. To understand their common origin, and indeed to have straight proofs of the main theorems, it is useful to come back to the more general transformation on abelian groups, seen in the Lie group part.

29.0.3   Reminder of the general transformation

Pontryagin dual

1. The Pontryagin dual $\widehat{G}$ of a topological, Hausdorff, locally compact, connected, abelian group (G,+) is the set of continuous morphisms : $\chi : G \rightarrow T$ where (T,x) is the group of complex number of module 1 : $T = \{ z \in \mathbb{C} : |z| = 1 \}$ with multiplication. $\widehat{G}$ is a topological locally compact abelian group with pointwise multiplication.

2. Representations of the group :

Any finite dimensional irreducible unitary representation of G is of the form : $\left(T, \chi \in \widehat{G} \right)$ and the map $\chi \left(g \right) = \exp i\theta \left(g \right)$ is a character

There is a bijective correspondance between the continuous unitary representations (H,ρ) of G in a Hilbert space H and the regular spectral measure P on $\widehat{G}$:

$P : S \to \mathcal{L}(H;H) :: P(\chi) = P^2(\chi) = P^*(\chi)$

$\rho(g) = \int_{\widehat{G}} \chi(g) P(\chi)$

The support of P is the smallest closed subset A of $\widehat{G}$ such that P(A)=1

Fourier transform

G has a Haar measure μ, left and right invariant, which is a Radon measure, so it is regular and locally finite. It is unique up to a scaling factor.

1. The Fourier transform is the linear and continuous map : $\hat{} : L^1(G,\mu,\mathbb{C}) \to C_{0\nu}\left(\widehat{G};\mathbb{C}\right) : \widehat{f}(\chi) = \int_G f(g)\overline{\chi(g)}\mu(g)$

Conversely, for each Haar measure μ, there is a unique Haar measure ν on $\widehat{G}$ such that :

$\forall f \in L^1(G,\mu,\mathbb{C}) : f(g) = \int_{\widehat{G}} \widehat{f}(\chi)\chi(g)\nu(\chi)$ for almost all g. If f is continuous then the identity stands for all g in G.

If $f \in C_{0c}(G;\mathbb{C})$ the Fourier transform is well defined and :

$\widehat{f} \in L^2\left(\widehat{G},\nu,\mathbb{C}\right)$ and $\int_G |f(g)|^2 \mu(g) = \int_{\widehat{G}} \left|\widehat{f}(\chi)\right|^2 \nu(\chi) \Leftrightarrow \|f\|_2 = \left\|\widehat{f}\right\|_2$

$C_{0c}(G;\mathbb{C})$ is a Banach space for which the Fourier transform is well defined, and is dense in $L^2(G,\mu,\mathbb{C})$. From there we have :

2. There is a unique extension $\mathcal{F}$ of the Fourier transform to a continuous linear map : $\mathcal{F} : L^2(G,\mu,\mathbb{C}) \to L^2\left(\widehat{G},\nu,\mathbb{C}\right)$ which are both Hilbert spaces. Thus there is an adjoint map $\mathcal{F}^* : L^2\left(\widehat{G},\nu,\mathbb{C}\right) \to L^2(G,\mu,\mathbb{C})$ which is continuous such that : $\int_G \left(\overline{f_1(g)}\right)((\mathcal{F}^*f_2)(g))\mu(g) = \int_{\widehat{G}} \left(\overline{\mathcal{F}f_1(\chi)}\right)(f_2(\chi))\nu(\chi)$

If $f \in L^1(G,\mu,\mathbb{C}) \cap L^2(G,\mu,\mathbb{C})$ we have : $\mathcal{F}(f) = \widehat{f}$ and $f(g) = \int_{\widehat{G}} \widehat{f}(\chi)\chi(g)\nu(\chi)$ thus

$\mathcal{F}^* = \mathcal{F}^{-1} : \mathcal{F}^*(h)(g) = \int_{\widehat{G}} h(\chi)\chi(g)\nu(\chi)$

3. Convolution : $\forall f,g \in L^1(G,\mu,\mathbb{C}) : \widehat{f*g} = \widehat{f} \times \widehat{g}$

The only linear multiplicative functionals λ on the algebra $\left(L^1(G,\mu,\mathbb{C}),*\right)$: $\lambda(f*g) = \lambda(f)\lambda(f)$ are the functionals induced by the Fourier transform : $\lambda(\chi) : L^1(G,\mu,\mathbb{C}) \to \mathbb{C} :: \lambda(\chi)(f) = \widehat{f}(\chi)$

4.

If G is finite then $\widehat{G}$ is a finite abelian group.

If G is a Lie group over the field K then it is finite dimensional and isomorphic to $(K/\mathbb{Z})^m \times K^{m-p}$ with $p = \dim span \ker \exp$

So we have two basic cases

G is compact

G is a torus T^m isomorphic to the group $\exp(K/\mathbb{Z})^m$, $\widehat{G}$ is discrete and conversely if $\widehat{G}$ is discrete, then G is compact.

$\widehat{G}$ is $\chi(\theta) = \exp i\langle z,\theta\rangle$ with $\langle z,\theta\rangle = \sum_{k=1}^m z_k\theta_k, z_k \in \mathbb{Z}$ so this is discrete group indexed on $z \in \mathbb{Z}^m$.

μ is proportional to the Lebesgue measure on the torus T^m

The Fourier transform reads : $\widehat{f}(z) = (2\pi)^{-m} \int_{T^m} f(\theta) \exp(-i\langle z, \theta\rangle) \mu$ for z$\in \mathbb{Z}$

There is a unique Haar measure $\nu = \sum_{\mathbb{Z}^m} \delta_z$ on $\mathbb{Z}^m$ which gives an inversion formula and $f(\theta) = \sum_{z\in\mathbb{Z}^m} \widehat{f}(z) \exp i\langle z, \theta\rangle$

μ is finite, $L^2(G, \mu, \mathbb{C}) \subset L^1(G, \mu, \mathbb{C})$ so that the Fourier transform extends to $L^2(G, \mu, \mathbb{C})$ and the Fourier transform is a unitary operator on L^2.

$\widehat{G}$ is an orthonormal subset of $L^2(G, \mathbb{C}, \mu)$ which can be taken as a Hilbert basis. We have the decomposition :

$\varphi \in L^2(G, \mathbb{C}, \mu) : \varphi = \sum_\chi \left(\int_G \overline{\chi(g)} \varphi(g) \mu(g) \right) \chi = \sum_\chi \widehat{\varphi}\chi$

So to sum up : there is no complicated issue of definition of the Fourier transform and its inverse, and the inverse Fourier transform is a series which coefficients $\widehat{f}(z)$.

G is not compact

We can restrict ourselves to the case G is a m dimensional real vector space E.

There is a bijective correspondance between the dual E* of E and $\widehat{G}$: E* $\rightarrow$ $\widehat{G}$ :: $\chi(x) = \exp i\lambda(x)$ and $\lambda(x) = \sum_{k=1}^m t_k x^k$

μ is proportional to the Lebesgue measure. The Fourier transform reads:

$\widehat{} : L^1(E, \mu, \mathbb{C}) \rightarrow C_{0\nu}(E^*; \mathbb{C}) : \widehat{\varphi}(\lambda) = \int_E \varphi(x) \exp(-i\lambda(x)) \mu(x)$

It has an extension $\mathcal{F}$ on $L^2(E, \mu, \mathbb{C})$ which is a unitary operator, but $\mathcal{F}$ is usually not given by any explicit formula.

The Fourier transform has an inverse defined on specific cases.

29.1 Fourier series

This is the "compact case" which, for functional analysis purpose, deals with periodic functions.

29.1.1 Periodic functions

For a fixed in $\mathbb{R}^m$ the set : $\mathbb{Z}a = \{za, z \in \mathbb{Z}\}$ is a closed subgroup so the set $G = \mathbb{R}/\mathbb{Z}\mathfrak{a}$ is a commutative compact group. The classes of equivalence are : $x \sim y \Leftrightarrow x - y = za, z \in \mathbb{Z}$.

A periodic function $f : \mathbb{R}^m \rightarrow \mathbb{C}$ with period a$\in \mathbb{R}^m$ is a function such that : $\forall x \in \mathbb{R}^m : f(x + a) = f(x)$.

So we can define periodic functions as functions $\varphi : G = \mathbb{R}/\mathbb{Z}\mathfrak{a} \rightarrow \mathbb{C}$.

With the change of variable : $x = \frac{a}{2\pi}\theta$ it is customary and convenient to write :

$T^m = \{(\theta_k)_{k=1}^m, 0 \leq \theta_k \leq 2\pi\}$

and we are left with functions f$\in C(T^m; \mathbb{C})$ with the same period $(2\pi, ..., 2\pi)$ defined on T^m

In the following, we have to deal with two kinds of spaces :

i) The space of periodic functions f$\in C\left(T^{m};\mathbb{C}\right)$: the usual spaces of functions, with domain in T^{m}, valued in $\mathbb{C}$. The Haar measure is proportional to the Lebesgue measure $d\theta = (\otimes d\xi)^{m}$ meaning that we integrate on $\xi \in [0, 2\pi]$ for each occurence of a variable.

ii) The space $C\left(\mathbb{Z}^{m};\mathbb{C}\right)$ of functions with domain in $\mathbb{Z}^{m}$ and values in $\mathbb{C}$: they are defined for any combinations of m signed integers. The appropriate measure (denoted ν above) is the Dirac measure $\delta_{z} = (\delta_{z})_{z\in\mathbb{Z}}$ and the equivalent of the integral in $L^{p}\left(\mathbb{Z}^{m}, \delta_{z}, \mathbb{C}\right)$ is a series from z=-∞ to z=+∞

29.1.2 Fourier series

See Taylor 1 p.177

Theorem 2315 *The **Fourier transform of periodic functions** is the map :*

$$\hat{\ }:L^{1}\left(T^{m}, d\theta, \mathbb{C}\right) \rightarrow C_{0\nu}\left(\mathbb{Z}^{m};\mathbb{C}\right) : \widehat{f}(z) = (2\pi)^{-m} \int_{T^{m}} f\left(\theta\right) \exp\left(-i\left\langle z, \theta\right\rangle\right) d\theta$$

.

The Fourier coefficients $\widehat{f}(z)$ vanish when $z\rightarrow \pm\infty$:
$$\forall Z \in \mathbb{Z}^{m} : \sup_{Z\in\mathbb{Z}^{m}}\left(1 + \sum_{k=1}^{m} z_{k}^{2}\right)^{Z}\left|\widehat{f}(z)\right| < \infty$$

Theorem 2316 *The space $L^{2}\left(\mathbb{Z}^{m}, \delta_{z}, \mathbb{C}\right)$ is a Hilbert space with the scalar product :* $\langle\varphi, \psi\rangle = \sum_{z\in\mathbb{Z}^{m}} \overline{\varphi\left(z\right)}\psi\left(z\right)$ *and the Hilbert basis :* $(\exp iz)_{z\in\mathbb{Z}^{m}}$

Theorem 2317 *The space $L^{2}\left(T^{m}, d\theta, \mathbb{C}\right)$ is a Hilbert space with the scalar product :* $\langle f, g\rangle = (2\pi)^{-m} \int_{T^{m}} \overline{f\left(\theta\right)}g\left(\theta\right) d\theta$ *and the Hilbert basis :* $(\exp iz\theta)_{z\in\mathbb{Z}^{m}}$

We have the identity : $\langle\exp iz_{1}\theta, \exp iz_{2}\theta\rangle = (2\pi)^{-m} \int_{T^{m}} e^{-i\langle z_{1}, \theta\rangle} e^{i\langle z_{2}, \theta\rangle} d\theta = \delta_{z_{1}, z_{2}}$

Theorem 2318 *The Fourier transform is a continuous operator between the Hilbert spaces :*
$\mathcal{F}:L^{2}\left(T^{m}, d\theta, \mathbb{C}\right) \rightarrow L^{2}\left(\mathbb{Z}^{m}, \delta_{z}, \mathbb{C}\right) :: \mathcal{F}\left(f\right)\left(z\right) = (2\pi)^{-m} \int_{T^{m}} f\left(\theta\right) \exp\left(-i\left\langle z, \theta\right\rangle\right) d\theta$
and its inverse is the adjoint map :
$\mathcal{F}^{*}:L^{2}\left(\mathbb{Z}^{m}, \delta_{z}, \mathbb{C}\right) \rightarrow L^{2}\left(T^{m}, d\theta, \mathbb{C}\right) :: \mathcal{F}^{*}\left(\varphi\right)\left(\theta\right) = \sum_{z\in\mathbb{Z}^{m}} \varphi\left(z\right) \exp i\left\langle z, \theta\right\rangle$
It is an isometry : $\langle\mathcal{F}f, \varphi\rangle_{L^{2}(\mathbb{Z}^{m}, \nu, \mathbb{C})} = \langle f, \mathcal{F}^{*}\varphi\rangle_{L^{2}(T^{m}, \mu, \mathbb{C})}$
Parseval identity : $\langle\mathcal{F}f, \mathcal{F}h\rangle_{L^{2}(\mathbb{Z}^{m}\delta_{z}, \mathbb{C})} = \langle f, h\rangle_{L^{2}(T^{m}, d\theta, \mathbb{C})}$
Plancherel identity : $\|\mathcal{F}f\|_{L^{2}(\mathbb{Z}^{m}, \delta_{z}, \mathbb{C})} = \|f\|_{L^{2}(T^{m}, d\theta, \mathbb{C})}$

Theorem 2319 *(Taylor 1 p.183) Any periodic function $f \in L^{2}\left(T^{m}, d\theta, \mathbb{C}\right)$ can be written as the series :*

$$f\left(\theta\right) = \sum_{z\in\mathbb{Z}^{m}} \mathcal{F}\left(f\right)\left(z\right) \exp i\left\langle z, \theta\right\rangle$$

Theorem 2320 *The series is still absolutely convergent if $f \in C_{r}\left(T^{m};\mathbb{C}\right)$ and $r > m/2$ or if $f \in C_{\infty}\left(T^{m};\mathbb{C}\right)$*

677

29.1.3 Operations with Fourier series

Theorem 2321 *If f is r differentiable : $\widehat{D_{\alpha_1...\alpha_r}f}(z) = (i)^r (z_{\alpha_1}...z_{\alpha_r}) \widehat{f}(z)$*

Theorem 2322 *For $f, g \in L^1(T^m, d\theta, \mathbb{C})$:*
$$\widehat{f}(z) \times \widehat{g}(z) = \widehat{(f * g)}(z) \text{ with } (f * g)(\theta) = (2\pi)^{-m} \int_{T^m} f(\zeta) g(\zeta - \theta) d\zeta$$
$$\left(\widehat{f \times g}\right)(z) = \sum_{\zeta \in \mathbb{Z}^m} \widehat{f}(z - \zeta) \widehat{g}(\zeta)$$

Theorem 2323 *Abel summability result (Taylor 1 p.180) : For any $f \in L^1(T^m, d\theta, \mathbb{C})$
the series*
$$J_r f(\theta) = \sum_{z \in \mathbb{Z}^m} \widehat{f}(z) r^{\|z\|} e^{i\langle z, \theta \rangle} \text{ with } \|z\| = \sum_{k=1}^m |z_k|$$
converges to f when $r \to 1$
i) if f is continuous and then the convergence is uniform on T^m
ii) or if $f \in L^p(T^m, d\theta, \mathbb{C}), 1 \leq p \leq \infty$

If m=1 $J_r f(\theta)$ can be written with $z = r \sin \theta$:
$$J_r f(\theta) = PI(f)(z) = \frac{1 - |z|^2}{2\pi} \int_{|\zeta|=1} \frac{f(\zeta)}{|\zeta - z|^2} d\zeta$$
and is called the Poisson integral.
This is the unique solution to the Dirichlet problem :
$$u \in C(\mathbb{R}^2; \mathbb{C}) : \Delta u = 0 \text{ in } |x^2 + y^2| < 1, \text{ u = f on } |x^2 + y^2| = 1$$

29.2 Fourier integrals

29.2.1 Definition

The Pontryagin dual of $G = (\mathbb{R}^m, +)$ is $\widehat{G} = \{\exp i\theta(g), t \in \mathbb{R}^{m*}\}$ so $\chi(g) = \exp i \sum_{k=1}^m t_i x_i = \exp i \langle t, x \rangle$

The Haar measure is proportional to the Lebesgue measure, which gives several common definitions of the Fourier transform, depending upon the location of a 2π factor. The chosen solution gives the same formula for the inverse (up to a sign). See Wikipedia "Fourier transform" for the formulas with other conventions of scaling.

Theorem 2324 *(Lieb p.125) The **Fourier transform of functions** is the
map :*
$$\hat{\ }:L^1(\mathbb{R}^m, dx, \mathbb{C}) \to C_{0\nu}(\mathbb{R}^m; \mathbb{C}) : \widehat{f}(t) = (2\pi)^{-m/2} \int_{\mathbb{R}^m} f(x) e^{-i\langle t, x \rangle} dx \text{ with }$$
$\langle x, t \rangle = \sum_{k=1}^m x_k t_k$
then $\widehat{f} \in L^\infty(\mathbb{R}^m, dx, \mathbb{C})$ and $\left\|\widehat{f}\right\|_\infty \leq \|f\|_1$
Moreover :
If $f \in L^2(\mathbb{R}^m, dx, \mathbb{C}) \cap L^1(\mathbb{R}^m, dx, \mathbb{C})$ then $\widehat{f} \in L^2(\mathbb{R}^m, dx, \mathbb{C})$ and $\left\|\widehat{f}\right\|_2 = \|f\|_1$
If $f \in S(\mathbb{R}^m)$ then $\widehat{f} \in S(\mathbb{R}^m)$
For $\widehat{f} \in L^1(\mathbb{R}^m, dx, \mathbb{C})$ there is an inverse given by :
$$\mathcal{F}^{-1}\left(\widehat{f}\right)(x) = (2\pi)^{-m/2} \int_{\mathbb{R}^m} \widehat{f}(t) e^{i\langle t, x \rangle} dt \text{ for almost all t. If f is continuous}$$
then the identity stands for all x in $\mathbb{R}^m$

The Fourier transform is a bijective, bicontinuous, map on $S(\mathbb{R}^m)$. As $S(\mathbb{R}^m) \subset L^p(\mathbb{R}^m, dx, \mathbb{C})$ for any p, it is also an unitary map (see below).

Warning ! the map is usually not surjective, there is no simple characterization of the image

Theorem 2325 *(Lieb p.128,130) There is a unique extension $\mathcal{F}$ of the Fourier transform to a continuous operator $\mathcal{F} \in \mathcal{L}\left(L^2(\mathbb{R}^m, dx, \mathbb{C}); L^2(\mathbb{R}^m, dx, \mathbb{C})\right)$.*

Moreover $\mathcal{F}$ is an unitary isometry :

Its inverse is the adjoint map : $\mathcal{F}^* = \mathcal{F}^{-1}$ with $\mathcal{F}^{-1}\left(\widehat{f}\right)(x) = \left(\mathcal{F}\widehat{f}\right)(-x)$.

Parserval formula : $\langle \mathcal{F}(f), g \rangle_{L^2} = \langle f, \mathcal{F}^*(g) \rangle_{L^2} \Leftrightarrow \int_{\mathbb{R}^m} \overline{\mathcal{F}(f)(t)} g(t) \, dt = \int_{\mathbb{R}^m} \overline{f(x)} (\mathcal{F}^*g)(x) \, dx$

Plancherel identity: $\|\mathcal{F}f\|_{L^2(\mathbb{R}^m, dx, \mathbb{C})} = \|f\|_{L^2(\mathbb{R}^m, dx, \mathbb{C})}$

However $\mathcal{F}$ is not given by any explicit formula but by an approximation with a sequence in $L^1 \cap L^2$ which is dense in L^2. Of course whenever f belongs also to $L^1(\mathbb{R}^m, dx, \mathbb{C})$ then $\mathcal{F} \equiv \hat{\ }$

Remark : the Fourier transform is defined for functions over $\mathbb{R}^m$. If O is an open subset of $\mathbb{R}^m$ then any function $f \in L^1(O, dx, \mathbb{C})$ can be extended to a function $\widetilde{f} \in L^1(\mathbb{R}^m, dx, \mathbb{C})$ with $\widetilde{f}(x) = f(x), x \in O, \widetilde{f}(x) = 0, x \neq O$ (no continuity condition is required).

29.2.2 Operations with Fourier integrals

Theorem 2326 *Derivatives : whenever the Fourier transform is defined :*
$\mathcal{F}(D_{\alpha_1 \ldots \alpha_r} f)(t) = i^r (t_{\alpha_1} .. t_{\alpha_r}) \mathcal{F}(f)(t)$
$\mathcal{F}(x^{\alpha_1} x^{\alpha_2} \ldots x^{\alpha_r} f) = i^r D_{\alpha_1 \ldots \alpha_r}(\mathcal{F}(f))$

with, of course, the usual notation for the derivative : $D_{(\alpha)} = D_{\alpha_1 \ldots \alpha_s} = \dfrac{\partial}{\partial \xi^{\alpha_1}} \dfrac{\partial}{\partial \xi^{\alpha_2}} \ldots \dfrac{\partial}{\partial \xi^{\alpha_s}}$

Theorem 2327 *(Lieb p.181) If $f \in L^2(\mathbb{R}^m, dx, \mathbb{C})$ then $\widehat{f} \in H^1(\mathbb{R}^m)$ iff the function : $g : \mathbb{R}^m \to \mathbb{R}^m :: g(t) = \|t\| \widehat{f}(t)$ is in $L^2(\mathbb{R}^m, dx, \mathbb{C})$ .And then :*
$\mathcal{F}(D_{\alpha_1 \ldots \alpha_r} f)(t) = (i)^r (t_{\alpha_1} .. t_{\alpha_r}) \mathcal{F}(f)$ and $\|f\|_2 = \int_{\mathbb{R}^m} \left|\widehat{f}(t)\right|^2 \left(1 + \|t\|^2\right) dt$

Theorem 2328 *Translation : For $f \in L^1(\mathbb{R}^m, dx, \mathbb{C}), a \in \mathbb{R}^m, \tau_a(x) = x - a, \tau_a^* f = f \circ \tau_a$*
$\mathcal{F}(\tau_a^* f)(t) = e^{-i\langle a, t \rangle} \mathcal{F}(f)(t)$
$\mathcal{F}\left(e^{i\langle a, x \rangle} f\right) = \tau_a^* \mathcal{F}(f)$

Theorem 2329 *Scaling : For $f \in L^1(\mathbb{R}^m, dx, \mathbb{C}), a \neq 0 \in \mathbb{R}, \lambda_a(x) = ax, \lambda_a^* f = f \circ \lambda_a$*
$\mathcal{F}(\lambda_a^* f) = \frac{1}{|a|} \lambda_{1/a}^* \mathcal{F}(f) \Leftrightarrow \mathcal{F}(\lambda_a^* f)(t) = \frac{1}{|a|} \mathcal{F}(f)\left(\frac{t}{a}\right)$

Theorem 2330 *For $f \in L^1(\mathbb{R}^m, dx, \mathbb{C}), L \in GL(\mathbb{R}^m; \mathbb{R}^m) :: y = [A]x$ with $detA \neq 0 : L_*(\mathcal{F}(f)) = |\det A| \mathcal{F}\left((L^t)^* f\right)$*

Proof. $L_* \left(\mathcal{F} \left(f \right) \right) = \left(L^{-1} \right)^* \left(\mathcal{F} \left(f \right) \right) = \left(\mathcal{F} \left(f \right) \right) \circ L^{-1} = \left(2\pi \right)^{-m/2} \int_{\mathbb{R}^m} e^{-i \left\langle A^{-1} t, x \right\rangle} f \left(x \right) dx$

$= \left(2\pi \right)^{-m/2} \int_{\mathbb{R}^m} e^{-i \left\langle t, \left(A^t \right)^{-1} x \right\rangle} f \left(x \right) dx = \left(2\pi \right)^{-m/2} \left| \det A \right| \int_{\mathbb{R}^m} e^{-i \langle t, y \rangle} f \left(A^t y \right) dy =$
$\left| \det A \right| \mathcal{F} \left(f \circ L^t \right)$ ■

Theorem 2331 *Conjugation :For $f \in L^1 \left(\mathbb{R}^m, dx, \mathbb{C} \right) : \overline{\mathcal{F} \left(f \right) \left(t \right)} = \mathcal{F} \left(\overline{f} \right) \left(-t \right)$*

Theorem 2332 *Convolution (Lieb p.132) If $f \in L^p \left(\mathbb{R}^m, dx, \mathbb{C} \right), g \in L^q \left(\mathbb{R}^m, dx, \mathbb{C} \right)$ with $1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}, 1 \leq p, q, r \leq 2$ then :*

$\mathcal{F} \left(f * g \right) = \left(2\pi \right)^{m/2} \mathcal{F} \left(f \right) \mathcal{F} \left(g \right)$
$\mathcal{F} \left(fg \right) = \left(2\pi \right)^{-m/2} \mathcal{F} \left(f \right) * \mathcal{F} \left(g \right)$

Some usual Fourier transforms for m=1: $\mathcal{F}(He^{-ax}) = \frac{1}{\sqrt{2\pi}} \frac{1}{a+it}$

$\mathcal{F} \left(e^{-ax^2} \right) = \frac{1}{\sqrt{2a}} e^{-\frac{t^2}{4a}}$
$\mathcal{F} \left(e^{-a|x|} \right) = \sqrt{\frac{2}{\pi}} \frac{a}{a^2 + t^2}$

Partial Fourier transform

Theorem 2333 *(Zuily p.121) Let $X = \left(x, y \right) \in \mathbb{R}^n \times \mathbb{R}^p = \mathbb{R}^m$*

The partial Fourier transform of $f \in L^1 \left(\mathbb{R}^{n+p}; dx, \mathbb{C} \right)$ on the first n components is the function :
$\widehat{f}(t, y) = \left(2\pi \right)^{-n/2} \int_{\mathbb{R}^n} f(x, y) e^{-i \langle t, x \rangle} dx \in C_{0\nu} \left(\mathbb{R}^m; \mathbb{C} \right)$
It is still a bijective map on $S(\mathbb{R}^m)$ and the inverse is :
$f(x, y) = \left(2\pi \right)^{-n/2} \int_{\mathbb{R}^n} \widehat{f}(t, y) e^{i \langle t, x \rangle} dt$

All the previous properties stand whenever we consider the Fourier transform on the first n components.

Fourier transforms of radial functions
A radial function in $\mathbb{R}^m$ is a function F(x)=f(r) where r=$\|x\|$
$\widehat{F}(t) = \left(2\pi \right)^{-m/2} \int_0^\infty f(r) \phi(r \|t\|) r^{n-1} dr = \|t\|^{1-\frac{m}{2}} \int_0^\infty f(r) J_{\frac{m}{2}-1}(r \|t\|) r^{\frac{m}{2}} dr$
with $\phi(u) = \int_{S^{m-1}} e^{iut} d\sigma_S = A_{m-2} \int_{-1}^1 e^{irs} \left(1 - s^2 \right)^{(m-3)/2} ds = \left(2\pi \right)^{m/2} u^{1-m/2} J_{\frac{m}{2}-1}(u)$
with the Bessel function defined for $\mathrm{Re}\, \nu > -1/2$:
$J_\nu(z) = \left(\Gamma \left(\frac{1}{2} \right) \Gamma \left(\nu + \frac{1}{2} \right) \right)^{-1} \left(\frac{z}{2} \right)^\nu \int_{-1}^1 \left(1 - t^2 \right)^{\nu - \frac{1}{2}} e^{izt} dt$
which is solution of the ODE :
$\left(\frac{d^2}{dr^2} + \frac{1}{r} \frac{d}{dr} + \left(1 - \frac{\nu^2}{r^2} \right) \right) J_\nu(r) = 0$

Paley-Wiener theorem

Theorem 2334 *(Zuily p.123) For any function $f \in C_{\infty c} \left(\mathbb{R}^m; \mathbb{C} \right)$ with support $Supp(f) = \{ x \in \mathbb{R}^m : \|x\| \leq r \}$ there is a holomorphic function F on $\mathbb{C}^m$ such that :*
$\forall x \in \mathbb{R}^m : F(x) = \widehat{f}(x)$

(1) $\forall n \in \mathbb{N}, \exists C_n \in \mathbb{R} : \forall z \in \mathbb{C}^m : |F(z)| \le C_n (1 + \|z\|)^{-n} e^{r|\text{Im } z|}$

Conversely for any holoorphic function F on $\mathbb{C}^m$ which meets the lattest condition (1) is met there is a function $f \in C_{\infty c}(\mathbb{R}^m; \mathbb{C})$ such that Supp(f)=$\{x \in \mathbb{R}^m : \|x\| \le r\}$ and $\forall x \in \mathbb{R}^m : \widehat{f}(x) = F(x)$

This theorem shows in particular that the Fourier transform of function with compact support is never compactly supported (except if it is null). This is a general feature of Fourier transform : $\widehat{f}$ is always "more spread out" than f. And conversely $\mathcal{F}^*$ is "more concentrated" than f.

Asymptotic analysis

Theorem 2335 *(Zuily p.127) For any functions $\varphi \in C_\infty(\mathbb{R}^m; \mathbb{R}), f \in C_{\infty c}(\mathbb{R}^m; \mathbb{C})$ the asymptotic value of $I(t) = \int_{\mathbb{R}^m} e^{it\varphi(x)} f(x) \, dx$ when $t \to +\infty$ is the following :*

i) If $\forall x \in Supp(f), \varphi'(x) \ne 0$ then $\forall n \in \mathbb{N}, \exists C_n \in \mathbb{R} : \forall t \ge 1 : |I(t)| \le C_n t^{-n}$

ii) If φ has a unique critical point $a \in Supp(f)$ and it is not degenerate $(\det \varphi"(a) \ne 0)$ then :
$\forall n \in \mathbb{N}, \exists (a_k)_{k=0}^n, a_k \in \mathbb{C}, \exists C_n > 0, r_n \in C(\mathbb{R}; \mathbb{R}) :$
$I(t) = r_n(t) + e^{it\varphi(a)} \sum_{k=0}^n a_k t^{-\frac{\pi}{2} - k}$
$|r_n(t)| \le C_n t^{-\frac{\pi}{2} - n}$
$a_0 = \frac{(2\pi)^{m/2}}{\sqrt{|\det \varphi"(a)|}} e^{i\frac{\pi}{4}\epsilon} f(a)$ *with $\epsilon =$sign $\det \varphi"(a)$*

29.3 Fourier transform of distributions

For an abelian topological group G endowed with a Haar measure μ the Fourier transform is well defined as a continuous linear map : $\hat{}:L^1(G, \mu, \mathbb{C}) \to C_{0\nu}(G; \mathbb{C})$ and its has an extension $\mathcal{F}$ to $L^2(G, \mu, \mathbb{C})$ as well.

So if V is some subspace of $L^1(G, \mu, \mathbb{C})$ or $L^2(G, \mu, \mathbb{C})$ we can define the Fourier transform on the space of distributions V' as :
$\mathcal{F} : V' \to V' :: \mathcal{F}S(\varphi) = S(\mathcal{F}(\varphi))$ whenever $\mathcal{F}(\varphi) \in V$

If G is a compact group $\mathcal{F}(\varphi)$ is a function on $\widehat{G}$ which is a discrete group and cannot belong to V (except if G is itself a finite group). So the procedure will work only if G is isomorphic to a finite dimensional vector space E, because $\widehat{G}$ is then isomorphic to E.

29.3.1 Definition

The general rules are :
i) $\mathcal{F}S(\varphi) = S(\mathcal{F}(\varphi))$ whenever $\mathcal{F}(\varphi) \in V$
ii) $\mathcal{F}(T(f)) = T(\mathcal{F}(f))$ whenever $S = T(f) \in V'$
Here T is one of the usual maps associating functions on $\mathbb{R}^m$ to distributions.

Theorem 2336 *Whenever the Fourier transform is well defined for a function f, and there is an associated distribution S=T(f), the Fourier transform of S is the distribution :* $\mathcal{F}\left(T\left(f\right)\right) = T\left(\mathcal{F}\left(f\right)\right) = \left(2\pi\right)^{-m/2} T_t\left(S_x\left(e^{-i\langle x,t\rangle}\right)\right)$

Proof. For $S = T(f), f \in L^1\left(\mathbb{R}^m, dx, \mathbb{C}\right)$

$\qquad S\left(\mathcal{F}\left(\varphi\right)\right) = T(f)\left(\mathcal{F}\left(\varphi\right)\right)\right) = \left(2\pi\right)^{-m/2} \int_{\mathbb{R}^m} f\left(x\right)\left(\int_{\mathbb{R}^m} e^{-i\langle x,t\rangle}\varphi\left(t\right) dt\right) dx$

$\qquad = \int_{\mathbb{R}^m} \varphi\left(t\right)\left(\left(2\pi\right)^{-m/2} \int_{\mathbb{R}^m} e^{-i\langle x,t\rangle} f\left(x\right) dx\right) dt = T\left(\widehat{f}\right)\left(\varphi\right)$

$\qquad = \left(2\pi\right)^{-m/2} \int_{\mathbb{R}^m} \varphi\left(t\right)\left(T_x(f)\left(e^{-i\langle x,t\rangle}\right)\right) dt$

$\qquad = \left(2\pi\right)^{-m/2} T_t\left(S_x\left(e^{-i\langle x,t\rangle}\right)\right)\left(\varphi\right) = \mathcal{F}S\left(\varphi\right)$ ∎

Which implies that it suffices to compute $S_x\left(e^{-i\langle x,t\rangle}\right)$ to get the Fourier transform of the distribution.

Theorem 2337 *The map :* $F : L^2\left(\mathbb{R}^m, dx, \mathbb{C}\right) \to L^2\left(\mathbb{R}^m, dx, \mathbb{C}\right) :: F\left(f\right)\left(x\right) = \left(2\pi\right)^{-m/2} T_t(f)\left(e^{-i\langle x,t\rangle}\right)$ *is a bijective isometry.*

Theorem 2338 *(Zuily p.114) The Fourier transform $\mathcal{F}$ and its inverse $\mathcal{F}$* are continuous bijective linear operators on the space $S\left(\mathbb{R}^m\right)'$ of tempered distributions*

$\qquad \mathcal{F}: S(\mathbb{R}^m)' \to S\left(\mathbb{R}^m\right)' :: \mathcal{F}\left(S\right)\left(\varphi\right) = S\left(\mathcal{F}\left(\varphi\right)\right)$

$\qquad \mathcal{F}^*: S(\mathbb{R}^m)' \to S\left(\mathbb{R}^m\right)' :: \mathcal{F}^*\left(S\right)\left(\varphi\right) = S\left(\mathcal{F}^*\left(\varphi\right)\right)$

$\mathcal{F}, \mathcal{F}$* on functions are given by the usual formulas.

Theorem 2339 *(Zuily p.117) The Fourier transform $\mathcal{F}$ and its inverse $\mathcal{F}$* are continuous linear operators on the space $C_{rc}\left(\mathbb{R}^m; \mathbb{C}\right)'_c$ of distributions with compact support.*

$\qquad \mathcal{F}: C_{rc}\left(\mathbb{R}^m; \mathbb{C}\right)'_c \to C_{rc}\left(\mathbb{R}^m; \mathbb{C}\right)'_c :: \mathcal{F}\left(S\right)\left(\varphi\right) = S\left(\mathcal{F}\left(\varphi\right)\right)$

$\qquad \mathcal{F}^*: C_{rc}\left(\mathbb{R}^m; \mathbb{C}\right)'_c \to C_{rc}\left(\mathbb{R}^m; \mathbb{C}\right)'_c :: \mathcal{F}^*\left(S\right)\left(\varphi\right) = S\left(\mathcal{F}^*\left(\varphi\right)\right)$

$\qquad$ *Moreover : if* $S \in C_{\infty c}\left(\mathbb{R}^m; \mathbb{C}\right)'_c \equiv C_\infty\left(\mathbb{R}^m; \mathbb{C}\right)'$ *then* $S_x\left(e^{-i\langle x,t\rangle}\right) \in C_\infty\left(\mathbb{R}^m; \mathbb{C}\right)$ *and can be extended to a holomorphic function in* $\mathbb{C}^m$

$\mathcal{F}, \mathcal{F}$* on functions are given by the usual formulas.

Theorem 2340 *Paley-Wiener-Schwartz (Zuily p.126): For any distribution* $S \in \left(C_{rc}\left(\mathbb{R}^m; \mathbb{C}\right)\right)'_c$ *with support Supp(S)=$\{x \in \mathbb{R}^m : \|x\| \leq \rho\}$ there is a holomorphic function F on $\mathbb{C}^m$ such that :*

$\qquad \forall \varphi \in C_{rc}\left(\mathbb{R}^m; \mathbb{C}\right) : T\left(F\right)\left(\varphi\right) = \left(\mathcal{F}\left(S\right)\right)\left(\varphi\right)$

$\qquad$ *(1)* $\exists C \in \mathbb{R} : \forall z \in \mathbb{C}^m : \left|F\left(z\right)\right| \leq C_n\left(1 + \|z\|\right)^{-r} e^{\rho|\text{Im } z|}$

$\qquad$ *Conversely for any holomorphic function F on $\mathbb{C}^m$ which meets the condition (1) there is a distribution $S \in \left(C_{rc}\left(\mathbb{R}^m; \mathbb{C}\right)\right)'_c$ with support Supp(S)=$\{x \in \mathbb{R}^m : \|x\| \leq \rho\}$ such that $\forall \varphi \in C_{rc}\left(\mathbb{R}^m; \mathbb{C}\right) : T\left(F\right)\left(\varphi\right) = \left(\mathcal{F}\left(S\right)\right)\left(\varphi\right)$*

$\qquad$ So $\mathcal{F}\left(S\right) = T\left(F\right)$ and as S∈$S(\mathbb{R}^m)'$: $S = \mathcal{F}^* T\left(F\right) = T\left(\mathcal{F}^* F\right) = \left(2\pi\right)^{-m/2} T_t\left(S_x\left(e^{i\langle x,t\rangle}\right)\right) \Rightarrow$ $F\left(t\right) = S_x\left(e^{i\langle x,t\rangle}\right)$

29.3.2 Properties

Theorem 2341 *Derivative : Whenever the Fourier transform and the derivative are defined :*

$$\mathcal{F}(D_{\alpha_1...\alpha_r}S) = i^r (t_{\alpha_1}...t_{\alpha_r})\, \mathcal{F}(S)$$
$$\mathcal{F}(x^{\alpha_1}...x^{\alpha_r}S) = i^r D_{\alpha_1...\alpha_r}(\mathcal{F}(S))$$

Theorem 2342 *(Zuily p.115) Tensorial product : For $S_k \in S(\mathbb{R}^m)'$, $k = 1..n$:$\mathcal{F}(S_1 \otimes S_2... \otimes S_n) = \mathcal{F}(S_1) \otimes ... \otimes \mathcal{F}(S_n)$*

Theorem 2343 *Pull-back : For $S \in S(\mathbb{R}^m)'$, $L \in GL(\mathbb{R}^m; \mathbb{R}^m) :: y = [A]\,x$ with $det A \neq 0 : \mathcal{F}(L^*S) = S(\mathcal{F}(L^t)^*)$*

Proof. $L^*S_y(\varphi) = |\det A|^{-1} S\left(\left(L^{-1}\right)^* \varphi\right)$

$$\mathcal{F}(L^*S_y)(\varphi) = L^*S_y(\mathcal{F}(\varphi)) = |\det A|^{-1} S_y\left(\left(L^{-1}\right)^*(\mathcal{F}(\varphi))\right)$$
$$= |\det A|^{-1} S_y\left(|\det A|\,\mathcal{F}(\varphi \circ L^t)\right) = S_y\left(\mathcal{F}\left(\left(L^t\right)^* \varphi\right)\right) \quad \blacksquare$$

Theorem 2344 *Homogeneous distribution : If S is homogeneous of order n in $S(\mathbb{R}^m)$ then $\mathcal{F}(S)$ is homogenous of order -n-m*

Theorem 2345 *Convolution :.For $S \in C_{\infty c}(\mathbb{R}^m; \mathbb{C})'_c$, $U \in S(\mathbb{R}^m)' : \mathcal{F}(U * S) = S_t\left(e^{-i\langle x,t\rangle}\right)\mathcal{F}(U)$*

which must be understood as the product of the function of x : $S_t\left(e^{-i\langle x,t\rangle}\right)$ by the distribution $\mathcal{F}(U)$, which is usually written, : $\mathcal{F}(U * S) = (2\pi)^{m/2}\,\mathcal{F}S \times \mathcal{F}(U)$, incorrectly, because the product of distributions is not defined.

Examples of Fourier transforms of distributions:

$$\mathcal{F}(\delta_0) = \mathcal{F}^*(\delta_0) = (2\pi)^{-m/2}\,T(1)\,;$$
$$\mathcal{F}(T(1)) = \mathcal{F}^*(T(1)) = (2\pi)^{m/2}\,\delta_0;$$
$$S = T\left(e^{i\epsilon a\|x\|^2}\right), a > 0, \epsilon = \pm 1 : \mathcal{F}(S) = \left(\tfrac{1}{2a}\right)^{m/2} e^{mi\epsilon\frac{\pi}{4}}\,T\left(e^{-i\epsilon\frac{\|x\|^2}{a}}\right)$$
$$S = vp\tfrac{1}{x} \in S(\mathbb{R})' : \mathcal{F}(S) = T\left(i\sqrt{2\pi}\,(2-H)\right) \text{ with the Heaviside function}$$
H

(Zuily p.127) Let be σ_r the Lebesgue measure on the sphere of radius r in $\mathbb{R}^m$.$\sigma_r \in (C_{\infty c}(\mathbb{R}^m; \mathbb{C}))'_c$

For m=3 $\mathcal{F}(\sigma_r) = \sqrt{\tfrac{2}{\pi}}\,rT\left(\tfrac{\sin r\|x\|}{\|x\|}\right)$

Conversely for any m>0 : $\text{supp}\mathcal{F}^*\left(T\left(\tfrac{\sin r\|x\|}{\|x\|}\right)\right) \subset \{\|x\| \leq r\}$

29.3.3 Extension of Sobolev spaces

We have seen the Sobolev spaces $H^r(O)$ for functions with r$\in \mathbb{N}$, extended to r<0 as their dual (with distributions). Now we extend to r$\in \mathbb{R}$.

Definition

Definition 2346 *For $m \in \mathbb{N}$, $s \in \mathbb{R}$, the Sobolev Space denoted $H^s(\mathbb{R}^m)$ is the subspace of distributions $S \in S(\mathbb{R}^m)'$ induced by a function $f : S = T(f)$ such that $f \in C(\mathbb{R}^m; \mathbb{C})$ with $\left(1 + \|x\|^2\right)^{s/2} \mathcal{F}(f) \in L^2(\mathbb{R}^m; dx, \mathbb{C})$*

$\forall s \in \mathbb{N}$, $\mathrm{H}^s(\mathbb{R}^m)$ coincides with $T(H^s(\mathbb{R}^m))$ where $\mathrm{H}^s(\mathbb{R}^m)$ is the usual Sobolev space

Theorem 2347 *Inclusions (Taylor 1 p.270, Zuily p.133,136): We have the following inclusions :*
i) $\forall s_1 \geq s_2 : H^{s_1}(\mathbb{R}^m) \subset H^{s_2}(\mathbb{R}^m)$ and the embedding is continuous
ii) $\delta_0 \subset H^s(\mathbb{R}^m)$ iff $s < -m/2$
iii) $\forall s < m/2 : T\left(L^1(\mathbb{R}^m; dx, \mathbb{C})\right) \subset H^s(\mathbb{R}^m)$
iv) $\forall s \in \mathbb{R} : T(S(\mathbb{R}^m)) \subset H^s(\mathbb{R}^m)$ and is dense in $H^s(\mathbb{R}^m)$
v) $\forall s \in \mathbb{R} : T(C_{\infty c}(\mathbb{R}^m; \mathbb{C})) \subset H^s(\mathbb{R}^m)$ and is dense in $H^s(\mathbb{R}^m)$
vi) $\forall s > m/2 + r, r \in \mathbb{N} : H^s(\mathbb{R}^m) \subset T(C_r(\mathbb{R}^m; \mathbb{C}))$
vi) $\forall s > m/2 : H^s(\mathbb{R}^m) \subset T(C_{0b}(\mathbb{R}^m; \mathbb{C}))$
vii) $\forall s > m/2 + \gamma, 0 < \gamma < 1 : H^s(\mathbb{R}^m) \subset T(C^\gamma(\mathbb{R}^m; \mathbb{C}))$ (Lipschitz functions)
vi) $C_{\infty c}(\mathbb{R}^m; \mathbb{C})'_c \subset \cup_{s \in \mathbb{R}} H^s(\mathbb{R}^m)$

Properties

Theorem 2348 *(Taylor 1 p.270, Zuily p.133,137) $H^s(\mathbb{R}^m)$ is a Hilbert space, its dual is a Hilbert space which is anti-isomorphic to $H^{-s}(\mathbb{R}^m)$ by : $\tau : H^{-s}(\mathbb{R}^m) \to (H^s(\mathbb{R}^m))'$ :: $\tau(U) = \mathcal{F}(\psi(-x))$ where $U = T(\psi)$*

The scalar product on $H^s(\mathbb{R}^m)$ is :
$S, U \in H^s(\mathbb{R}^m), S = T(\varphi), U = T(\psi)$:
$$\langle S, U \rangle = \left\langle \left(1 + \|x\|^2\right)^{s/2} \mathcal{F}(\varphi), \left(1 + \|x\|^2\right)^{s/2} \mathcal{F}(\psi) \right\rangle_{L^2}$$
$$= \int_{\mathbb{R}^m} \left(1 + \|x\|^2\right)^s \overline{\mathcal{F}(\varphi)} \mathcal{F}(\psi) \, dx$$

Theorem 2349 *(Zuily p.135) Product with a function : $\forall s \in \mathbb{R} : \forall \varphi \in S(\mathbb{R}^m), S \in H^s(\mathbb{R}^m) : \varphi S \in H^s(\mathbb{R}^m)$*

Theorem 2350 *(Zuily p.135) Derivatives : $\forall s \in \mathbb{R}, \forall S \in H^s(\mathbb{R}^m), \forall \alpha_1, ... \alpha_r : D_{\alpha_1 ... \alpha_r} S \in H^{s-r}(\mathbb{R}^m), \|D_{\alpha_1 ... \alpha_r} S\|_{H^{s-r}} \leq \|S\|_{H^s}$*

Theorem 2351 *Pull back : For $L \in G\mathcal{L}(\mathbb{R}^m; \mathbb{R}^m)$ the pull back L^*S of $S \in S(\mathbb{R}^m)'$ is such that $L^*S \in S(\mathbb{R}^m)'$ and $\forall s \in \mathbb{R} : L^* : H^s(\mathbb{R}^m) \to H^s(\mathbb{R}^m)$*

Theorem 2352 *(Zuily p.141) Trace :The map : $Tr_m : S(\mathbb{R}^m) \to S(\mathbb{R}^{m-1})$ :: $Tr_m(\varphi)(x_1,..x_{m-1}) = \varphi(x_1,..x_{m-1},0)$ is continuous and $\forall s > \frac{1}{2}$ there is a unique extension to a continuous linear surjective map : $Tr_m : H^s(\mathbb{R}^m) \to H^{s-\frac{1}{2}}(\mathbb{R}^{m-1})$*

Theorem 2353 *(Zuily p.139) If K is a compact of $\mathbb{R}^m$, $H^s(K)$ the subset of $H^s(\mathbb{R}^m)$ such that $S = T(f)$ with $Supp(f)$ in K, then $\forall s' > s$ the map $H^s(K) \to H^{s'}(\mathbb{R}^m)$ is compact.*

Sobolev spaces on manifolds

(Taylor 1 p.282) The space $H^s(M)$ of Sobolev distributions on a m dimensional manifold M is defined as the subset of distributions $S \in C_{\infty c}(M;\mathbb{C})'$ such that, for any chart $(O_i, \psi_i)_{i \in I}$ of M, $\psi_i(O_i) = U_i$ where U_i is identified with its embedding in $\mathbb{R}^m$ and any $\varphi \in C_{\infty c}(O_i;\mathbb{C}) : (\psi_i^{-1})^* \varphi S \in H^s(U_i)$.

If M is compact then we have a Sobolev space $H^s(M)$ with some of the properties of $H^s(\mathbb{R}^m)$.

The same construct can be implemented to define Sobolev spaces on compact manifolds with boundary.

30 DIFFERENTIAL OPERATORS

Differential operators are the key element of any differential equation, meaning a map relating an unknown function and its partial derivative to some known function and subject to some initial conditions. And indeed solving a differential equation can often be summarized to finding the inverse of a differential operator.

When dealing with differential operators, meaning with maps involving both a function and its derivative, a common hassle comes from finding a simple definition of the domain of the operator : we want to keep the fact that the domain of y and y' are somewhat related. The simplest solution is to use the jet formalism. It is not only practical, notably when we want to study differential operators on fiber bundles, but it gives a deeper insight of the meaning of locality, thus making a better understanding of the difference between differential operators and pseudo differential operators.

30.0.4 Definitions

Differential operators have been defined in a general settings in the Fiber bundle part. For practical purpose we will use the following (fully consistent with the previous one) :

Definition 2354 *A **r order differential operator** is a fibered manifold, base preserving, morphism $D : J^r E_1 \to E_2$ between two smooth complex finite dimensional vector bundles $E_1(M,V_1,\pi_1), E_2(M,V_2,\pi_2)$ on the same real manifold.*

As a special case we have scalar differential operators with $V_1 = V_2 = \mathbb{C}$ and $J^r E_1 = J^r (M; \mathbb{C})$

Definition 2355 *A r order scalar differential operator* on a space F of r *differentiable complex functions on a manifold M is a map* : $D : J^r F \to C(M; \mathbb{C})$

Comments

The jet formalism can seem a bit artificial, but it is useful:

i) It gives an intrinsic (coordinate free) definition. Remind that any section on a fiber bundle is defined through a family of maps on each open of a trivialization (as any tensor field) with specific transition laws.

ii) It gives a simple and direct way to deal with the domain of definition of D.

iii) It is the simplest way to deal with such matters as symbol, adjoint and transpose of an operator

iv) In physics physical laws must be equivariant under some gauge transformations. Meaning that such a differential operator must be natural, and we have some powerful theorems about such operators. At least D should be covariant in a change of chart on M, meaning that D should be equivariant under the right action of $GL^r(\mathbb{R}, m)$ on $T_m^r(V)$:

$\rho : T_m^r(V) \times GL^r(\mathbb{R}, m) \to T_m^r(V) :: \rho(j_0^r z, j_0^r \varphi) = j_0^r (z \circ \varphi)$ with $\varphi \in Diff_r(\mathbb{R}^m; \mathbb{R}^m)$

Locality

$J^r E_1, E_2$ are two fibered manifolds with the same base M, so $\forall x \in M : \pi_2 \circ D = \pi_1$. That we will write more plainly : D(x) maps fiberwise $Z(x) \in J^r E_1(x)$ to $Y(x) \in E_2(x)$

So a differential operator is *local*, in the meaning that it can be fully computed from data related at one point and these data involve not only the value of the section at x, but also the values at x of its partial derivatives up to the r order. This is important in numerical analysis, because generally one can find some algorithm to compute a function from the value of its partial derivative. This is fundamental here as we will see that almost all the properties of D use locality. We will see that pseudo differential operators are not local.

Sections on $J^r E_1$ and on E_1

D maps sections of $J^r E_1$ to sections of E_2 :

$D : \mathfrak{X}(J^r E_1) \to \mathfrak{X}(E_2) :: D(Z)(x) = D(x)(Z(x))$

D, by itself, does not involve any differentiation.

$J^r E_1$ is a vector bundle $J^r E_1(M, J_0^r(\mathbb{R}^m, V_1)_0, \pi^r)$. A section Z on $J^r E_1$ is a map $M \to J^r E_1$ and reads :

$Z = (z, z_{\alpha_1 \ldots \alpha_s}, 1 \le \alpha_k \le m, s = 1 \ldots r)$ with $z \in E_1$, $z_{\alpha_1 \ldots \alpha_s} \in V_1$ symmetric in all lower indices, dimM=m

D reads in a holonomic basis of E_2 : $D(x)(Z) = \sum_{i=1}^{n_2} D^i(x, z, z_\alpha, z_{\alpha\beta}, \ldots z_{\alpha_1 \ldots \alpha_r}) e_{2i}(x)$

Any class r section on E_1 gives rise to a section on $J^r E_1$ thus we have a map : $\widehat{D} : \mathfrak{X}_r(E_1) \to \mathfrak{X}(E_2) :: \widehat{D}(U) = D(J^r U)$ and this is $\widehat{D} = D \circ J^r$ which involves the derivatives.

if $X \in \mathfrak{X}_r(E_1), X = \sum_{i=1}^{n_1} X^i(x) e_{1i}(x)$ then $z^i_{\alpha_1...\alpha_s}(x) = D_{\alpha_1...\alpha_s} X^i(x)$ with

$$D_{(\alpha)} = D_{\alpha_1...\alpha_s} = \frac{\partial}{\partial \xi^{\alpha_1}} \frac{\partial}{\partial \xi^{\alpha_2}} ... \frac{\partial}{\partial \xi^{\alpha_s}}$$ where the $\alpha_k = 1...m$ can be identical

and $(\xi^\alpha)_{\alpha=1}^m$ are the coordinates in a chart of M.

Topology

$\mathfrak{X}_r(E_1), \mathfrak{X}(J^r E_1)$ are Fréchet spaces, $\mathfrak{X}(E_2)$ can be endowed with one of the topology see previously, $\mathfrak{X}_0(E_2)$ is a Banach space. The operator will be assumed to be continuous with the chosen topology. Indeed the domain of a differential operator is quite large (and can be easily extended to distributions), so usually more interesting properties will be found for the restriction of the operator to some subspaces of $\mathfrak{X}(J^r E_1)$

Manifold with boundary

If M is a manifold with boundary in N, then $\overset{\circ}{M}$ is an open subset and a submanifold of N. The restriction of a differential operator D from $\mathfrak{X}(J^r E_1)$ to the sections of $\mathfrak{X}(J^r E_1)$ with support in $\overset{\circ}{M}$ is well defined. It can de defined on the boundary ∂M if D is continuous on N.

Composition of differential operators

Differential operators can be composed as follows :

$D_1 : \mathfrak{X}(J^{r_1} E_1) \to \mathfrak{X}(E_2)$

$D_2 : \mathfrak{X}(J^{r_2} E_2) \to \mathfrak{X}(E_3)$

$D_2 \circ J^{r_2} \circ D_1 = \widehat{D}_2 \circ D_1 : \mathfrak{X}(J^{r_1} E_1) \to \mathfrak{X}(E_3)$

It implies that $D_1(\mathfrak{X}(E_2)) \subset \mathfrak{X}_{r_2}(E_2)$. The composed operator is in fact of order $r_1 + r_2 : \widehat{D}_3 : \mathfrak{X}_{r_1+r_2}(E_1) \to \mathfrak{X}(E_3) :: \widehat{D}_3 = \widehat{D}_2 \circ \widehat{D}_1$

Parametrix

Definition 2356 *A parametrix for a differential operator :* $D : J^r E \to E$ *on the vector bundle E is a map :* $Q : \mathfrak{X}(E) \to \mathfrak{X}(E)$ *such that* $Q \circ \widehat{D} - Id$ *and* $\widehat{D} \circ Q - Id$ *are compact maps* $\mathfrak{X}(E) \to \mathfrak{X}(E)$

Compact maps "shrink" a set. A parametrix can be seen as a proxy for an inverse map : $Q \approx \widehat{D}^{-1}$. The definition is relative to the topology on $\mathfrak{X}(E)$. Usually Q is not a differential operator but a pseudo-differential operator, and it is not unique.

Operators depending on a parameter

Quite often one meets family of differential operators depending on a scalar parameter t (usually the time) : $D(t) : J^r E_1 \to E_2$ where t belongs to $\mathbb{R}$.

One can extend a manifold M to the product TxM, and similarly a vector bundle $E(M, V, \pi)$ to a vector bundle $E_T(\mathbb{R} \times M, V, \pi_T)$.

An atlas $(O_a, \varphi_a)_{a \in A}$ of E becomes $(T \times O_a, \psi_a)_{a \in A} : \psi_a((t, x), u) = (\varphi_a(x, u), t)$ with transitions maps : $\psi_{ba}((t, x)) = \varphi_{ba}(x)$. The projection is : $\pi_T((\varphi_a(x, u), t)) = (t, x)$. All elements of the fiber at (t,x) of E_T share the same time t so the vector space structure on the fiber at t is simply that of V. E_T is still a vector bundle with fibers isomorphic to V. A section $X \in \mathfrak{X}(E_T)$ is then identical to a map : $\mathbb{R} \to \mathfrak{X}(E)$.

The r jet prolongation of E_T is a vector bundle $J^r E_T \left(\mathbb{R} \times M, J_0^r\left(\mathbb{R}^{m+1}, V_1\right)_0, \pi_T\right)$. We need to enlarge the derivatives to account for t. A section $X \in \mathfrak{X}(E_T)$ has a r jet prolongation and for any two sections the values of the derivatives are taken at the same time.

A differential operator between two vector bundles depending on a parameter is then a base preserving map : $D : J^r E_{T1} \to E_{T2}$ so $D(t, x) : J^r E_{T1}(t, x) \to E_{T2}(t, x)$. It can be seen as a family of operators D(t), but the locality condition imposes that the time is the same both in $Z_1(t)$ and $Z_2(t) = D(t)(Z_1(t))$

30.1 Linear differential operators

30.1.1 Definitions

Vector bundles are locally vector spaces, so the most "natural" differential operator is linear.

Definition 2357 *A r order **linear differential operator** is a linear, base preserving morphism $D : \mathfrak{X}(J^r E_1) \to \mathfrak{X}(E_2)$ between two smooth complex finite dimensional vector bundles $E_1(M, V_1, \pi_1), E_2(M, V_2, \pi_2)$ on the same real manifold.*

Components expression

$\forall x \in M : D(x) \in \mathcal{L}(J^r E_1(x); E_2(x))$

It reads in a holonomic basis $(e_{2,i}(x))_{i=1}^n$ of E_2

$D(J^r z(x)) = \sum_{s=0}^r \sum_{\alpha_1 \ldots \alpha_s = 1}^m \sum_{i=1}^{n_2} \sum_{j=1}^{n_1} A(x)_j^{i, \alpha_1 \ldots \alpha_s} z_{\alpha_1 \ldots \alpha_s}^j(x) e_{2,i}(x)$ where $A(x)_j^{i, \alpha_1 \ldots \alpha_s}, z_{\alpha_1 \ldots \alpha_s}^j(x) \in \mathbb{C}$ where $z_{\alpha_1 \ldots \alpha_s}^i$ is symmetric in all lower indices.

Or equivalently :

$D(J^r z(x)) = \sum_{s=0}^r \sum_{\alpha_1 \ldots \alpha_s = 1}^m A(x)^{\alpha_1 \ldots \alpha_s} Z_{\alpha_1 \ldots \alpha_s}(x)$ where $Z_{\alpha_1 \ldots \alpha_s}(x) = \sum_{j=1}^{n_1} z_{\alpha_1 \ldots \alpha_s}^i(x) e_{1,i}(x) \in E_1(x), A(x)^{\alpha_1 \ldots \alpha_s} \in \mathcal{L}(E_1(x); E_2(x))$

x appear only through the maps $A(x)_\alpha$, D is linear in all the components.

A scalar linear differential operator reads :

$D(J^r z(x)) = \sum_{s=0}^r \sum_{\alpha_1 \ldots \alpha_s = 1}^m A(x)^{\alpha_1 \ldots \alpha_s} z_{\alpha_1 \ldots \alpha_s}$ where $A(x)^{\alpha_1 \ldots \alpha_s}, z_{\alpha_1 \ldots \alpha_s} \in \mathbb{C}$

Quasi-linear operator

A differential operator si said to be **quasi-linear** if it reads :

$D(J^r z(x)) = \sum_{s=0}^{r} \sum_{\alpha_1 \ldots \alpha_s=1}^{m} A(x,z)^{\alpha_1 \ldots \alpha_s} Z_{\alpha_1 \ldots \alpha_s}(x)$ where the coefficients depend on the value of z only, and not $z^i_{\alpha_1 \ldots \alpha_s}$ (so they depend on X and not its derivatives).

Domain and range of D

We assume that the maps : $A^{\alpha_1 \ldots \alpha_s} : M \to \mathcal{L}(V_1; V_2)$ are smooth.

So on $\mathfrak{X}_0(J^r E_1)$ we have $D(\mathfrak{X}_0(J^r E_1)) \subset \mathfrak{X}_0(E_2)$ and $\widehat{D}(\mathfrak{X}_r(E_1)) \subset \mathfrak{X}_0(E_2)$.

If there is a Radon measure μ on M, and if the coefficients $A(x)^{i,\alpha_1 \ldots \alpha_s}_j$ are bounded on M (notably if M is compact) then we have :

$D(L^p(M, \mu, J^r E_1)) \subset L^p(M, \mu, E_2)$ and $\widehat{D}(W^{r,p}(E_1)) \subset L^p(M, \mu, E_2)$

Any linear differential operator can also be seen as acting on sections of E_1 :

$D^{\#}(x) : E_1(x) \to J^r E_2(x) :: D^{\#}(x)(X) = (A(x)^{\alpha_1 \ldots \alpha_s} X, s = 0..r, 1 \leq \alpha_k \leq m)$
so $Z^{\alpha_1 \ldots \alpha_s} = A(x)^{\alpha_1 \ldots \alpha_s} X$

30.1.2 Differential operators and distributions

It is customary to extend scalar differential operators from functions to "generalized functions", meaning distributions. This is a subject more subtle than it seems, so it is useful to explain how it works, as it can be extended to distributions on a vector bundle. Notice that it works only for linear differential operators, as the product of two distributions is not defined.

Reminder of distributions on vector bundles:

A distribution on a vector bundle E is a functional : $S : \mathfrak{X}_{\infty,c}(E) \to \mathbb{C}$

The action of a distribution $S \in \mathfrak{X}_{\infty,c}(E)'$ on a section $Z \in \mathfrak{X}_{\infty,c}(J^r E)$ is the map :

$S : \mathfrak{X}_{\infty,c}(J^r E) \to (\mathbb{C}, \mathbb{C}^{sm}, s = 1..r) :: S(Z_{\alpha_1 \ldots \alpha_s}(x), s = 0..r, \alpha_j = 1...m) = (S(Z_{\alpha_1 \ldots \alpha_s}), s = 0..r, \alpha_j = 1...m)$

The derivative of $S \in \mathfrak{X}_{\infty,c}(E)'$ with respect to $(\xi^{\alpha_1}, ..., \xi^{\alpha_s})$ on M is the distribution :

$(D_{\alpha_1 .. \alpha_s} S) \in \mathfrak{X}_{\infty,c}(E)': \forall X \in \mathfrak{X}_{\infty,c}(E) : (D_{\alpha_1 .. \alpha_s} S)(X) = S(D_{\alpha_1 .. \alpha_s} X)$

The r jet prolongation of $S \in \mathfrak{X}_{\infty,c}(E)'$ is the map $J^r S$ such that $\forall X \in \mathfrak{X}_{\infty,c}(E) : J^r S(X) = S(J^r X)$

$J^r S : \mathfrak{X}_{\infty,c}(E) \to (\mathbb{C}, \mathbb{C}^{sm}, s = 1..r) :: J^r S(X) = (S(D_{\alpha_1 .. \alpha_s} X), s = 0...r, \alpha_j = 1...m)$

If the distribution $S \in \mathfrak{X}_{\infty,c}(E)'$ is induced by the m form $\lambda \in \Lambda_m(M; E')$ then its r jet prolongation : $J^r(T(\lambda)) = T(J^r \lambda)$ with $J^r \lambda \in \Lambda_m(M; J^r E')$

Definition 2358 *Let $E_1(M, V_1, \pi_1), E_2(M, V_2, \pi_2)$ be two smooth complex finite dimensional vector bundles on the same real manifold. A differential operator for the distributions $\mathfrak{X}_{\infty c}(E_2)'$ is a map : $D' : J^r \mathfrak{X}_{\infty c}(E_2)' \to \mathfrak{X}_{\infty c}(E_1)'$ such that : $\forall S_2 \in \mathfrak{X}_{\infty c}(E_2)', X_1 \in \mathfrak{X}_{\infty c}(E_1) : D' J^r S_2(X_1) = S_2(D J^r(X_1))$ for a differential operator $D : \mathfrak{X}(J^r E_1) \to \mathfrak{X}(E_2)$*

Theorem 2359 *To any linear differential operator D between the smooth finite dimensional vector bundles on the same real manifold $E_1\left(M, V_1, \pi_1\right), E_2\left(M, V_2, \pi_2\right)$ is associated a differential operator for the distributions : $D': J^r\mathfrak{X}_{\infty c}\left(E_2\right)' \to \mathfrak{X}_{\infty c}\left(E_1\right)'$ which reads :$D(Z) = \sum_{s=0}^{r}\sum_{\alpha_1\ldots\alpha_s=1}^{m} A\left(x\right)^{\alpha_1\ldots\alpha_s} z_{\alpha_1\ldots\alpha_s} \to D' = \sum_{s=0}^{r}\sum_{\alpha_1\ldots\alpha_s=1}^{m}\left(A\left(x\right)^{\alpha_1\ldots\alpha_s}\right)^t D_{\alpha_1\ldots\alpha_s}$*

Proof. The component expression of D is :

$\mathrm{D}(Z) = \sum_{s=0}^{r}\sum_{\alpha_1\ldots\alpha_s=1}^{m} A\left(x\right)^{\alpha_1\ldots\alpha_s} z_{\alpha_1\ldots\alpha_s}$ where $A\left(x\right)^{\alpha_1\ldots\alpha_s} \in \mathcal{L}\left(E_1\left(x\right); E_2\left(x\right)\right)$

Define :

$D'J^r S_2\left(X_1\right) = \sum_{s=0}^{r}\sum_{\alpha_1\ldots\alpha_s=1}^{m}\left(A^{\alpha_1\ldots\alpha_s}\right)^* S_2\left(D_{\alpha_1\ldots\alpha_s}X_1\right)$

with the pull-back $\left(A^{\alpha_1\ldots\alpha_s}\right)^* S_2$ of $S_2 :: \left(A^{\alpha_1\ldots\alpha_s}\right)^* S_2\left(D_{\alpha_1\ldots\alpha_s}X_1\right)$

$= S_2\left(A\left(x\right)^{\alpha_1\ldots\alpha_s}\left(D_{\alpha_1\ldots\alpha_s}X_1\right)\right)$

$D'J^r S_2\left(X_1\right) = \sum_{s=0}^{r}\sum_{\alpha_1\ldots\alpha_s=1}^{m} S_2\left(A\left(x\right)^{\alpha_1\ldots\alpha_s}\left(D_{\alpha_1\ldots\alpha_s}X_1\right)\right)$

$= S_2\left(\sum_{s=0}^{r}\sum_{\alpha_1\ldots\alpha_s=1}^{m} A\left(x\right)^{\alpha_1\ldots\alpha_s}\left(D_{\alpha_1\ldots\alpha_s}X_1\right)\right) = S_2\left(DJ^r\left(X_1\right)\right)$

$A^{\alpha_1\ldots\alpha_s} \in \mathcal{L}\left(E_1\left(x\right); E_2\left(x\right)\right) \Rightarrow \left(A^{\alpha_1\ldots\alpha_s}\right)^* = \left(A^{\alpha_1\ldots\alpha_s}\right)^t \in \mathcal{L}\left(E_2'\left(x\right); E_1'\left(x\right)\right)$

So : $D'J^r S_2 = \sum_{s=0}^{r}\sum_{\alpha_1\ldots\alpha_s=1}^{m}\left(A^{\alpha_1\ldots\alpha_s}\right)^t S_2\circ D_{\alpha_1\ldots\alpha_s}$

$= \sum_{s=0}^{r}\sum_{\alpha_1\ldots\alpha_s=1}^{m}\left(A^{\alpha_1\ldots\alpha_s}\right)^t D_{\alpha_1\ldots\alpha_s}S_2$ ∎

If $E_1 = E_2$ and D is such that : $A\left(x\right)^{\alpha_1\ldots\alpha_s} = a\left(x\right)^{\alpha_1\ldots\alpha_s} I$ with $a\left(x\right)^{\alpha_1\ldots\alpha_s}$ a function and I the identity map then

$S_2\left(A\left(x\right)^{\alpha_1\ldots\alpha_s}\left(D_{\alpha_1\ldots\alpha_s}X_1\left(x\right)\right)\right) = a\left(x\right)^{\alpha_1\ldots\alpha_s} S_2\left(D_{\alpha_1\ldots\alpha_s}X_1\left(x\right)\right)$

$= a\left(x\right)^{\alpha_1\ldots\alpha_s} D_{\alpha_1\ldots\alpha_s}S_2\left(X_1\left(x\right)\right)$

$D'J^r S_2 = \sum_{s=0}^{r}\sum_{\alpha_1\ldots\alpha_s=1}^{ms} a\left(x\right)^{\alpha_1\ldots\alpha_s} D_{\alpha_1\ldots\alpha_s}S_2$

so D' reads with the same coefficients as D.

For the distributions $S \in \mathfrak{X}_{\infty,c}\left(E_2\right)'$ induced by m forms $\lambda_2 \in \Lambda_m\left(M; E_2'\right)$ then :

$S_2 = T\left(\lambda_2\right) \Rightarrow D'J^rT\left(\lambda_2\right)\left(X_1\right) = T\left(\lambda_2\right)\left(DJ^r\left(X_1\right)\right) = D'T\left(J^r\lambda_2\right)\left(X_1\right)$

$D'J^rT\left(\lambda_2\right) = \sum_{s=0}^{r}\sum_{\alpha_1\ldots\alpha_s=1}^{m}\left(A^{\alpha_1\ldots\alpha_s}\right)^t T\left(D_{\alpha_1\ldots\alpha_s}\lambda_2\right)$

$= \sum_{s=0}^{r}\sum_{\alpha_1\ldots\alpha_s=1}^{m} T\left(\left(A^{\alpha_1\ldots\alpha_s}\right)^t D_{\alpha_1\ldots\alpha_s}\lambda_2\right)$

$= T\left(\sum_{s=0}^{r}\sum_{\alpha_1\ldots\alpha_s=1}^{m}\left(A^{\alpha_1\ldots\alpha_s}\right)^t D_{\alpha_1\ldots\alpha_s}\lambda_2\right)$

that we can write : $D'J^rT\left(\lambda_2\right) = T\left(D^t\lambda_2\right)$

with $D^t = \sum_{s=0}^{r}\sum_{\alpha_1\ldots\alpha_s=1}^{m}\left(A^{\alpha_1\ldots\alpha_s}\right)^t D_{\alpha_1\ldots\alpha_s}$

Example : $E_1 = E_2, r = 1$

$D\left(J^1X\right) = \sum_{ij}\left(\left[A\right]^i_j X^j + \left[B^\alpha\right]^i_j \partial_\alpha X^j\right) e_i\left(x\right)$

$D'T\left(J^1\lambda\right) = T\left(\sum_{ij}\left(\left[A\right]^i_j \lambda_i + \left[B^\alpha\right]^i_j \partial_\alpha\lambda_i\right) e^j\left(x\right)\right)$

In particular when the operator is scalar :

Theorem 2360 *To any scalar linear differential operator D acting on a space of functions $F\subset C_r\left(M;\mathbb{C}\right)$ over a smooth manifold the associated linear differential operator D' acting on the distributions F' has same coefficients as D. And for the distributions induced by m forms $\lambda \in \Lambda_m\left(M;\mathbb{C}\right)$ we have :*
$D'J^rT\left(\left(\lambda_0 d\xi^1 \wedge \ldots \wedge d\xi^m\right)\right) = T\left(\left(D\lambda_0\right) d\xi^1 \wedge \ldots \wedge d\xi^m\right)$

So it is customary to look at λ as a function $\lambda \in C\left(M;\mathbb{C}\right)$. However, to be rigorous, we cannot forget that λ_0 is defined through a family of functions which change according to the cover of M.

30.1.3 Adjoint of a differential operator

Definition

The definition of the adjoint of a differential operator follows the general definition of the adjoint of a linear map with respect to a scalar product.

Definition 2361 *A map $\widehat{D}^{*} : \mathfrak{X}\left(E_2\right) \to \mathfrak{X}\left(E_1\right)$ is the adjoint of a linear differential operator D between the smooth finite dimensional vector bundles on the same real manifold $E_1\left(M, V_1, \pi_1\right), E_2\left(M, V_2, \pi_2\right)$ endowed with scalar products G_1, G_2 on sections if :*

$$\forall X_1 \in \mathfrak{X}\left(E_1\right), X_2 \in \mathfrak{X}\left(E_2\right) : G_2\left(\widehat{D}\left(X_1\right), X_2\right) = G_1\left(X_1, \widehat{D}^{*}X_2\right)$$

Definition 2362 *A linear differential operator on a vector bundle is said to be **self adjoint** if $D = D^{*}$*

An adjoint map is not necessarily a differential operator. If D has bounded smooth coefficients, the map : $\widehat{D} = D \circ J^r : W^{2,r}\left(E\right) \to W^{2,r}\left(E\right)$ is continuous on the Hilbert space $W^{2,r}\left(E\right)$, so $\widehat{D}$ has an adjoint $\widehat{D}^{*} \in \mathcal{L}\left(W^{2,r}\left(E\right); W^{2,r}\left(E\right)\right)$. However this operator is not necessarily local. Indeed if D is scalar, it is also a pseudo-differential operator, and as such its adjoint is a pseudo-differential operator, whose expression is complicated (see next section).

Condition of existence

In the general conditions of the definition :

i) The scalar products induce antilinear morphisms :

$\Theta_1 : \mathfrak{X}\left(E_1\right) \to \mathfrak{X}\left(E_1\right)' :: \Theta_1\left(X_1\right)\left(Y_1\right) = G_1\left(X_1, Y_1\right)$

$\Theta_2 : \mathfrak{X}\left(E_2\right) \to \mathfrak{X}\left(E_2\right)' :: \Theta_2\left(X_2\right)\left(Y_2\right) = G_2\left(X_2, Y_2\right)$

They are injective, but not surjective, because the vector spaces are infinite dimensional.

ii) To the operator : $\widehat{D} : \mathfrak{X}\left(E_1\right) \to \mathfrak{X}\left(E_2\right)$ the associated operator on distributions (which always exists) reads:

$$\widehat{D}' : \mathfrak{X}\left(E_2\right)' \to \mathfrak{X}\left(E_1\right)' :: \widehat{D}'S_2\left(X_1\right) = S_2\left(\widehat{D}\left(X_1\right)\right)$$

iii) Assume that, at least on some vector subspace F of $\mathfrak{X}\left(E_2\right)$ there is $X_1 \in \mathfrak{X}\left(E_1\right) :: \Theta_1\left(X_1\right) = \widehat{D}' \circ \Theta_2\left(X_2\right)$ then the operator : $\widehat{D}^{*} = \Theta_1^{-1} \circ \widehat{D}' \circ \Theta_2$ is such that :

$$G_1\left(X_1, \widehat{D}^{*}X_2\right) = G_1\left(X_1, \Theta_1^{-1} \circ \widehat{D}' \circ \Theta_2\left(X_2\right)\right)$$

$$= \overline{G_1\left(\Theta_1^{-1} \circ \widehat{D}' \circ \Theta_2\left(X_2\right), X_1\right)} = \overline{\widehat{D}' \circ \Theta_2\left(X_2\right)\left(X_1\right)}$$

$$= \overline{\Theta_2\left(X_2\right)\left(\widehat{D}\left(X_1\right)\right)} = \overline{G_2\left(X_2, \widehat{D}\left(X_1\right)\right)} = G_2\left(\widehat{D}\left(X_1\right), X_2\right)$$

$\widehat{D}^{*} = \Theta_1^{-1} \circ \widehat{D}' \circ \Theta_2$ is C-linear, this is an adjoint map and a differential operator. As the adjoint, when it exists, is unique, then $\widehat{D}^{*}$ is the adjoint of D

on F. So whenever we have an inverse Θ_1^{-1} we can define an adjoint. This leads to the following theorem, which encompasses the most usual cases.

Fundamental theorem

Theorem 2363 *A linear differential operator D between the smooth finite dimensional vector bundles on the same real manifold $E_1\left(M, V_1, \pi_1\right), E_2\left(M, V_2, \pi_2\right)$ endowed with scalar products G_1, G_2 on sections defined by scalar products g_1, g_2 on the fibers and a volume form ϖ_0 on M has an adjoint which is a linear differential operator whith same order as D.*

The procedure above can be implemented because Θ_1 is inversible.
Proof. Fiberwise the scalar products induce antilinear isomorphisms :

$\tau : E\left(x\right) \to E'\left(x\right) :: \tau\left(X\right)\left(x\right) = \sum_{ij} g_{ij}\left(x\right) \overline{X}^i\left(x\right) e^j\left(x\right) \Rightarrow \tau\left(X\right)\left(Y\right) = \sum_{ij} g_{ij} \overline{X}^i Y^j = g\left(X, Y\right)$

$\tau^{-1} : E'\left(x\right) \to E\left(x\right) :: \tau\left(\lambda\right)\left(x\right) = \sum_{ij} \overline{g}^{ki}\left(x\right) \overline{\lambda}_k\left(x\right) e_i\left(x\right) \Rightarrow g\left(\tau\left(\lambda\right), Y\right) = \sum_{ij} g_{ij} g^{ki} \lambda_k Y^j = \lambda\left(Y\right)$

with $\overline{g}^{ki} = g^{ik}$ for a hermitian form.
The scalar products for sections are : $G\left(X, Y\right) = \int_M g\left(x\right)\left(X\left(x\right), Y\left(x\right)\right) \varpi_0$
and we have the maps :

$T : \Lambda_m\left(M; E'\right) \to \mathfrak{X}\left(E\right)' :: T\left(\lambda\right)\left(Y\right) = \int_M \lambda\left(x\right)\left(Y_1\left(x\right)\right) \varpi_0 = \int_M g\left(x\right)\left(X\left(x\right), Y\left(x\right)\right) \varpi_0$

$\Theta : \mathfrak{X}\left(E_1\right) \to \mathfrak{X}\left(E\right)' :: \Theta\left(X\right) = T\left(\tau\left(X\right) \otimes \varpi_0\right)$

If $D(Z) = \sum_{s=0}^{r} \sum_{\alpha_1...\alpha_s=1}^{m} A\left(x\right)^{\alpha_1...\alpha_s} z_{\alpha_1...\alpha_s}$ where $A\left(x\right)^{\alpha_1...\alpha_s} \in \mathcal{L}\left(E_1\left(x\right); E_2\left(x\right)\right)$

$\widehat{D}' \circ \Theta_2\left(X_2\right) = D' J^r T_1\left(\tau_2\left(X_2\right) \otimes \varpi_0\right) = T_1\left(D^t \tau_2\left(X_2\right) \otimes \varpi_0\right)$

$D^t \tau_2\left(X_2\right) = \sum_{s=0}^{r} \sum_{\alpha_1...\alpha_s=1}^{m} \left(A^{\alpha_1...\alpha_s}\right)^t \left(D_{\alpha_1...\alpha_s}\left(\sum_{ij} g_{2ij}\left(x\right) \overline{X_2}^i\left(x\right)\right) e_2^j\left(x\right)\right)$

$\left(A^{\alpha_1...\alpha_s}\right)^t \left(D_{\alpha_1...\alpha_s}\left(\sum_{ij} g_{2ij}\left(x\right) \overline{X_2}^i\left(x\right)\right) e_2^j\left(x\right)\right)$

$= \sum_j \left[A^{\alpha_1...\alpha_s}\right]_j^k D_{\alpha_1...\alpha_s}\left(\sum_{kl} g_{2lk}\left(x\right) \overline{X_2}^l\left(x\right)\right) e_1^j\left(x\right)$

$= \sum_{kj} \left[A^{\alpha_1...\alpha_s}\right]_j^k \Upsilon_{k\alpha_1...\alpha_s}\left(J^s \overline{X_2}\right) e_1^j\left(x\right)$

where $\Upsilon_{k\alpha_1...\alpha_s}$ is a s order linear differential operator on $\overline{X_2}$

$D^t \tau_2\left(X_2\right) = \sum_{s=0}^{r} \sum_{\alpha_1...\alpha_s=1}^{m} \sum_{jk} \left[A^{\alpha_1...\alpha_s}\right]_j^k \Upsilon_{k\alpha_1...\alpha_s}\left(J^s \overline{X_2}\right) e_1^j\left(x\right)$

$\widehat{D}^*\left(X_2\right) = \sum_{ijk} g_1^{ik} \sum_{s=0}^{r} \sum_{\alpha_1...\alpha_s=1}^{m} \overline{\left[A^{\alpha_1...\alpha_s}\right]_k^j} \Upsilon_{j\alpha_1...\alpha_s}\left(J^s \overline{X_2}\right) e_{1i}\left(x\right)$

So we have a r order linear differential operator on X_2. ∎

Example : $E_1 = E_2, r = 1$

$D\left(J^1 X\right) = \sum_{ij} \left(\left[A\right]_j^i X^j + \left[B^\alpha\right]_j^i \partial_\alpha X^j\right) e_i\left(x\right)$

$D'T\left(J^1 \lambda\right) = T\left(\sum_{ij} \left(\left[A\right]_j^i \lambda_i + \left[B^\alpha\right]_j^i \partial_\alpha \lambda_i\right) e^j\left(x\right)\right)$

$\tau\left(X\right)\left(x\right) = \sum_{pq} g_{pq}\left(x\right) \overline{X}^p\left(x\right) e^q\left(x\right)$

$D'T\left(J^1 \tau\left(X\right)\right) = T\left(\sum \left(\left[A\right]_j^q g_{pq} \overline{X}^p + \left[B^\alpha\right]_j^q \left(\left(\partial_\alpha g_{pq}\right) \overline{X}^p + g_{pq} \partial_\alpha \overline{X}^p\right)\right) e^j\left(x\right)\right)$

$\widehat{D}^*\left(X\right) = \sum g^{ij} \left(\overline{\left[A\right]}_j^q \overline{g}_{pq} X^p + \overline{\left[B^\alpha\right]}_j^q \left(\left(\partial_\alpha \overline{g}_{pq}\right) X^p + \overline{g}_{pq} \partial_\alpha X^p\right)\right) e_i\left(x\right)$

692

$$\widehat{D}^{*}(X) = \sum \left(\left(\left[g^{-1} \right] \left[A \right]^{*} \left[g \right] + \left[g^{-1} \right] \left[B^{\alpha} \right]^{*} \left[\partial_{\alpha} g \right] \right) X + \left(\left[g^{-1} \right] \left[B^{\alpha} \right]^{*} \left[g \right] \right) \partial_{\alpha} X \right)^{i} e_{i}(x)$$

Notice that the only requirement on M is a volume form, which can come from a non degenerate metric, but not necessarily. And this metric is not further involved.

30.1.4 Symbol of a linear differential operator

Definition

Definition 2364 *The **symbol** of a linear r order differential operator D :* $\mathfrak{X}(J^{r}E_{1}) \to \mathfrak{X}(E_{2})$ *is the r order symmetric polynomial map :*
$P(x) : \mathbb{R}^{m*} \to E_{2}(x) \otimes E_{1}(x)^{*} :: P(x)(u)$
$= \sum_{i=1}^{n_{1}} \sum_{j=1}^{n_{2}} \sum_{s=0}^{r} \sum_{\alpha_{1}...\alpha_{s}} [A(x)^{\alpha_{1}...\alpha_{s}}]_{i}^{j} u_{\alpha_{1}}..u_{\alpha_{s}} e_{2j}(x) \otimes e_{1}^{i}(x)$ *with $e_{1}^{i}(x), e_{2,j}(x)$ holonomic bases of $E_{1}(x)^{*}, E_{2}(x)$ and $u = (u_{1},...u_{m}) \in \mathbb{R}^{m*}$*
*The r order part of $P(x)(u)$ is the **principal symbol** of D : $\sigma_{D}(x,u) \in$* $\mathcal{L}(E_{1}(x); E_{2}(x))$

$\sigma_{D}(x,u) = \sum_{i=1}^{n_{1}} \sum_{j=1}^{n_{2}} \sum_{\alpha_{1}...\alpha_{r}} A(x)_{i}^{j,\alpha_{1}...\alpha_{r}} u_{\alpha_{1}}..u_{\alpha_{r}} e_{2j}(x) \otimes e_{1}^{i}(x)$
Explanation : A linear operator is a map : $D(x) \in \mathcal{L}(J^{r}E_{1}(x); E_{2}(x)) = J^{r}E_{1}(x)^{*} \otimes E_{2}(x)$. $J^{r}E_{1}(x)$ can be identified to $E_{1}(x) \otimes \sum_{s=0}^{r} \odot^{s}\mathbb{R}^{m}$ so we can see D as a tensor $D \in \sum_{s=0}^{r} \odot^{s}\mathbb{R}^{m} \otimes E_{1}^{*} \otimes E_{2}$ which acts on vectors of $\mathbb{R}^{m*}$.
Conversely, given a r order symmetric polynomial it defines uniquely a linear differential operator.
Formally, it sums up to replace $\frac{\partial}{\partial \xi^{\alpha}}$ by u_{α}.
Notice that : $X \in \mathfrak{X}(E_{1}) : P(x)(u)(X) = D(Z)$
with $Z = \sum_{i=1}^{n_{1}} \sum_{s=0}^{r} \sum_{\alpha_{1}...\alpha_{s}} u_{\alpha_{1}}..u_{\alpha_{s}} X^{i}(x) e_{1i}(x)$

Composition of operators

Theorem 2365 *If D_{1}, D_{2} are two order r differential on vector bundles on the same manifold, the principal symbol of their compose $D_{1} \circ D_{2}$ is a 2r order symmetric polynomial map given by : $[\sigma_{D_{1} \circ D_{2}}(x,u)] = [\sigma_{D_{1}}(x,u)][\sigma_{D_{2}}(x,u)]$*

it is not true for the other components of the symbol
Proof. $\sigma_{D_{1} \circ D_{2}}(x,u)$
$= \sum_{i,j,k=1}^{n} \sum_{\alpha_{1}...\alpha_{r}} \sum_{\beta_{1}...\beta_{r}} A_{1}(x)_{k}^{j,\alpha_{1}...\alpha_{r}} A_{2}(x)_{i}^{k,\beta_{1}...\beta_{r}} u_{\alpha_{1}}..u_{\alpha_{r}} u_{\beta_{1}}..u_{\beta_{r}} e_{3j}(x) \otimes e_{1}^{i}(x)$
In matrix notation:
$[\sigma_{D_{1} \circ D_{2}}(x,u)] = \left(\sum_{\alpha_{1}...\alpha_{r}} u_{\alpha_{1}}..u_{\alpha_{r}} [A_{1}(x)]^{\alpha_{1}...\alpha_{r}} \right) \left(\sum_{\beta_{1}...\beta_{r}} u_{\beta_{1}}..u_{\beta_{r}} [A_{2}(x)]^{\beta_{1}...\beta_{r}} \right)$
∎

Image of a differential operator

A constant map $L : \mathfrak{X}(E_1) \to \mathfrak{X}(E_2)$ between two vector bundles $E_1(M, V_1, \pi_1)$, $E_2(M, V_2, \pi_2)$ on the same manifold is extended to a map $\widehat{L} : \mathfrak{X}(J^r E_1) \to \mathfrak{X}(J^r E_2) :: \widehat{L} = J^r \circ L$

$\widehat{L}\left(Z^i_{\alpha_1...\alpha_s} e^{\alpha_1..\alpha_s}_{1i}\right) = \sum_j [L]^i_j Z^j_{\alpha_1...\alpha_s} e^{\alpha_1..\alpha_s}_{2i}$

A differential operator $D : J^r E_1 \to E_1$ gives a differential operator : $D_2 : J^r E_1 \to E_2 :: D_2 = D \circ J^r \circ L = \widehat{D} \circ \widehat{L}$

The symbol of D_2 is : $\sigma_{D_2}(x, u) = \sigma_D(x, u) \circ L \in \mathcal{L}(E_1(x); E_2(x))$

$P(x)(u) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{s=0}^{r} \sum_{\alpha_1...\alpha_s} [A(x)^{\alpha_1...\alpha_s}]^j_k [L]^k_i u_{\alpha_1}..u_{\alpha_s} e_{2j}(x) \otimes e^i_1(x)$

Adjoint

Theorem 2366 *If we have scalar products $\langle\rangle$ on two vector bundles $E_1(M, V_1, \pi_1)$, $E_2(M, V_2, \pi_2)$, and two linear r order differential operators $D_1 : \mathfrak{X}(J^r E_1) \to \mathfrak{X}(E_2), D_2 : \mathfrak{X}(J^r E_2) \to \mathfrak{X}(E_1)$ such that D_2 is the adjoint of D_1 then : $P_{D_1}(x, u) = P_{D_2}(x, u)^*$ the adjoint with respect to the scalar products.*

Proof. We have : $\forall X_1 \in \mathfrak{X}(J^r E_1), X_2 \in \mathfrak{X}(J^r E_2) : \langle D_1 J^r X_1, X_2 \rangle_{E_2} = \langle X_1, D_2 J^r X_2 \rangle_{E_1}$

take : $X_1(x) = X_1 \exp\left(\sum_{\alpha=1}^m u_\alpha \xi^\alpha\right)$ with X some fixed vector in V_1.

$D_{\alpha_1..\alpha_s} X_1(x) = u_{\alpha_1}..u_{\alpha_s} X_1 \exp\left(\sum_{\alpha=1}^m u_\alpha \xi^\alpha\right)$

And similarly for X_2

$D_1 J^r X_1 = P_{D_1}(x, u) X_1 \exp\left(\sum_{\alpha=1}^m u_\alpha \xi^\alpha\right), D_1 J^r X_2 = P_{D_2}(x, u) X_2 \exp\left(\sum_{\alpha=1}^m u_\alpha \xi^\alpha\right)$

$\forall X_1 \in V_1, X_2 \in X_2 : \langle P_{D_1}(x, u) X_1, X_2 \rangle = \langle X_1, P_{D_2}(x, u) X_2 \rangle$ ∎

Elliptic operator

Definition 2367 *For a linear r order differential operator $D : \mathfrak{X}(J^r E) \to \mathfrak{X}(E)$ with principal symbol $\sigma_D(x, u)$ over a vector bundle $E(M, V, \pi)$:*

*The **characteristic set** of D is the set : $Char(D) \subset M \times \mathbb{R}^m$ where $\sigma_D(x, u)$ is an isomorphism.*

*D is said to be **elliptic (**or weakly elliptic**)** if its principal symbol $\sigma_D(x, u)$ is an isomorphism whenever $u \neq 0$.*

If V is a Hilbert space with inner product $\langle\rangle$:

*D is said to be **strongly** (or uniformly) **elliptic** if:*

$\exists C \in \mathbb{R} : \forall x \in M, \forall u, v \in T_x M^*, \langle v, v \rangle = 1 : \langle \sigma_D(x, u) v, v \rangle \geq C \langle v, v \rangle$

*D is said to be **semi elliptic** if:*

$\forall x \in M, \forall u, v \in T_x M^*, \langle v, v \rangle = 1 : \langle \sigma_D(x, u) v, v \rangle \geq 0$

Theorem 2368 *A necessary condition for an operator to be strongly elliptic is dim(M) even.*

Theorem 2369 *The composition of two weakly elliptic operators is still a weakly operator.*

Theorem 2370 *A weakly elliptic operator on a vector bundle with compact base has a (non unique) parametrix.*

As a consequence the kernel of such an operator is finite dimensional in the space of sections and is a Fredholm operator.

Theorem 2371 *If is D weakly elliptic, then D*D, DD* are weakly elliptic.*

Proof. by contradiction : let be $u \neq 0, X \neq 0 : \sigma_{D^* \circ D}(x, u) X = 0$
$\sigma_{D^* \circ D}(x, u) = \sigma_{D^*}(x, u) \sigma_D(x, u) = \sigma_D(x, u)^* \sigma_D(x, u)$ ∎

Definition 2372 *A linear differential operator D' on a space of distributions is said to be **hypoelliptic** if the singular support of D'(S) is included in the singular support of S.*

So whenever $f \in W$ then $\exists g \in W : T(g) = D(T(f))$. If D is strongly elliptic then the associated operator D' is hypoelliptic. The laplacian and the heat kernel are hypoelliptic.
Remark : contrary at what we could expect a hyperbolic operator is not a differential operator such that the principal symbol is degenerate (see PDE).

Index of an operator
There is a general definition of the index of a linear map (see Banach spaces). For differential operators we have :

Definition 2373 *A linear differential operator $D : F_1 \to \mathfrak{X}(E_2)$ between two vector bundle E_1, E_2 on the same base, with $F_1 \subset \mathfrak{X}(J^r E_1)$ is a Fredholm operator if $\ker D$ and $\mathfrak{X}(E_2)/D(F_1)$ are finite dimensional. The index (also called the analytical index) of D is then : Index(D)=$\dim \ker D - \dim \mathfrak{X}(E_2)/D(F_1)$*

Notice that the condition applies to the full vector space of sections (and not fiberwise).

Theorem 2374 *A weakly elliptic operator on a vector bundle with compact base is Freholm*

This is the starting point for a set of theorems such that the Atiyah-Singer index theorem. For a differential operator D:$J^r E_1 \to E_2$ between vector bundles on the same base manifold M, one can define a topological index (this is quite complicated) which is an integer deeply linked to topological invariants of M. The most general theorem is the following :

Theorem 2375 *Teleman index theorem : For any abstract elliptic operator on a closed, oriented, topological manifold, its analytical index equals its topological index.*

It means that the index of a differential operator depends deeply of the base manifold.
We have more specifically the followings :

Theorem 2376 *(Taylor 2 p.264) If D is a linear elliptic first order differential operator between two vector bundles E_1, E_2 on the same compact manifold, then $D : W^{2,r+1}(E_1) \to W^{2,r}(E_2)$ is a Fredholm operator, $\ker D$ is independant of r, and $D^t : W^{2,-r}(E_1) \to W^{2,-r-1}(E_2)$ has the same properties. If D_s is a family of such operators, continuously dependant of the parameter s, then the index of D_s does not depend on s.*

Theorem 2377 *(Taylor 2 p.266) If D is a linear elliptic first order differential operator on a vector bundle with base a compact manifold with odd dimension then Index(D)=0*

On any oriented manifold M the exterior algebra can be split between the forms of even E or odd F order. The sum $D=d + \delta$ of the exterior differential and the codifferential exchanges the forms between E and F. If M is compact the topological index of D is the Euler characteristic of the Hodge cohomology of M, and the analytical index is the Euler class of the manifold. The index formula for this operator yields the Chern-Gauss-Bonnet theorem.

30.1.5 Linear differential operators and Fourier transform

Theorem 2378 *For a linear scalar differential operator D over $\mathbb{R}^m$: $D : C_r(\mathbb{R}^m; \mathbb{C}) \to C_0(\mathbb{R}^m; \mathbb{C})$:*

for $f \in S(\mathbb{R}^m)$: $Df = (2\pi)^{-m/2} \int_{\mathbb{R}^m} P(x, it)\widehat{f}(t)e^{i\langle t,x \rangle}dt$ where P is the symbol of D

As $\widehat{f} \in S(\mathbb{R}^m)$ and $P(x, it)$ is a polynomial in t, then $P(x, it)\widehat{f}(t) \in S(\mathbb{R}^m)$ and $Df = \mathcal{F}_t^*\left(P(x, it)\widehat{f}(t)\right)$

As $\widehat{f} \in S(\mathbb{R}^m)$ and the induced distribution $T\left(\widehat{f}\right) \in S(\mathbb{R}^m)'$ we have :
$Df = (2\pi)^{-m/2} T\left(\widehat{f}\right)_t \left(P(x, it)e^{i\langle t,x \rangle}\right)$

Proof. The Fourier transform is a map : $\mathcal{F}:S(\mathbb{R}^m) \to S(\mathbb{R}^m)$ and

$D_{\alpha_1...\alpha_s}f = \mathcal{F}^*\left(i^r (t_{\alpha_1}..t_{\alpha_s})\widehat{f}\right)$
So : $Df = \sum_{s=0}^r \sum_{\alpha_1...\alpha_s=1}^m A(x)^{\alpha_1...\alpha_s} D_{\alpha_1...\alpha_s}f(x)$
$= (2\pi)^{-m/2} \sum_{s=0}^r \sum_{\alpha_1...\alpha_s=1}^m i^s A(x)^{\alpha_1...\alpha_s} \int_{\mathbb{R}^m} t_{\alpha_1}...t_{\alpha_s}\widehat{f}(t)e^{i\langle t,x \rangle}dt$
$= (2\pi)^{-m/2} \int_{\mathbb{R}^m} \left(\sum_{s=0}^r \sum_{\alpha_1...\alpha_s=1}^m A(x)^{\alpha_1...\alpha_s} (it_{\alpha_1})...(it_{\alpha_s})\right)\widehat{f}(t)e^{i\langle t,x \rangle}dt$
$= (2\pi)^{-m/2} \int_{\mathbb{R}^m} P(x, it)\widehat{f}(t)e^{i\langle t,x \rangle}dt$ ∎

Theorem 2379 *For a linear scalar differential operator D over $\mathbb{R}^m$ and for $f \in S(\mathbb{R}^m), A(x)^{\alpha_1...\alpha_s} \in L^1(\mathbb{R}^m, dx, \mathbb{C})$:*

$Df = (2\pi)^{-m} \int_{\mathbb{R}^m} \sum_{s=0}^r \sum_{\alpha_1...\alpha_s=1}^m \left(\widehat{A}^{\alpha_1...\alpha_s} * \left((it_{\alpha_1})...(it_{\alpha_s})\widehat{f}\right)\right) e^{i\langle t,x \rangle}dt$

Proof. $A(x)^{\alpha_1...\alpha_s} D_{\alpha_1...\alpha_s}f(x)$
$= \mathcal{F}^*\left((2\pi)^{-m/2} \mathcal{F}(A(x)^{\alpha_1...\alpha_s}) * \mathcal{F}(D_{\alpha_1...\alpha_s}f(x))\right)$

696

$$= (2\pi)^{-m/2} \mathcal{F}^* \left(\widehat{A}(x)^{\alpha_1...\alpha_s} * \left((it_{\alpha_1})...(it_{\alpha_s}) \widehat{f}(x) \right) \right)$$
$$= (2\pi)^{-m} \int_{\mathbb{R}^m} \widehat{A}(x)^{\alpha_1...\alpha_s} * \left((it_{\alpha_1})...(it_{\alpha_s}) \widehat{f}(x) \right) e^{i\langle t,x\rangle} dt \quad \blacksquare$$

Theorem 2380 *For a linear differential operator D' on the space of tempered distributions $S(\mathbb{R}^m)'$: $D'S = \mathcal{F}^* \left(P(x,it)\widehat{S} \right)$ where P is the symbol of D'*

Proof. The Fourier transform is a map : $\mathcal{F}$:$S(\mathbb{R}^m)' \to S(\mathbb{R}^m)'$ and $D_{\alpha_1...\alpha_s}S = \mathcal{F}^* \left(i^r (t_{\alpha_1}..t_{\alpha_s}) \widehat{S} \right)$

So : $D'S = \sum_{s=0}^r \sum_{\alpha_1...\alpha_s=1}^m A(x)^{\alpha_1...\alpha_s} D_{\alpha_1...\alpha_s}S$

$= \sum_{s=0}^r \sum_{\alpha_1...\alpha_s=1}^m A(x)^{\alpha_1...\alpha_s} \mathcal{F}^* \left(it_{\alpha_1}...it_{\alpha_s}\widehat{S} \right)$

$= \mathcal{F}^* \left(\sum_{s=0}^r \sum_{\alpha_1...\alpha_s=1}^m A(x)^{\alpha_1...\alpha_s} (it_{\alpha_1})...(it_{\alpha_s}) \widehat{S} \right)$

$= \mathcal{F}^* \left(P(x,it)\widehat{S} \right) \quad \blacksquare$

Whenever $S = T(f), f \in L^1(\mathbb{R}^m, dx, \mathbb{C}) : \widehat{S} = T\left(\widehat{f}\right) = (2\pi)^{-m/2} T_t \left(S_x \left(e^{-i\langle x,t\rangle} \right) \right)$
and we get back the same formula as 1 above.

30.1.6 Fundamental solution

Fundamental solutions and Green functions are ubiquitous tools in differential eqations, which come in many flavors. We give here a general definition, which is often adjusted with respect to the problem in review. Fundamental solutions for classic operators are given in the present section and in the section on PDE.

Definition

Definition 2381 *A **fundamental solution** at a point y$\in$ M of a scalar linear differential operator D' on distributions F' on a space of functions F over a manifold M is a distribution $U_y \in F'$ such that : $D'U_y = \delta_y$*

*If there is a map $T : W \to F'$, a **Green's function** for D' is a map $G \in C(M;W)$ such that : $\forall y \in M : T(G(y))$ is a fundamental solution at y.*

It is usually written as G(x,y) and : $D'(x) T_y(G(x,y)) = \delta_y(x)$
Notice that :
i) the result is always a Dirac distribution, so we need a differential operator acting on distributions
ii) a Dirac distribution is characterized by a support in one point. On a manifold it can be any point.
iii) δ_y must be in F'

Fundamental solution of PDE

Its interest, and its name, come from the following:.

Theorem 2382 *If U_0 is a fundamental solution at 0 for a scalar linear differential operator D' on $C_{\infty c}(O;\mathbb{C})'$ with an open O of $\mathbb{R}^m$, then for any $S\in \left(C_{\infty c}(O;\mathbb{C})'\right)_c$, $U_0 * S$ is a solution of D'X=S*

Proof. $D'(U_0 * S) = D'(U_0) * S = \delta_0 * S = S$ ∎

Theorem 2383 *If F is a Fréchet space of complex functions in an open O of $\mathbb{R}^m$, D a linear differential operator on F and U(y) a fundamental solution at y of the associated differential operator D', then for any compactly supported function f, $u = U(y)_t (f(x + y - t))$ is a solution of $Du = f$*

Proof. As we are in $\mathbb{R}^m$ we can use convolution. The convolution of $U(y)$ and T(f) is well defined, and :

$D'(U(y) * T(f) * \delta_{-y}) = (D'U(y)) * T(f) * \delta_{-y} = \delta_y * \delta_{-y} * T(f) = \delta_0 * T(f) = T(f)$

$U(y) * T(f) * \delta_{-y} = U(y) * T_x(f(x + y)) = T(U(y)_t (f(x + y - t))) = T(u)$

and $u \in C_\infty(O;\mathbb{C})$ so $D'T(u) = T(Du)$ and

$T(f) = T(Du) \Rightarrow Du = f$ ∎

If there is a Green's function then : $u(x) = \int_O g(x,y) f(y) dy$

Operators depending on a parameter

Let F be a Fréchet space of complex functions on a manifold M, J some interval in $\mathbb{R}$. We consider functions in F depending on a parameter t. We have a family D(t) of linear scalar differential operators $D(t) : F \to F$, depending on the same parameter t∈ J.

We can see this as a special case of the above, with $J \times M$ as manifold. A fundamental solution at (t,y) with y some fixed point in M is a family U(t,y) of distributions acting on $C_{\infty c}(J \times M;\mathbb{C})'$ and such that : $D'(t) U(t,y) = \delta_{(t,y)}$.

Fourier transforms and fundamental solutions

The Fourier transform gives a way to fundamental solutions for scalar linear differential operators.

Let D' be a linear differential operator on the space of tempered distributions $S(\mathbb{R}^m)'$ then : $D'S = \mathcal{F}^*\left(P(x,it)\widehat{S}\right)$ where P is the symbol of D'. If U is a fundamental solution of D' at 0 : D'U=$\delta_0 = \mathcal{F}^*\left(P(x,it)\widehat{U}\right)$ then $P(x,it)\widehat{U} = (2\pi)^{-m/2} T(1)$ which gives usually a convenient way to compute U.

If D' is a linear differential operator on $\mathbb{R}^m$ with constant coefficients there is a fundamental solution :

$P(x,it)\widehat{U} = (2\pi)^{-m/2} T(1)$ reads : $\left(\sum_{s=0}^r \sum_{\alpha_1...\alpha_s=1}^m A^{\alpha_1...\alpha_s}(-i)^s t^\alpha\right)\widehat{U} = (2\pi)^{-m/2}$ from where we can get $\widehat{U}$ and U by inverse Fourier transform.

Example :

$D = \frac{\partial^2}{\partial x^2}$ then $U = xH$ with H the Heaviside function. It is easy to check that :

$\frac{\partial}{\partial x}(xH) = H + x\delta_0$

$\frac{\partial^2}{\partial x^2}(xH) = 2\delta_0 + x\delta_0'$

$\frac{\partial^2}{\partial x^2}(xH)(\varphi) = 2\varphi(0) - \delta_0\left(\frac{\partial}{\partial x}(x\varphi)\right) = 2\varphi(0) - \delta_0(\varphi + x\varphi') = \varphi(0) = \delta_0(\varphi)$

If D$= \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$ then $U = \sum_{i=1}^{n} x_i H(x_i)$

Parametrix

A parametrix is a proxy" for a fundamental solution.

Let F be a Fréchet space of complex functions, F' the associated space of distributions, W a space of functions such that : $T : W \to F'$, a scalar linear differential operator D' on F'. A **parametrix** for D' is a distribution $U \in F'$: $D'(U) = \delta_y + T(u)$ with $u \in W$.

30.1.7 Connection and differential operators

(see Connections on vector bundle)

Covariant derivative as a differential operator

A covariant derivative induced by a linear connection on a vector bundle $E(M, V, \pi)$ is a map acting on sections of E.

$\nabla : \mathfrak{X}(E) \to \Lambda_1(M; E) : \nabla S = \sum_{i\alpha}\left(\partial_\alpha X^i + \Gamma_{\alpha j}^i(x) X^j\right) e_i(x) \otimes dx^\alpha$

We can define the tensorial product of two vector bundles, so we have $TM^* \otimes E$ and we can consider the covariant derivative as a linear differential operator :

$\widehat{\nabla} : \mathfrak{X}(J^1 E) \to \mathfrak{X}(E \otimes TM^*) : \widehat{\nabla} Z = \sum_{i\alpha}\left(Z_\alpha^i + \Gamma_{\alpha j}^i(x) Z^j\right) e_i(x) \otimes dx^\alpha$

and for a section of E, that is a vector field : $\widehat{\nabla} \circ J^1 = \nabla$

Notice that $E \otimes TM^*$ involves the cotangent bundle, but $J^1 E = J^1 E(E, \mathbb{R}^{m*} \otimes V, \pi)$ does not :

$X \in E(x) \otimes T_x M^* : X = \sum_{i\alpha} X_\alpha^i e_i(x) \otimes dx^\alpha$

$Z \in J^1 E(x) : Z = \left(\left(\sum_i Z^i e_i(x), \sum_{i\alpha} Z_\alpha^i e_i^\alpha(x)\right)\right)$

We can define higher order connections in the same way as r linear differential operators acting on E :

$\nabla^r : \mathfrak{X}(J^r E) \to \Lambda_1(M; E) : \nabla^r Z = \sum_{s=0}^r \sum_{\beta_1..\beta_s} \Gamma_{\alpha j}^{\beta_1..\beta_s i}(x) Z_{\beta_1..\beta_s}^j e_i(x) \otimes dx^\alpha$

$\nabla^r X = \sum_{s=0}^r \sum_{\beta_1..\beta_s} \Gamma_{\alpha j}^{\beta_1..\beta_s i}(x) \left(\partial_{\beta_1..\beta_s}^j X^j\right) e_i(x) \otimes dx^\alpha$

Exterior covariant derivative

1. The curvature of the connection is a linear map, but not a differential operator :

$\widehat{\Omega} : \mathfrak{X}(E) \to \Lambda_2(M; E) :: \widehat{\Omega}(x)(X(x))$

$= \sum_{\alpha\beta} \sum_{j\in I}\left(-\partial_\alpha \Gamma_{j\beta}^i(x) + \sum_{k\in I} \Gamma_{j\alpha}^k(x)\Gamma_{k\beta}^i(x)\right) X^j(x) dx^\alpha \wedge dx^\beta \otimes e_i(x)$

2. The exterior covariant derivative is the map :

$\nabla_e : \Lambda_r(M;E) \to \Lambda_{r+1}(M;E) :: \nabla_e \varpi = \sum_i \left(d\varpi^i + \sum_j \Gamma^i_{\alpha j} d\xi^\alpha \wedge \varpi^j \right) e_i(x)$

This is a first order linear differential operator : $d\varpi^i$ is the ordinary exterior differential on M.

For r=0 we have just $\nabla_e = \widehat{\nabla}$

For r=1 the differential operator reads :

$\widehat{\nabla}_e : J^1(E \otimes TM^*) \to E \otimes \Lambda_2 TM^* ::$

$\widehat{\nabla}_e \left(Z^i_\alpha e_i(x) \otimes dx^\alpha, Z^i_{\alpha,\beta} e^\beta_i(x) \otimes dx^\alpha \right) = \sum_{i\alpha\beta} \left(Z^i_{\alpha,\beta} + \Gamma^i_{\alpha j} Z^j_\beta \right) e_i(x) \otimes dx^\alpha \wedge dx^\beta$

and we have : $\nabla_e(\nabla X) = -\widehat{\Omega}(X)$ which reads : $\widehat{\nabla}_e \circ J^1 \circ \widehat{\nabla}(X) = -\widehat{\Omega}(X)$

$\widehat{\nabla}_e J^1 \left(\nabla_\beta X \otimes dx^\beta \right) = \widehat{\nabla}_e \left((\nabla_\beta X^i) e_i(x) \otimes dx^\beta, \partial_\alpha \nabla_\beta X^i e_i(x) \otimes dx^\alpha \otimes dx^\beta \right)$

$= \sum_{i\alpha\beta} \left(\partial_\alpha \nabla_\beta X^i + \Gamma^i_{\alpha j} (\nabla_\beta X^i) \right) e_i(x) \otimes dx^\alpha \wedge dx^\beta = -\widehat{\Omega}(X)$

3. If we apply two times the exterior covariant derivative we get :

$\nabla_e(\nabla_e \varpi) = \sum_{ij} \left(\sum_{\alpha\beta} R^i_{j\alpha\beta} d\xi^\alpha \wedge d\xi^\beta \right) \wedge \varpi^j \otimes e_i(x)$

Where $R = \sum_{\alpha\beta} \sum_{ij} R^i_{j\alpha\beta} d\xi^\alpha \wedge d\xi^\beta \otimes e^j(x) \otimes e_i(x)$ and $R^i_{j\alpha\beta} = \partial_\alpha \Gamma^i_{j\beta} + \sum_k \Gamma^i_{k\alpha} \Gamma^k_{j\beta}$

$R \in \Lambda_2(M; E^* \otimes E)$ has been called the Riemannian curvature. This is not a differential operator (the derivatives of the section are not involved) but a linear map.

Adjoint

1. The covariant derivative along a vector field W on M is a differential operator on the same vector bundle :

$\widehat{\nabla}_W : \mathfrak{X}(J^1 E) \to \mathfrak{X}(E) ::: \widehat{\nabla}_W Z = \sum_{i\alpha} \left(Z^i_\alpha + \Gamma^i_{\alpha j}(x) Z^j \right) W^\alpha e_i(x)$ with W some fixed vector field in TM.

2. If there is a scalar product g on E and a volume form on M the adjoint map of $\widehat{\nabla}_W$ is defined as seen above (Adjoint).

$[A] = \sum_\alpha [\Gamma_\alpha] W^\alpha$

$[B_\alpha] = W^\alpha [I]$

$\widehat{D}^*(X) = \sum \left(\left([g^{-1}][A]^*[g] + [g^{-1}][B^\alpha]^*[\partial_\alpha g] \right) X + \left([g^{-1}][B^\alpha]^*[g] \right) \partial_\alpha X \right)^i e_i(x)$

$\widehat{D}^*(X) = \sum \left(\left([g^{-1}][\Gamma_\alpha]^*[g] + [g^{-1}][\partial_\alpha g] \right) X + \partial_\alpha X \right)^i \overline{W}^\alpha e_i(x)$

W is real so : $\nabla^* : \mathfrak{X}(E) \to \mathfrak{f}_1(M;E)$

$\nabla^*(X) = \sum \left([g^{-1}] \left(([\Gamma_\alpha]^* + [\partial_\alpha g][g^{-1}]) X + \partial_\alpha X \right) [g] \right)^i e_i(x) \otimes d\xi^\alpha$

is such that :

$\forall W \in \mathfrak{X}(TM), \forall Z \in J^1 E(x), Y \in E(x) : \left\langle \widehat{\nabla}_W Z, Y \right\rangle_{E(x)} = \left\langle Z, \widehat{\nabla}^*_W Y \right\rangle_{J^1 E(x)}$

$\int_M g(x)(\nabla X, Y) \varpi_0 = \int_M g(x)(X, \nabla^* Y) \varpi_0$

3. The symbol of $\widehat{\nabla}$ and $\widehat{\nabla}^*$ are:

$P(x)(u) = \sum_\alpha \Gamma^i_{\alpha j} e_i(x) \otimes dx^\alpha \otimes e^j(x) + u_\alpha e_i(x) \otimes dx^\alpha \otimes e^i(x)$

$P^*(x)(u) = \sum_\alpha \left([g^{-1}][\Gamma_\alpha]^*[g] + [g^{-1}][\partial_\alpha g] \right)^i_j e_i(x) \otimes dx^\alpha \otimes e^j(x) + u_\alpha e_i(x) \otimes dx^\alpha \otimes e^i(x)$

so : $\sigma_\nabla(x, u) = \sigma_{\nabla^*}(x, u) = \sum_\alpha \sum_{i=1}^n u_\alpha e_i(x) \otimes dx^\alpha \otimes e^i(x)$

30.1.8 Dirac operators

Definition of a Dirac operator

Definition 2384 *A first order linear differential operator D on a vector bundle endowed with a scalar product g is said to be a **Dirac operator** if it is weakly elliptic and the principal symbol of $D^* \circ D$ is scalar*

Meaning that $\sigma_{D^* \circ D}(x, u) = \gamma(x, u)\, Id \in \mathcal{L}(E(x); E(x))$
D* is the adjoint of D with respect to g

Properties of a Dirac operator

Theorem 2385 *A Dirac operator D on a vector bundle E with base M induces a riemmanian metric on TM* :*
$G^*(x)(u, v) = \frac{1}{2}(\gamma(x, u + v) - \gamma(x, u) - \gamma(x, v))$

Proof. $\sigma_{D^* \circ D}(x, u) = \sigma_{D^*}(x, u)\, \sigma_D(x, u) = \sigma_D(x, u)^*\, \sigma_D(x, u) = \gamma(x, u)\, Id$

As D is elliptic, $D^* \circ D$ is elliptic and :$\gamma(x, u) \neq 0$ whenever u$\neq 0$

By polarization $G^*(x)(u, v) = \frac{1}{2}(\gamma(x, u + v) - \gamma(x, u) - \gamma(x, v))$ defines a bilinear symmetric form G^* on $\otimes^2 TM^*$ which is definite positive and induces a riemannian metric on TM^* and so a riemannian metric G on TM. ∎

Notice that G is not directly related to g on V (which can have a dimension different from M). E can be a complex vector bundle.

With $D(J^1 Z(x)) = B(x) Z(x) + \sum_{\alpha=1}^{m} A(x)^\alpha Z_\alpha(x)$

$G^*(x)(u, v) I_{m \times m} = \frac{1}{2} \sum_{\alpha, \beta} \left[A(x)^* \right]^\alpha \left[A(x) \right]^\beta (u_\alpha v_\beta + u_\beta v_\alpha) = \sum_{\alpha, \beta} \left[A(x)^* \right]^\alpha \left[A(x) \right]^\beta u_\alpha v_\beta = \left[\sigma_D(x, u) \right]^* \left[\sigma_D(x, v) \right]$

Theorem 2386 *If the Dirac operator D on the vector bundle $E(M, V, \pi)$ endowed with a scalar product g is self adjoint there is an algebra morphism $\Upsilon : Cl(TM^*, G^*) \to \mathcal{L}(E; E) :: \Upsilon(x, u \cdot v) = \sigma_D(x, u) \circ \sigma_D(x, v)$*

Proof. i) Associated to the vector space (TM*(x),G*(x)) there is a Clifford algebra Cl(TM*(x),G*(x))

ii) G*(x) is a Riemmanian metric, so all the Clifford algebras Cl(TM*(x),G*(x)) are isomorphic and we have a Clifford algebra Cl(TM*,G*) over M, which is isomorphic to Cl(TM,G)

iii) $\mathcal{L}(E(x); E(x))$ is a complex algebra with composition law

iv) the map :

$L : T_x M^* \to \mathcal{L}(E(x); E(x)) :: L(u) = \sigma_D(x, u)$ is such that :

$L(u) \circ L(v) + L(v) \circ L(u) = \sigma_D(x, u) \circ \sigma_D(x, v) + \sigma_D(x, v) \circ \sigma_D(x, u) = 2 G^*(x)(u, v) I_{m \times m}$

so, following the universal property of Clifford algebra, there exists a unique algebra morphism :

$\Upsilon_x : Cl(T_x M^*, G^*(x)) \to \mathcal{L}(E(x); E(x))$ such that $L = \Upsilon_x \circ \imath$ where $\imath$ is the canonical map : $\imath : T_x M^* \to Cl(T_x M^*, G^*(x))$

and $Cl\left(T_xM^*, G^*\left(x\right)\right)$ is the Clifford algebra over T_xM^* endowed with the bilinear symmetric form $G^*\left(x\right)$.

v) The Clifford product of vectors translates as : $\Upsilon\left(x, u \cdot v\right) = \sigma_D\left(x, u\right) \circ \sigma_D\left(x, v\right)$ ∎

Dirac operators on a spin bundle

Conversely, it we have a connection on a spin bundle E we can build a differential operator on $\mathfrak{X}\left(E\right)$ which is a Dirac like operator. This procedure is important in physics, and quite clever, so it is useful to detail all the steps.

A reminder of a theorem (see Fiber bundle - connections)

For any representation (V,r) of the Clifford algebra $Cl(\mathbb{R}, r, s)$ and principal bundle Sp(M,$Spin\left(\mathbb{R}, r, s\right)$,$\pi_S$) the associated bundle E=Sp[V,r] is a spin bundle. Any principal connection Ω on Sp with potential $\grave{A}$ induces a linear connection with form $\Gamma = r\left(\sigma\left(\grave{A}\right)\right)$ on E and covariant derivative ∇ . Moreover, the representation $[\mathbb{R}^m,\mathbf{Ad}]$ of $Spin\left(\mathbb{R}, r, s\right)$,$\pi_S$) leads to the associated vector bundle F=Sp[$\mathbb{R}^m$,$\mathbf{Ad}$] and Ω induces a linear connection on F with covariant derivative $\widehat{\nabla}$. There is the relation :
$$\forall X \in \mathfrak{X}\left(F\right), U \in \mathfrak{X}\left(E\right) : \nabla\left(r\left(X\right)U\right) = r\left(\widehat{\nabla}X\right)U + r\left(X\right)\nabla U$$

The ingredients are :

The Lie algebra $o(\mathbb{R}, r, s)$ with a basis $(\overrightarrow{\kappa_\lambda})_{\lambda=1}^q$ and r+s = m

$(\mathbb{R}^m, \eta)$ endowed with the symmetric bilinear form η of signature (r,s) on $\mathbb{R}^m$ and its orthonormal basis $(\varepsilon_\alpha)_{\alpha=1}^m$

σ is the isomorphism :$o(\mathbb{R}, r, s) \to T_1SPin(\mathbb{R}, r, s) :: \sigma\left(\overrightarrow{\kappa}\right) = \sum_{\alpha\beta} [\sigma]_\beta^\alpha \varepsilon_\alpha \cdot \varepsilon_\beta$ with $[\sigma] = \frac{1}{4}[J][\eta]$ where $[J]$ is the mxm matrix of $\overrightarrow{\kappa}$ in the standard representation of $o(\mathbb{R}, r, s)$

The representation (V,r) of $Cl\left(\mathbb{R}, r, s\right)$, with a basis $(e_i)_{i=1}^n$ of V, is defined by the nxn matrices $\gamma_\alpha = r\left(\varepsilon_\alpha\right)$ and : $\gamma_i\gamma_j + \gamma_j\gamma_i = 2\eta_{ij}I, \eta_{ij} = \pm 1$.

The linear connection on E is :
$$\Gamma\left(x\right) = r\left(\sigma\left(\grave{A}_a\right)\right) = \sum_{\alpha\lambda ij} \grave{A}_\alpha^\lambda [\theta_\lambda]_j^i d\xi^\alpha \otimes e_i\left(x\right) \otimes e^j\left(x\right)$$
with $[\theta_\lambda] = \frac{1}{4}\sum_{kl} \left([J_\lambda][\eta]\right)_l^k \left([\gamma_k][\gamma_l]\right)$
and its covariant derivative :
$$U \in \mathfrak{X}\left(E\right) : \nabla U = \sum_{i\alpha}\left(\partial_\alpha U^i + \sum_{\lambda j} \grave{A}_\alpha^\lambda [\theta_\lambda]_j^i U^j\right) d\xi^\alpha \otimes e_i\left(x\right)$$
The connection on Sp induces a linear connection on F :
$$\widehat{\Gamma}\left(x\right) = \left(\mathbf{Ad}\right)'|_{s=1}\left(\sigma\left(\grave{A}\left(x\right)\right)\right) = \sum_\lambda \grave{A}_\alpha^\lambda\left(x\right)[J_\lambda]_j^i \varepsilon_i\left(x\right) \otimes \varepsilon^j\left(x\right)$$
and its covariant derivative :
$$X \in \mathfrak{X}\left(F\right) : \widehat{\nabla}X = \sum_{i\alpha}\left(\partial_\alpha X^i + \sum_{\lambda j} \grave{A}_\alpha^\lambda [J_\lambda]_j^i X^j\right) d\xi^\alpha \otimes \varepsilon_i\left(x\right)$$

Definition 2387 *The Dirac operator is the first order differential operator :*
$\widehat{D} : J^1E \to E :: \widehat{D}Z = \sum_\alpha [\gamma^\alpha]\left(Z_\alpha + [\Gamma_\alpha]Z\right)$

Proof. i) The Clifford algebra of the dual Cl($\mathbb{R}^{m*}, \eta$) is isomorphic to $Cl(\mathbb{R}, r, s)$. (V*,r*) is a representation of Cl($\mathbb{R}^{m*}, \eta$) . It is convenient to take as generators : $r^*\left(\varepsilon^i\right) = \gamma^i = \eta_{ii}\gamma_i \Leftrightarrow [\gamma^i] = -[\gamma_i]^{-1}$

ii) Because **Ad** preserves η the vector bundle F is endowed with a scalar product g for which the holonomic basis $\varepsilon_i(x)$ is orthonormal : $g(x)(\varepsilon_{a\alpha}(x), \varepsilon_{a\beta}(x)) = \eta(\varepsilon_\alpha, \varepsilon_\beta) = \eta_{\alpha\beta}$

iii) F is m dimensional, and the holonomic basis of F can be expressed in components of the holonomic basis of M : $\varepsilon_i(x) = [L(x)]_i^\beta \partial_\beta \xi$.

The scalar product g on F is expressed as :

$g(x)_{\alpha\beta} = g(x)(\partial_\alpha \xi, \partial_\beta \xi) = \sum \left[L(x)^{-1}\right]_\alpha^i \left[L(x)^{-1}\right]_\beta^j g(x)(\varepsilon_i(x), \varepsilon_j(x)) =$

$\sum_{ij} \left[L(x)^{-1}\right]_\alpha^i \left[L(x)^{-1}\right]_\beta^j \eta_{ij}$

and the associated scalar product on TM* is $g^*(x)^{\alpha\beta} = g^*(x)\left(d\xi^\alpha, d\xi^\beta\right) = \sum_{ij} [L(x)]_i^\alpha [L(x)]_j^\beta \eta^{ij}$

iv) So we can define the action :

$R^*(x) : T_x M^* \times E(x) \to E(x) :: R^*(d\xi^\alpha)(e_i(x)) = r^*\left(\sum_i [L(x)]_k^\alpha \varepsilon^k\right)(e_i) = \sum_{kj} [L(x)]_k^\alpha \left[\gamma^k\right]_i^j e_j(x)$

or with $[\gamma^\alpha] = \sum_k [L(x)]_k^\alpha \left[\gamma^k\right] : R^*(d\xi^\alpha)(e_i(x)) = \sum_j [\gamma^\alpha]_i^j e_j(x)$

v) The Dirac operator is defined as :

$DU = \sum_{i\alpha} \left(\nabla_\alpha U^i\right) R^*(d\xi^\alpha)(e_i(x))$

$DU = \sum_{ij\alpha} [\gamma^\alpha]_i^j \left[\nabla_\alpha U^i\right] e_j(x)$

Written as a usual differential operator :

$\widehat{D} : J^1 E \to E :: \widehat{D}Z = \sum_{ij\alpha} [\gamma^\alpha]_j^i \left(Z_\alpha^j + \Gamma_{\alpha k}^j Z^k\right) e_i(x) = \sum_\alpha [\gamma^\alpha]\left(Z_\alpha + [\Gamma_\alpha] Z\right)$

∎

Theorem 2388 *D is a first order weakly elliptic differential operator and the principal symbol of D^2 is scalar*

Proof. i) The symbol of D is :

$P(x, u) = \sum_{\alpha, i, j} ([\gamma^\alpha][\Gamma_\alpha] + [\gamma^\alpha] u_\alpha)_i^j e_j(x) \otimes e^i(x)$

and its principal symbol :

$\sigma_D(x, u) = \sum_\alpha u_\alpha [\gamma^\alpha] = \sum_\alpha u_\alpha R^*(d\xi^\alpha) = R^*(u)$.

As r is an algebra morphism $R^*(u) = 0 \Rightarrow u = 0$. Thus D is weakly elliptic, and DD*,D*D are weakly elliptic.

ii) As we have in the Clifford algebra :

$u, v \in T_x M^* : u \cdot v + v \cdot u = 2g^*(x)(u, v)$

and R^* is an algebra morphism :

$R^*(u \cdot v + v \cdot u) = R^*(u) \circ R^*(v) + R^*(v) \circ R^*(u) = 2g^*(u, v) Id$

$\sigma_D(x, u) \circ \sigma_D(x, v) + \sigma_D(x, v) \circ \sigma_D(x, u) = 2g^*(u, v) Id$

$\sigma_{D \circ D}(x, u) = \sigma_D(x, u) \circ \sigma_D(x, u) = g^*(u, u) Id$ ∎

So D^2 is a scalar operator : it is sometimes said that D is the "square root" of the operator with symbol $g^*(u, v) Id$

Theorem 2389 *The operator D is self adjoint with respect to the scalar product G on E iff :*

$\sum_\alpha ([\gamma^\alpha][\Gamma_\alpha] + [\Gamma_\alpha][\gamma^\alpha]) = \sum_\alpha [\gamma^\alpha]\left[G^{-1}\right][\partial_\alpha G]$

If the scalar product G is induced by a scalar product on V then the condition reads :

$$\sum_\alpha \left([\gamma^\alpha] [\Gamma_\alpha] + [\Gamma_\alpha] [\gamma^\alpha] \right) = 0$$

Proof. i) The volume form on M is $\varpi_0 \sqrt{|\det g|} d\xi^1 \wedge ... \wedge d\xi^m$. If we have a scalar product fiberwise on E(x) with matrix $[G(x)]$ then, applying the method presented above (see Adjoint), the adjoint of D is with $[A] = \sum_\alpha [\gamma^\alpha] [\Gamma_\alpha], [B^\alpha] = [\gamma^\alpha]$:

$$D^* (X) = \sum \left(\left(\left[G^{-1} \right] \sum_\alpha [\Gamma_\alpha]^* [\gamma^\alpha]^* [G] + \left[G^{-1} \right] [\gamma^\alpha]^* [\partial_\alpha G] \right) X \right.$$
$$\left. + \left(\left[G^{-1} \right] [\gamma^\alpha]^* [G] \right) \partial_\alpha X^i e_i (x) \right)$$

The operator D is self adjoint $D = D^*$ iff both :

$$\forall \alpha : [\gamma^\alpha] = \left[G^{-1} \right] [\gamma^\alpha]^* [G]$$
$$\sum_\alpha [\gamma^\alpha] [\Gamma_\alpha] = \sum_\alpha \left[G^{-1} \right] [\Gamma_\alpha]^* [\gamma^\alpha]^* [G] + \left[G^{-1} \right] [\gamma^\alpha]^* [\partial_\alpha G]$$

ii) As $[L]$ is real the first condition reads :

$$\forall \alpha : \sum_j [L(x)]_j^\alpha \left[\gamma^j \right]^* [G(x)] = [G(x)] \sum_j [L(x)]_j^\alpha \left[\gamma^j \right]$$

By multiplication with $\left[L(x)^{-1} \right]_\alpha^i$ and summing we get : $\left[\gamma^i \right]^* = [G(x)] \left[\gamma^i \right] [G(x)]^{-1}$

Moreover :

$$\partial_\alpha \left([G(x)] \left[\gamma^i \right] [G(x)]^{-1} \right) = 0$$
$$= (\partial_\alpha [G(x)]) \left[\gamma^i \right] [G(x)]^{-1} - [G(x)] \left[\gamma^i \right] [G(x)]^{-1} (\partial_\alpha [G(x)]) [G(x)]^{-1}$$
$$[G(x)]^{-1} (\partial_\alpha [G(x)]) \left[\gamma^i \right] = \left[\gamma^i \right] [G(x)]^{-1} (\partial_\alpha [G(x)])$$

iii) With $[\gamma^\alpha]^* = [G] [\gamma^\alpha] \left[G^{-1} \right]$ the second condition gives :

$$\sum_\alpha [\gamma^\alpha] [\Gamma_\alpha] = \sum_\alpha \left[G^{-1} \right] [\Gamma_\alpha]^* [G] [\gamma^\alpha] + [\gamma^\alpha] \left[G^{-1} \right] [\partial_\alpha G]$$
$$= \left(\sum_\alpha \left[G^{-1} \right] [\Gamma_\alpha]^* [G] + \left[G^{-1} \right] [\partial_\alpha G] \right) [\gamma^\alpha]$$

iv) $[\Gamma_\alpha] = \sum_\lambda \grave{A}_\alpha^\lambda [\theta_\lambda]$

$$[\theta_\lambda] = \frac{1}{4} \sum_{kl} ([J_\lambda] [\eta])_l^k [\gamma_k] [\gamma_l] = \frac{1}{4} \sum_{kl} [J_\lambda]_l^k [\gamma_k] \left[\gamma^l \right]$$
$$[\Gamma_\alpha] = \frac{1}{4} \sum_{kl} \sum_\lambda \grave{A}_\alpha^\lambda [J_\lambda]_l^k [\gamma_k] \left[\gamma^l \right] = \frac{1}{4} \sum_{kl} \left[\widehat{\Gamma}_\alpha \right]_l^k [\gamma_k] \left[\gamma^l \right]$$

À and J are real so :

$$[\Gamma_\alpha]^* = \frac{1}{4} \sum_{kl} \left[\widehat{\Gamma}_\alpha \right]_l^k \left[\gamma^l \right]^* [\gamma_k]^* = \frac{1}{4} [G(x)] \left(\sum_{kl} \left[\widehat{\Gamma}_\alpha \right]_l^k \left[\gamma^l \right] [\gamma_k] \right) [G(x)]^{-1}$$

Using the relation $\left[\widehat{\Gamma}_\alpha \right]^t [\eta] + [\eta] \left[\widehat{\Gamma}_\alpha \right] = 0$ (see Metric connections on fiber bundles):

$$\sum_{kl} \left[\widehat{\Gamma}_\alpha \right]_l^k \left[\gamma^l \right] [\gamma_k] = - \sum_{kl} \left[\widehat{\Gamma}_\alpha \right]_k^l \eta_{kk} \eta_{ll} \left[\gamma^l \right] [\gamma_k] = - \sum_{kl} \left[\widehat{\Gamma}_\alpha \right]_k^l [\gamma_l] \left[\gamma^k \right] =$$
$$- \sum_{kl} \left[\widehat{\Gamma}_\alpha \right]_l^k [\gamma_k] \left[\gamma^l \right] = - [\Gamma_\alpha]$$
$$[\Gamma_\alpha]^* = - [G(x)] [\Gamma_\alpha] [G(x)]^{-1}$$

v) The second condition reads:

$$\sum_\alpha [\gamma^\alpha] [\Gamma_\alpha] = \left[G^{-1} \right] \sum_\alpha \left(- [G] [\Gamma_\alpha] [G]^{-1} [G] + [\partial_\alpha G] \right) [\gamma^\alpha]$$
$$= \sum_\alpha \left(- [\Gamma_\alpha] + \left[G^{-1} \right] [\partial_\alpha G] \right) [\gamma^\alpha]$$
$$\sum_\alpha [\gamma^\alpha] [\Gamma_\alpha] + [\Gamma_\alpha] [\gamma^\alpha] = \sum_\alpha \left[G^{-1} \right] [\partial_\alpha G] [\gamma^\alpha] = \sum_\alpha [\gamma^\alpha] \left[G^{-1} \right] [\partial_\alpha G]$$

vi) If G is induced by a scalar product on V then $[\partial_\alpha G] = 0$ ∎

30.2 Laplacian

The laplacian comes in mathematics with many flavors and definitions (we have the connection laplacian, the Hodge laplacian, the Lichnerowicz laplacian, the Bochner laplacian,...). We will follow the more general path, requiring the minimum assumptions for its definition. So we start form differential geometry and the exterior differential (defined on any smooth manifold) and codifferential (defined on any manifold endowed with a metric) acting on forms. From there we define the laplacian Δ as an operator acting on forms over a manifold. As a special case it is an operator acting on functions on a manifold, and furthermore as an operator on functions in $\mathbb{R}^m$.

In this subsection :

M is a pseudo riemannian manifold : real, smooth, m dimensional, endowed with a non degenerate scalar product g (when a definite positive metric is required this will be specified as usual) which induces a volume form ϖ_0. We will consider the vector bundles of r forms complex valued $\Lambda_r(M;\mathbb{C})$ and $\Lambda(M;\mathbb{C}) = \oplus_{r=0}^m \Lambda_r(M;\mathbb{C})$

The metric g can be extended fiberwise to a hermitian map G_r on $\Lambda_r(M;\mathbb{C})$ and to an inner product on the space of sections of the vector bundle : $\lambda, \mu \in \Lambda_r(M;\mathbb{C}) : \langle \lambda, \mu \rangle_r = \int_M G_r(x)(\lambda(x), \mu(x))\varpi_0$ which is well defined for $\lambda, \mu \in L^2(M, \varpi_0, \Lambda_r(M;\mathbb{C}))$.

We denote $\epsilon = (-1)^p$ where p is the number of - in the signature of g. So $\epsilon = 1$ for a riemannian manifold.

Most of the applications are in differential equations and involve initial value conditions, so we will also consider the case where M is a manifold with boundary, embedded in a real, smooth, m dimensional manifold (it is useful to see the precise definition of a manifold with boundary).

30.2.1 Laplacian acting on forms

(see Differential geometry for the principles)

Hodge dual

The Hodge dual $*\lambda_r$ of $\lambda \in L^2(M, \varpi_0, \Lambda_r(M;\mathbb{C}))$ is m-r-form, denoted $*\lambda$, such that :

$\forall \mu \in L^2(M, \varpi_0, \Lambda_r(M;\mathbb{C})) : *\lambda_r \wedge \mu = G_r(\lambda, \mu)\varpi_0$

with $G_r(\lambda, \mu)$ the scalar product of r forms.

This is an anti-isomorphism $* : L^2(M, \varpi_0, \Lambda_r(M;\mathbb{C})) \to L^2(M, \varpi_0, \Lambda_{m-r}(M;\mathbb{C}))$ and

$*^{-1}\lambda_r = \epsilon(-1)^{r(m-r)} * \lambda_r \Leftrightarrow ** \lambda_r = \epsilon(-1)^{r(m-r)}\lambda_r$

Values of the Hodge dual (see Differential geometry).

Codifferential

It is common to define the codifferential as the "formal adjoint" of the exterior differential. We will follow a more rigorous path.

1. The exterior differential d is a first order differential operator on $\Lambda(M;\mathbb{C})$:

d:$\Lambda(M;\mathbb{C}) \to \Lambda_{r+1}(M;\mathbb{C})$

With the jet symbolism :

$\widehat{d} : J^1(\Lambda_r(M;\mathbb{C})) \to \Lambda_{r+1}(M;\mathbb{C}) :: \widehat{d}(Z_r) = \sum_{\{\alpha_1...\alpha_r\}} \sum_\beta Z^\beta_{\alpha_1...\alpha_r} d\xi^\beta \wedge d\xi^{\alpha_1} \wedge ...d\xi^{\alpha_r}$

The symbol of $\widehat{d}$ is :

$P(x,d)(u)(\mu_r) = \left(\sum_\beta u_\beta d\xi^\beta\right) \wedge \left(\sum_{\{\alpha_1...\alpha_r\}} \mu_r d\xi^{\alpha_1} \wedge ...d\xi^{\alpha_r}\right) = u \wedge \mu$

2. The codifferential is the operator :

$\delta : \mathfrak{X}_1(\Lambda_{r+1}(M;\mathbb{C})) \to \mathfrak{X}_0(\Lambda_r(M;\mathbb{C})) ::$

$\delta\lambda_r = \epsilon(-1)^{r(m-r)+r} * d * \lambda_r = (-1)^r * d *^{-1} \lambda_r$

With the jet symbolism :

$\widehat{\delta} : J^1(\Lambda_{r+1}(M;\mathbb{C})) \to \Lambda_r(M;\mathbb{C}) :: \widehat{\delta}Z_r = \epsilon(-1)^{r(m-r)+r} * \widehat{d} \circ J^1(*Z_{r+1})$

The symbol of $\widehat{\delta}$: $P(x,\delta)(u)\mu_{r+1} = -i_{u^*}\mu_{r+1}$ with $u^* = \sum g^{\alpha\beta}\overline{u}_\beta \partial x_\alpha$

Proof. As $\langle d\lambda, \mu\rangle = \langle\lambda, \delta\mu\rangle$ (see below) we have :

$\langle P(x,d)(u)\lambda_r, \mu_{r+1}\rangle_{r+1} = \langle\lambda_r, P(x,\delta)(u)\mu_{r+1}\rangle_r = \langle u \wedge \lambda_r, \mu_{r+1}\rangle_{r+1}$

$\sum_{\{\alpha_1...\alpha_{r+1}\}} (-1)^{k-1} \overline{u}_{\alpha_k} \overline{\lambda}_{\alpha_1...\widehat{\alpha_k}...\alpha_{r+1}} \mu^{\alpha_1...\alpha_{r+1}}$

$= \sum_{\{\alpha_1...\alpha_r\}} \overline{\lambda}_{\alpha_1.....\alpha_r} [P(x,\delta)(u)\mu_{r+1}]^{\alpha_1..\alpha_r}$

$(-1)^{k-1} \overline{u}_{\alpha_k} \mu^{\alpha_1...\alpha_{r+1}} = [P(x,\delta)(u)\mu_{r+1}]^{\alpha_1..\alpha_r}$

$[P(x,\delta)(u)\mu_{r+1}]_{\alpha_1...\alpha_r} = \sum_k (-1)^{k-1} \sum_{\alpha_k} \overline{u}^{\alpha_k} \mu_{\alpha_1...\alpha_{r+1}}$ ∎

As $\sigma_{\widehat{\delta}}(x,u)\lambda_r = \epsilon(-1)^{r(m-r)+r} * \sigma_{\widehat{d}}(x,u)*\lambda_r = \epsilon(-1)^{r(m-r)+r} * (u \wedge (*\lambda_r))$

we have :

$* (u \wedge (*\lambda_r)) = -\epsilon(-1)^{r(m-r)+r} i_{u^*}\lambda_r$

3. Properties :

$d^2 = \delta^2 = 0$

For $f \in C(M;\mathbb{R}) : \delta f = 0$

For $\mu_r \in \Lambda_r TM^*$:

$*\delta\mu_r = (-1)^{m-r-1} d * u_r$

$\delta * \mu_r = (-1)^{r+1} * d\mu_r$

For r=1 : $\delta\left(\sum_i \lambda_\alpha dx^\alpha\right) = (-1)^m \frac{1}{\sqrt{|\det g|}} \sum_{\alpha,\beta=1}^m \partial_\alpha\left(g^{\alpha\beta}\lambda_\beta \sqrt{|\det g|}\right)$

4. The codifferential is the adjoint of the exterior differential :

d, δ are defined on $\Lambda(M;\mathbb{C})$, with value in $\Lambda(M;\mathbb{C})$. The scalar product is well defined if $dX, \delta X \in L^2(M, \varpi_0, \Lambda_r(M;\mathbb{C}))$

So we can consider the scalar products on $W^{1,2}(\Lambda_r(M;\mathbb{C}))$:

$\lambda \in W^{1,2}(\Lambda_r(M;\mathbb{C})), \mu \in W^{1,2}(\Lambda_{r+1}(M;\mathbb{C})) : \langle\delta\mu, \lambda\rangle = \int_M G_r(\delta\mu, \lambda)\varpi_0, \langle\mu, d\lambda\rangle = \int_M G_{r+1}(\mu, d\lambda)\varpi_0$

We have the identity for any manifold M :

$\langle\delta\mu, \lambda\rangle - \langle\mu, d\lambda\rangle = (-1)^{r-m} \int_M d(*\mu \wedge \lambda)$

Theorem 2390 *If M is a manifold with boundary :*$\langle\delta\mu, \lambda\rangle - \langle\mu, d\lambda\rangle = (-1)^{r-m} \int_{\partial M} *\mu \wedge \lambda$

Proof. $d(*\mu \wedge \lambda) = (d*\mu) \wedge \lambda + (-1)^{m-r-1} *\mu \wedge d\lambda = (-1)^{r-m} *\delta\mu \wedge \lambda + (-1)^{m-r-1} *\mu \wedge d\lambda$

with $*\delta\mu = (-1)^{m-r} d * \mu$

$d(*\mu \wedge \lambda) = (-1)^{r-m}(*\delta\mu \wedge \lambda - *\mu \wedge d\lambda)$

$$= (-1)^{r-m} \left(G_r \left(\delta\mu, \lambda \right) - G_{r+1} \left(\mu, d\lambda \right) \right) \varpi_0$$

$$d \left(*\mu \wedge \lambda \right) \in \Lambda_m \left(M; \mathbb{C} \right)$$

$$(-1)^{r-m} \int_M \left(G_r \left(\delta\mu, \lambda \right) - G_{r+1} \left(\mu, d\lambda \right) \right) \varpi_0 = \langle \delta\mu, \lambda \rangle - \langle \mu, d\lambda \rangle = \int_M d \left(*\mu \wedge \lambda \right) = \int_{\partial M} *\mu \wedge \lambda$ ∎

Theorem 2391 *The codifferential is the adjoint of the exterior derivative with respect to the interior product on* $W^{1,2} \left(\Lambda \left(M; \mathbb{C} \right) \right)$

in the following meaning : $\langle d\lambda, \mu \rangle = \langle \lambda, \delta\mu \rangle$

Notice that this identity involves $\widehat{d} \circ J^1 = d, \widehat{\delta} \circ J^1 = \delta$ so we have d*=δ

Proof. Starting from : $\langle \delta\mu, \lambda \rangle - \langle \mu, d\lambda \rangle = \int_M \left(G_r \left(\delta\mu, \lambda \right) - G_{r+1} \left(\mu, d\lambda \right) \right) \varpi_0 = (-1)^{r-m} \int_M d \left(*\mu \wedge \lambda \right)$

With the general assumptions about the manifold M (see Topological considerations) there is a cover of M such that its closure gives an increasing, countable, sequence of compacts. The space $W_c^{1,2} \left(\Lambda_r \left(M; \mathbb{C} \right) \right)$ of r forms on M, continuous with compact support is dense in $W^{1,2} \left(\Lambda_r \left(M; \mathbb{C} \right) \right)$, which is a Banach space. So any form can be estimated by a convergent sequence of compactly supported forms, and for them we can find a manifold with boundary N in M such that :

$\langle \delta\mu, \lambda \rangle - \langle \mu, d\lambda \rangle = \int_N \left(G_r \left(\delta\mu, \lambda \right) - G_{r+1} \left(\mu, d\lambda \right) \right) \varpi_0 = (-1)^{r-m} \int_{\partial N} *\mu \wedge \lambda = 0$

So the relation above extends to $W^{1,2} \left(\Lambda_r \left(M; \mathbb{C} \right) \right).$ ∎

This theorem is true only if M is without boundary but, with our precise definition, a manifold with boundary is not a manifold.

Theorem 2392 *The operator* $d + \delta$ *is self adjoint on* $W^{1,2} \left(\Lambda_r \left(M; \mathbb{C} \right) \right)$

Laplacian

On the pseudo Riemannian manifold M the Laplace-de Rahm (also called Hodge laplacian) operator is :

$\Delta : \mathfrak{X}_2 \left(\Lambda_r \left(M; \mathbb{C} \right) \right) \to \mathfrak{X}_0 \left(\Lambda_r \left(M; \mathbb{C} \right) \right) :: \Delta = - \left(\delta d + d\delta \right) = - \left(d + \delta \right)^2$

Remark : one finds also the definition $\Delta = \left(\delta d + d\delta \right).$

We have :

$\langle \Delta\lambda, \mu \rangle = - \langle \left(\delta d + d\delta \right) \lambda, \mu \rangle = - \langle \delta d\lambda, \mu \rangle - \langle \left(d\delta \right) \lambda, \mu \rangle = - \langle d\lambda, d\mu \rangle - \langle \delta\lambda, \delta\mu \rangle$

Theorem 2393 *(Taylor 1 p.163) The principal symbol of* Δ *is scalar :* $\sigma_\Delta \left(x, u \right) = - \left(\sum_{\alpha\beta} g^{\alpha\beta} u_\alpha u_\beta \right) Id$

It follows that the laplacian (or $-\Delta$) is an elliptic operator iff the metric g is definite positive.

Theorem 2394 *The laplacian* Δ *is a self adjoint operator with respect to* G_r *on* $W^{2,2} \left(\Lambda_r \left(M; \mathbb{C} \right) \right)$:

$\forall \lambda, \mu \in \Lambda_r \left(M; \mathbb{C} \right) : \langle \Delta\lambda, \mu \rangle = \langle \lambda, \Delta\mu \rangle$

Proof. $\langle \Delta\lambda, \mu \rangle = - \langle \left(\delta d + d\delta \right) \lambda, \mu \rangle = - \langle \delta d\lambda, \mu \rangle - \langle \left(d\delta \right) \lambda, \mu \rangle = - \langle d\lambda, d\mu \rangle - \langle \delta\lambda, \delta\mu \rangle = - \langle \lambda, \delta d\mu \rangle - \langle \lambda, d\delta\mu \rangle = \langle \lambda, \Delta\mu \rangle$ ∎

Theorem 2395 *If the metric g sur M is definite positive then the laplacian on $W^{2,2}(\Lambda_r(M;\mathbb{C}))$ is such that :*

i) its spectrum is a locally compact subset of $\mathbb{R}$

ii) its eigen values are real, and constitute either a finite set or a sequence converging to 0

iii) it is a closed operator : if $\mu_n \to \mu$ in $W^{1,2}(\Lambda_r(M;\mathbb{C}))$ then $\Delta\mu_n \to \Delta\mu$

Proof. If $X \in W^{2,2}(\Lambda_r(M;\mathbb{C})) \Leftrightarrow X \in \mathfrak{X}_2(\Lambda_r(M;\mathbb{C})), J^2X \in L^2(M,\varpi_0,\Lambda_r(M;\mathbb{C}))$ then $\delta dX, d\delta X \in L^2(M,\varpi_0,\Lambda_r(M;\mathbb{C}))$ and $\Delta X \in L^2(M,\varpi_0,\Lambda_r(M;\mathbb{C}))$. So the laplacian is a map : $\Delta : W^{2,2}(\Lambda_r(M;\mathbb{C})) \to L^2(M,\varpi_0,\Lambda_r(M;\mathbb{C}))$

If the metric g is definite positive, then $W^{2,2}(\Lambda_r(M;\mathbb{C})) \subset L^2(M,\varpi_0,\Lambda_r(M;\mathbb{C}))$ are Hilbert spaces and as Δ is self adjoint the properties are a consequence of general theorems on operators on C*-algebras. $\blacksquare$

Harmonic forms

Definition 2396 *On the pseudo Riemannian manifold M a r-form is said to be **harmonic** if $\Delta\mu = 0$*

Then for $\mu \in \Lambda_r(M;\mathbb{C})$: $\Delta\mu = 0 \Leftrightarrow d\delta\mu = \delta d\mu = 0,$

Theorem 2397 *(Taylor 1 p.354) On a Riemannian compact manifold M the set of harmonic r-form is finite dimensional, isomorphic to the space of order r cohomology $H^r(M)$*

The space $H^r(M)$ is defined in the Differential Geometry part (see cohomology). It is independant of the metric g, so the set of harmonic forms is the same whatever the (riemannian) metric.

Theorem 2398 *(Taylor 1 p.354) On a Riemannian compact manifold M $\forall\mu \in W^{k,2}(\Lambda_r(M;\mathbb{C}))$: $\mu = d\delta G\mu + \delta dG\mu + P_r\mu$ where $G : W^{k,2}(\Lambda_r(M;\mathbb{C})) \to W^{k+2,2}(\Lambda_r(M;\mathbb{C}))$ and P_r is the orthogonal projection of $L^2(M,\Lambda_r(M;\mathbb{C}),\varpi_0)$ onto the space of r harmonic forms. The three terms are mutually orthogonal in $L^2(M,\Lambda_r(M;\mathbb{C}),\varpi_0)$*

This is called the Hodge decomposition. There are many results on this subject, mainly when M is a manifold with boundary in $\mathbb{R}^m$ (see Axelsson).

Inverse of the laplacian

We have a stronger result if the metric is riemannian and M compact :

Theorem 2399 *(Taylor 1 p.353) On a smooth Riemannian compact manifold M the operator :*

$\Delta : W^{1,2}(\Lambda_r(M;\mathbb{C})) \to W^{-1,2}(\Lambda_r(M;\mathbb{C}))$

is such that : $\exists C_0, C_1 \geq 0 : -\langle\mu,\Delta\mu\rangle \geq C_0\|\mu\|^2_{W^{1,2}} - C_1\|\mu\|^2_{W^{-1,2}}$

The map : $-\Delta + C_1 : H^1(\Lambda_r(M;\mathbb{C})) \to H^{-1}(\Lambda_r(M;\mathbb{C}))$ is bijective and its inverse is a self adjoint compact operator on $L^2(M,\Lambda_r(M;\mathbb{C}),\varpi_0)$

The space $H^{-1}(\Lambda_r(M;\mathbb{C})) = W^{-1,2}(\Lambda_r(M;\mathbb{C}))$ is defined as the dual of $\overline{\mathfrak{X}_{\infty c}(\Lambda_r TM^* \otimes \mathbb{C})}$ in $W^{1,2}(\Lambda_r(M;\mathbb{C}))$

30.2.2 Scalar Laplacian

When acting on complex valued functions on a pseudo riemannian manifold M (as above) the laplacian has specific properties.

Coordinates expressions

Theorem 2400 *On a smooth m dimensional real pseudo riemannian manifold (M,g) :*
For $f \in C_2(M;\mathbb{C}) : \Delta f = (-1)^{m+1} div\,(gradf)$

Theorem 2401 $\Delta f = (-1)^{m+1} \sum_{\alpha,\beta=1}^{m} g^{\alpha\beta}\partial_{\alpha\beta}^2 f + (\partial_\beta f)\left(\partial_\alpha g^{\alpha\beta} - \frac{1}{2}\sum_{\lambda\mu} g_{\lambda\mu}\partial_\alpha g^{\mu\lambda}\right)$

So that if $g^{\alpha\beta} = \eta^{\alpha\beta} = Cte : \Delta f = \sum_{\alpha,\beta=1}^{m} \eta^{\alpha\beta}\partial_{\alpha\beta}^2 f$ and in euclidean space we have the usual formula : $\Delta f = (-1)^{m+1}\sum_{\alpha=1}^{m} \frac{\partial^2 f}{\partial x_\alpha^2}$

The principal symbol of Δ is $\sigma_\Delta(x,u) = (-1)^{m+1}\sum_{\alpha,\beta=1}^{m} g^{\alpha\beta} u_\alpha u_\beta$

Proof. $\delta f = 0 \Rightarrow \Delta f = -\delta df$

$\Delta f = -\delta\left(\sum_\alpha \partial_\alpha f dx^\alpha\right) = -(-1)^m \frac{1}{\sqrt{|\det g|}}\sum_{\alpha,\beta=1}^{m} \partial_\alpha\left(g^{\alpha\beta}\partial_\beta f \sqrt{|\det g|}\right)$

$\Delta f = (-1)^{m+1}\sum_{\alpha,\beta=1}^{m} g^{\alpha\beta}\partial_{\alpha\beta}^2 f + (\partial_\beta f)\left(\partial_\alpha g^{\alpha\beta} + g^{\alpha\beta}\frac{\partial_\alpha\sqrt{|\det g|}}{\sqrt{|\det g|}}\right)$

$\frac{\partial_\alpha\sqrt{\epsilon\det g}}{\sqrt{\epsilon\det g}} = \frac{1}{2}\frac{\epsilon\partial_\alpha\det g}{\epsilon\det g} = \frac{1}{2}\frac{1}{\det g}(\det g)Tr\left([\partial_\alpha g][g^{-1}]\right) = \frac{1}{2}\sum_{\lambda\mu} g^{\mu\lambda}\partial_\alpha g_{\lambda\mu} = -\frac{1}{2}\sum_{\lambda\mu} g_{\lambda\mu}\partial_\alpha g^{\mu\lambda}$ ∎

The last term can be expressed with the Lévy Civita connection :

Theorem 2402 $\Delta f = (-1)^{m+1}\sum_{\alpha,\beta=1}^{m} g^{\alpha\beta}\left(\partial_{\alpha\beta}^2 f - \sum_\gamma \Gamma_{\alpha\beta}^\gamma \partial_\gamma f\right)$

Proof. $\sum_\gamma \Gamma_{\gamma\alpha}^\gamma = \frac{1}{2}\frac{\partial_\alpha|\det g|}{|\det g|} = \frac{\partial_\alpha\left(\sqrt{|\det g|}\right)}{\sqrt{|\det g|}}$

$\sum_\alpha \partial_\alpha g^{\alpha\beta} + g^{\alpha\beta}\sum_\gamma \Gamma_{\gamma\alpha}^\gamma = \sum_{\alpha\gamma} -g^{\beta\alpha}\Gamma_{\gamma\alpha}^\gamma - g^{\alpha\gamma}\Gamma_{\alpha\gamma}^\beta + g^{\alpha\beta}\Gamma_{\gamma\alpha}^\gamma = -\sum_{\alpha\gamma} g^{\alpha\gamma}\Gamma_{\alpha\gamma}^\beta$

$\Delta f = (-1)^{m+1}\sum_{\alpha,\beta=1}^{m} g^{\alpha\beta}\partial_{\alpha\beta}^2 f - (\partial_\beta f)\sum_\gamma\left(g^{\alpha\gamma}\Gamma_{\alpha\gamma}^\beta\right)$

$\Delta f = (-1)^{m+1}\sum_{\alpha,\beta=1}^{m} g^{\alpha\beta}\left(\partial_{\alpha\beta}^2 f - \sum_\gamma \Gamma_{\alpha\beta}^\gamma \partial_\gamma f\right)$ ∎

If we write : $p = gradf$ then $\Delta f = (-1)^{m+1} divp$

Proof. $p = gradf \Leftrightarrow p^\alpha = g^{\alpha\beta}\partial_\beta f \Leftrightarrow \partial_\alpha f = g_{\alpha\beta}p^\beta$

$\Delta f = (-1)^{m+1}\sum_{\alpha,\beta\gamma=1}^{m} g^{\alpha\beta}\left(p^\gamma\partial_\beta g_{\alpha\gamma} + g_{\alpha\gamma}\partial_\beta p^\gamma - \Gamma_{\alpha\beta}^\gamma g_{\gamma\eta}p^\eta\right)$

$\Delta f = (-1)^{m+1}\sum_{\alpha,\beta\gamma=1}^{m}\left(p^\gamma g^{\alpha\beta}\left(\sum_\eta g_{\gamma\eta}\Gamma_{\alpha\beta}^\eta + g_{\alpha\eta}\Gamma_{\beta\gamma}^\eta\right) + \partial_\beta p^\beta - \Gamma_{\alpha\beta}^\gamma g^{\alpha\beta}g_{\gamma\eta}p^\eta\right)$

$\Delta f = (-1)^{m+1}\sum_{\alpha,\beta\gamma\eta=1}^{m}\left(g^{\alpha\beta}\left(p^\gamma g_{\gamma\eta}\Gamma_{\alpha\beta}^\eta - \Gamma_{\alpha\beta}^\eta g_{\gamma\eta}p^\gamma\right) + p^\gamma\Gamma_{\beta\gamma}^\beta + \partial_\beta p^\beta\right)$

$\Delta f = (-1)^{m+1}\sum_{\alpha=1}^{m}\left(\partial_\alpha p^\alpha + \sum_\beta p^\alpha\Gamma_{\beta\alpha}^\beta\right) = \sum_{\alpha=1}^{m}\nabla_\alpha p^\alpha$ ∎

Warning ! When dealing with the scalar laplacian, meaning acting on functions over a manifold, usually one drops the constant $(-1)^{m+1}$. So the last expression gives the alternate definition : $\Delta f = div\,(grad\,(f))$. We will follow this convention.

The riemannian Laplacian in $\mathbb{R}^m$ with spherical coordinates has the following expression:

$\Delta = \left(\frac{\partial^2}{\partial r^2} + \frac{m-1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \Delta_{S^{n-1}} \right)$

Wave operator

As can be seen from the principal symbol Δ is elliptic iff the metric g on M is riemannian, which is the usual case. When the metric has a signature (p + ,q -), as for Lorentz manifolds, we have a **d'Alambertian** denoted $\Box$. If q=1 usually there is a folliation of M in space like hypersurfaces S_t, p dimensional manifolds endowed with a riemannian metric, over which one considers a purely riemannian laplacian Δ_x . So $\Box$ is split in Δ_x and a "time component" which can be treated as $\frac{\partial^2}{\partial t^2}$, and the functions are then $\varphi(t,x) \in C(\mathbb{R}; C(S_t; \mathbb{C}))$ .This is the "wave operator" which is seen in the PDE sections.

Domain of the laplacian

As many theorems about the laplacian use distributions it is necessary to understand how we get from on side of the question (functions) to the other (distributions).

1. The scalar riemannian laplacian acts on functions. If $f \in C_r(M; \mathbb{C})$ then Δf is still a function if $r \geq 2$. Thus Δf is in L^2 if f$\in W^{r,2}(M) = H^r(M)$ with r > 1.

2. The associated operator acting on distributions $\mu \in C_{\infty c}(M; \mathbb{C})'$ is $\Delta' \mu(\varphi) = \mu(\Delta \varphi)$. It has same coefficients as Δ

$\Delta' \mu = \sum_{\alpha,\beta=1}^{m} g^{\alpha\beta} \partial_{\alpha\beta}^2 \mu + \left(\partial_\alpha g^{\alpha\beta} - \frac{1}{2} \sum_{\lambda\mu} g_{\lambda\mu} \partial_\alpha g^{\mu\lambda} \right) (\partial_\beta \mu)$

3. Distributions acting on a space of functions can be defined through m forms, with all the useful properties (notably the derivative of the distribution is the derivative of the function). As noticed above the operator D' on distributions can be seen as D acting on the unique component of the m-form : $D' J^r T \left(\left(\lambda_0 d\xi^1 \wedge ... \wedge d\xi^m \right) \right) = T \left((D\lambda_0) d\xi^1 \wedge ... \wedge d\xi^m \right)$ so it is usual to "forget" the $d\xi^1 \wedge ... \wedge d\xi^m$ part.

If $\lambda_0 \in L^p_{loc}(O, dx, \mathbb{C}), 1 \leq p \leq \infty$ then $T(\lambda_0) \in C_{\infty c}(O; \mathbb{C})'$ so we can consider $\Delta' T(\lambda_0) = T(\lambda_0) \Delta$.

The Sobolev spaces have been extended to distributions : $H^{-r}(M)$ is a vector subspace of $C_{\infty c}(M; \mathbb{C})'$ identified to the distributions : $H^{-r}(M) = \left\{ \mu \in C_{\infty c}(M; \mathbb{C})'; \mu = \sum_{\|\alpha\| \leq r} D_\alpha T(f_\alpha), f_\alpha \in L^2(M, \varpi_0, \mathbb{C}) \right\}$. The spaces $H^k(M), k \in \mathbb{Z}$ are Hilbert spaces.

If $f \in H^r(M)$ then $\Delta' T(f) \in H^{r-2}(M)$. It is a distribution induced by a function if r > 1 and then $\Delta' T(f) = T(\Delta f)$

4. Δ is defined in $H^2(M) \subset L^2(M, \varpi_0, \mathbb{C})$, which are Hilbert spaces. It has an adjoint Δ^* on $L^2(M, \varpi_0, \mathbb{C})$ defined as :

$\Delta^* \in L\left(D(\Delta^*); L^2(M, \varpi_0, \mathbb{C}) \right) :: \forall u \in H^2(M), v \in D(\Delta^*) : \langle \Delta u, v \rangle = \langle u, \Delta^* v \rangle$ and $H^2(M) \subset D(\Delta^*)$ so Δ is symmetric. We can look for extending the domain beyond $H^2(M)$ in $L^2(M, \varpi_0, \mathbb{C})$ (see Hilbert spaces). If the exten-

sion is self adjoint and unique then Δ is said to be essentially self adjoint. We have the following :

Theorem 2403 *(Gregor'yan p.6) On a riemannian connected smooth manifold M Δ has a unique self adjoint extension in $L^2(M,\varpi_0,\mathbb{C})$ to the domain $\left\{ f \in \overline{C_{\infty c}(M;\mathbb{C})} : \Delta'T(f) \in T(L^2(M,\varpi_0,\mathbb{C})) \right\}$ where the closure is taken in $L^2(M,\varpi_0,\mathbb{C})$. If M is geodesically complete then Δ is essentially self adjoint on $C_{\infty c}(M;\mathbb{C})$*

Theorem 2404 *(Taylor 2 p.82-84) On a riemannian compact manifold with boundary M Δ has a self adjoint extension in $L^2\left(\overset{\circ}{M},\mathbb{C},\varpi_0\right)$ to the domain*

$\left\{ f \in H_c^1\left(\overset{\circ}{M}\right) : \Delta'T(f) \in T\left(L^2\left(\overset{\circ}{M},\mathbb{C},\varpi_0\right)\right) \right\}$. *If M is smooth Δ is essentially self adjoint on the domains*

  $f \in C_\infty(M;\mathbb{C}) : f = 0$ *on ∂M and on* $f \in C_\infty(M;\mathbb{C}) : \frac{\partial f}{\partial n} = 0$ *on ∂M (n is the normal to the boundary)*

Theorem 2405 *(Zuily p.165) On a compact manifold with smooth boundary M in $\mathbb{R}^m$:*
$$\left\{ f \in H_c^1\left(\overset{\circ}{M}\right) : \Delta'T(f) \in T\left(L^2\left(\overset{\circ}{M},\mathbb{C},\varpi_0\right)\right) \right\} \equiv H_c^1\left(\overset{\circ}{M}\right) \cap H^2\left(\overset{\circ}{M}\right)$$

Green's identity

Theorem 2406 *On a pseudo riemannian manifold : for $f,g \in W^{2,2}(M)$:$\langle df, dg\rangle = -\langle f, \Delta g\rangle = -\langle \Delta f, g\rangle$*

Proof. Indeed : $\underline{\langle df, dg\rangle = \langle f, \delta dg\rangle}$ and $\delta g = 0 \Rightarrow \Delta g = -\delta dg$
  $\langle df, dg\rangle = \overline{\langle dg, df\rangle} = -\overline{\langle g, \Delta f\rangle} = -\langle \Delta f, g\rangle$ ∎
  As a special case : $\langle df, df\rangle = -\langle f, \Delta f\rangle = \|df\|^2$
  As a consequence :

Theorem 2407 *On a pseudo riemannian manifold with boundary M : for $f,g \in W^{2,1}(M)$: $\langle \Delta f, g\rangle - \langle f, \Delta g\rangle = \int_{\partial M}\left(\frac{\partial f}{\partial n}g - f\frac{\partial g}{\partial n}\right)\varpi_1$*

  where $\frac{\partial f}{\partial n} = f'(p)n$ and n is a unitary normal and ϖ_1 the volume form on ∂M induced by ϖ_0

Spectrum of the laplacian
  The spectrum of Δ is the set of complex numbers λ such that $\Delta - \lambda I$ has no bounded inverse. So it depends on the space of functions on which the laplacian is considered : regularity of the functions, their domain and on other conditions which can be imposed, such as "the Dirichlet condition". The eigen values are isolated points in the spectrum, so if the spectrum is discrete it coincides

with the eigenvalues. The eigen value 0 is a special case : the functions such that $\Delta f = 0$ are harmonic and cannot be bounded (see below). So if they are compactly supported they must be null and thus cannot be eigenvectors and 0 is not an eigenvalue.

One key feature of the laplacian is that the spectrum is different if the domain is compact or not. In particular the laplacian has eigen values iff the domain is relatively compact. The eigen functions are an essential tool in many PDE.

Theorem 2408 *(Gregor'yan p.7) In any non empty relatively compact open subset O of a riemannian smooth manifold M the spectrum of $-\Delta$ on*
$\left\{ f \in \overline{C_{\infty c}\left(O;\mathbb{C}\right)} : \Delta'T\left(f\right) \in T\left(L^2\left(O,\varpi_0,\mathbb{C}\right)\right)\right\}$ *is discrete and consists of an increasing sequence $(\lambda_n)_{n=1}^{\infty}$ with $\lambda_n \geq 0$ and $\lambda_n \to_{n \to \infty} \infty$ .If $M\backslash\overline{O}$ is non empty then $\lambda_1 > 0$.*

If the eigenvalues are counted with their multiplicity we have the Weyl's formula: $\lambda_n \sim C_m \left(\frac{n}{Vol(O)}\right)^{2/m}$ with : $C_m = (2\pi)^2 \left(\frac{m\Gamma\left(\frac{m}{2}\right)}{2\pi^{m/2}}\right)^{m/2}$ and $\text{Vol(O)} = \int_O \varpi_0$

Theorem 2409 *(Taylor 1 p.304-316) On a riemannian compact manifold with boundary M :*

i) The spectrum of $-\Delta$ on $\left\{ f \in H_c^1\left(\overset{\circ}{M}\right) : \Delta'T\left(f\right) \in T\left(L^2\left(\overset{\circ}{M},\varpi_0,\mathbb{C}\right)\right)\right\}$ is discrete and consists of an increasing sequence $(\lambda_n)_{n=1}^{\infty}$ with $\lambda_n \geq 0$ and $\lambda_n \to_{n\to\infty} \infty$ .If the boundary is smooth, then $\lambda_1 > 0$

ii) The eigenvectors e_n of $-\Delta$ belong to $C_\infty\left(M;\mathbb{C}\right)$ and constitute a countable Hilbertian basis of $L^2\left(\overset{\circ}{M},\mathbb{C},\varpi_0\right)$. If the boundary is smooth, then $e_1 \in H_c^1\left(\overset{\circ}{M}\right)$, e_1 is nowhere vanishing in the interior of M, and if e_1=0 on ∂M then the corresponding eigenspace is unidimensional.

Notice that by definition M is closed and includes the boundary.

Theorem 2410 *(Zuily p.164) On a bounded open subset O of $\mathbb{R}^m$ the eigenvalues of $-\Delta$ are an increasing sequence $(\lambda_n)_{n=1}^{\infty}$ with $\lambda_n > 0$ and $\lambda_n \to_{n\to\infty} \infty$. The eigenvectors $(e_n)_{n=1}^{\infty} \in H_c^1\left(O\right), \|e_n\|_{H^1(O)} = \lambda_n$ and can be chosen to constitute an orthonomal basis of $L^2\left(O,dx,\mathbb{C}\right)$*

When O is a sphere centered in 0 then the e_n are the spherical harmonics : polynomial functions which are harmonic on the sphere (see Representation theory).

If the domain is an open non bounded of $\mathbb{R}^m$ the spectrum of $-\Delta$ is $[0,\infty]$ and the laplacian has no eigen value.

Fundamental solution for the laplacian

On a riemannian manifold a fundamental solution of the operator $-\Delta$ is given by a Green's function through the heat kernel function p(t,x,y) which is itself given through the eigenvectors of $-\Delta$ (see Heat kernel)

Theorem 2411 *(Gregor'yan p.45) On a riemannian manifold (M,g) if the Green's function $G(x,y) = \int_0^\infty p(t,x,y)\,dt$ is such that $\forall x \neq y \in M : G(x,y) < \infty$ then a fundamental solution of $-\Delta$ is given by $T(G) : -\Delta_x TG(x,y) = \delta_y$*

It has the following properties :

i) $\forall x,y \in M : G(x,y) \geq 0$,

ii) $\forall x : G(x,.) \in L^1_{loc}(M,\varpi_0,\mathbb{R})$ is harmonic and smooth for $x \neq y$

iii) it is the minimal non negative fundamental solution of $-\Delta$ on M.

The condition $\forall x \neq y \in M : G(x,y) < \infty$ is met (one says that G is finite) if :

i) If M is a non empty relatively compact open subset of a riemannian manifold N such that $N \setminus \overline{M}$ is non empty

ii) or if the smallest eigen value $\lambda_{\min}$ of $-\Delta$ is > 0

On $\mathbb{R}^m$ the fundamental solution of $\Delta U = \delta_y$ is : $U(x) = T_y(G(x,y))$ where G(x,y) is the function :

$m \geq 3 : G(x,y) = \frac{1}{(2-m)A(S_{m-1})} \|x - y\|^{2-m}$ where $A(S_{m-1}) = \frac{2\pi^{m/2}}{\Gamma\left(\frac{m}{2}\right)}$ is the Lebesgue surface of the unit sphere in $\mathbb{R}^m$.

$m = 3 : G(x,y) = -\frac{1}{4\pi}\frac{1}{\|x-y\|}$

$m = 2 : G(x,y) = -\frac{1}{2\pi}\ln\|x - y\|$

Inverse of the laplacian   The inverse of the laplacian exists if the domain is bounded. It is given by the Green's function.

Theorem 2412 *(Taylor 1 p.304-316) On a riemannian compact manifold with boundary M :*

$\Delta : H^1_c\left(\overset{\circ}{M}\right) \to H^{-1}\left(\overset{\circ}{M}\right)$ *is a bijective map*

The inverse of Δ is then a compact, self adjoint differential operator on $L^2\left(\overset{\circ}{M},\mathbb{C},\varpi_0\right)$

Theorem 2413 *(Gregor'yan p.45) On a riemannian manifold (M,g), if the smallest eigen value λ_1 of $-\Delta$ on M is > 0 then the operator : $(Gf)(x) = \int_M G(x,y) f(y)\varpi_0(y)$ is the inverse operator of $-\Delta$ in $L^2(M,\varpi_0,\mathbb{R})$*

The condition is met if O is a non empty relatively compact open subset of a riemannian manifold (M,g) such that $M \setminus \overline{O}$ is non empty. Then $\forall f \in L^2(O,\varpi_0,\mathbb{R}), \varphi(x) = \int_O G(x,y) f(y)\varpi_0(y)$ is the unique solution of $-\Delta\varphi = f$

Theorem 2414 *(Zuily p.163) On a bounded open subset O of $\mathbb{R}^m$:*

$-\Delta : \left\{f \in H^1_c(O) : \Delta' T(f) \in T\left(L^2(O,\varpi_0,\mathbb{C})\right)\right\} \to T\left(L^2(O,\varpi_0,\mathbb{C})\right)$ *is an isomorphism*

Its inverse is an operator compact, positive and self adjoint. Its spectrum is comprised of 0 and a sequence of positive numbers which are eigen values and converges to 0.

Maximum principle

Theorem 2415 *(Taylor 1 p.309) , For a first order linear real differential operator D with smooth coefficients on $C_1\left(M;\mathbb{R}\right)$, with M a connected riemannian compact manifold with boundary :*

i) $-\Delta + D : H_c^1\left(\overset{\circ}{M}\right) \to H^{-1}\left(\overset{\circ}{M}\right)$ is a Freholm operator of index zero, thus it is surjective iff it is injective.

ii) If ∂M is smooth, for $f \in C_1\left(\overline{M};\mathbb{R}\right) \cap C_2\left(\overset{\circ}{M};\mathbb{R}\right)$ such that $(D+\Delta)(f) \geq 0$ on $\overset{\circ}{M}$, and $y \in \partial M : \forall x \in \overset{\circ}{M} : f(y) \geq f(x)$ then $f'(y)n > 0$ where n is an outward pointing normal to ∂M.

iii) If ∂M is smooth, for $f \in C_0\left(\overline{M};\mathbb{R}\right) \cap C_2\left(\overset{\circ}{M};\mathbb{R}\right)$ such that $(D+\Delta)(f) \geq 0$ on $\overset{\circ}{M}$, then either f is constant or $\forall x \in \overset{\circ}{M} : f(x) < \sup_{y \in \partial M} f(y)$

Theorem 2416 *(Taylor 1 p.312) For a first order scalar linear differential operator D on $C_1(O;\mathbb{R})$ with smooth coefficients, with O an open, bounded open subset of $\mathbb{R}^m$ with boundary $\partial O = \overset{.}{O}$:*

If $f \in C_0\left(\overline{O};\mathbb{R}\right) \cap C_2(O;\mathbb{R})$ and $(D+\Delta)(f) \geq 0$ on O then $\sup_{x \in O} f(x) = \sup_{y \in \partial O} f(y)$

Furthermore if $(D+\Delta)(f) = 0$ on ∂O then $\sup_{x \in O} |f(x)| = \sup_{y \in \partial M} |f(y)|$

Harmonic functions in $\mathbb{R}^m$

Definition 2417 *A function $f \in C_2(O;\mathbb{C})$ where O is an open subset of $\mathbb{R}^m$ is said to be **harmonic** if : $\Delta f = \sum_{j=1}^m \frac{\partial^2 f}{(\partial x^i)^2} = 0$*

Theorem 2418 *(Lieb p.258) If $S \in C_{\infty c}(O;\mathbb{R})$, where O is an open subset of $\mathbb{R}^m$, is such that $\Delta S = 0$, there is a harmonic function f: $S = T(f)$*

Harmonic functions have very special properties : they are smooth, defined uniquely by their value on a hypersurface, have no extremum except on the boundary.

Theorem 2419 *A harmonic function is indefinitely R-differentiable*

Theorem 2420 *(Taylor 1 p.210) A harmonic function $f \in C_2(\mathbb{R}^m;\mathbb{C})$ which is bounded is constant*

Theorem 2421 *(Taylor 1 p.189) On a smooth manifold with boundary M in* $\mathbb{R}^m$, *if* $u, v \in C_\infty\left(\overline{M}; \mathbb{R}\right)$ *are such that* $\Delta u = \Delta v = 0$ *in* $\overset{\circ}{M}$ *and u=v=f on* ∂M *then u=v on all of M.*

The result still stands if M is bounded and u,v continuous on ∂M.

Theorem 2422 *A harmonic function* $f \in C_2(O; \mathbb{R})$ *on an open connected subset O of* $\mathbb{R}^m$ *has no interior minimum of maximum unless it is constant. In particular if O is bounded and f continuous on the border* ∂O *of O then* $\sup_{x \in O} f(x) = \sup_{y \in \partial O} f(y)$ *. If f is complex valued :* $\sup_{x \in O} |f(x)| = \sup_{y \in \partial O} |f(y)|$

The value of harmonic functions equals their averadge on balls :

Theorem 2423 *(Taylor 1 p.190) Let B(0,r) the open ball in* $\mathbb{R}^m$, $f \in C_2(B(0,r); \mathbb{R}) \cap C_0(B(0,r); \mathbb{R})$, $\Delta f = 0$ *then* $f(0) = \frac{1}{A(B(0,r))} \int_{\partial B(0,r)} f(x)\, dx$

Harmonic radial functions : In $\mathbb{R}^m$ a harmonic function which depends only of $r = \sqrt{x_1^2 + .. + x_m^2}$ must satisfy the ODE :

$f(x) = g(r) : \Delta f = \frac{d^2 g}{dr^2} + \frac{n-1}{r}\frac{dg}{dr} = 0$
and the solutions are : $m \neq 2 : g = Ar^{2-m} + B$; $m = 2 : g = -A \ln r + B$
Because of the singularity in 0 the laplacian is :
$m \neq 2 : \Delta' T(f) = T(\Delta f) - (m-2)2\pi^{m/2}/\Gamma(m/2)\delta_0$; $m = 2 : \Delta' T(f) = T(\Delta f) - 2\pi\delta_0$

Harmonic functions in $\mathbb{R}^2$

Theorem 2424 *A harmonic function in an open O of* $\mathbb{R}^2$ *is smooth and real analytic.*

Theorem 2425 *If* $f : \Omega \to \mathbb{C}$ *is holomorphic on an open* Ω *in* $\mathbb{C}$, *then the function* $\widehat{f}(x,y) = f(x+iy)$ *is harmonic :* $\frac{\partial^2 \widehat{f}}{\partial x^2} + \frac{\partial^2 \widehat{f}}{\partial y^2} = 0$. *Its real and imaginary components are also harmonic.*

This is the direct consequence of the Cauchy relations. Notice that this result does not stand for $O \subset \mathbb{R}^{2n}, n > 1$

Theorem 2426 *(Schwartz III p.279) A harmonic function P*$\in C_2(O; \mathbb{R})$ *on a simply connected open O in* $\mathbb{R}^2$ *:*
i) is in an infinitely many ways the real part of an holomorphic function : $f \in H\left(\widehat{O}; \mathbb{C}\right)$ *where* $\widehat{O} = \{x + iy, x, y \in \Omega\}$ *. and f is defined up to an imaginary constant.*
ii) has no local extremum or is constant
iii) if B(a,R) is the largest disc centered in a, contained in $O, \gamma = \partial B$ *then* $f(a) = \frac{1}{2\pi R}\oint_\gamma f = \frac{1}{2\pi R^2}\int\int_B f(x,y)dxdy$

30.3 The heat kernel

The heat kernel comes from the process of heat dissipation inside a medium which follows a differential equation like $\frac{\partial u}{\partial t} - \Delta_x u = 0$. This operator has many fascinating properties. In many ways it links the volume of a region of a manifold with its area, and so is a characteristic of a riemannian manifold.

30.3.1 Heat operator

Heat operator

Definition 2427 *The* **heat operator** *on a riemannian manifold (M,g) with metric g is the scalar linear differential operator* $D(t,x) = \frac{\partial u}{\partial t} - \Delta_x u$ *acting on functions on* $(\mathbb{R}_+ \times M, g)$.

According to our general definition a fundamental solution of the operator D(t,x) at a point $(t,y) \in \mathbb{R}_+ \times M$ is a distribution U(t,y) depending on the parameters (t,y) such that : $D'U(t,y) = \delta_{t,y}$. To account for the semi-continuity in t=0 one imposes :

$\forall \varphi \in C_{\infty c}(\mathbb{R}_+ \times M; \mathbb{R}) : D'U(t,y)\varphi(t,x) \to_{t \to 0_+} \varphi(t,y)$

A fundamental solution is said to be regular if : $U(t,y) = T(p(t,x,y))$ is induced by a smooth map : p:M$\to C_\infty(\mathbb{R}_+; \mathbb{R})$::u(t,x)=p(t,x,y) such that : $p(t,x,y) \geq 0, \forall t \geq 0 : \int_M u(t,x)\varpi_0 \leq 1$. where ϖ_0 is the volume form induced by g.

A fundamental regular solution can then be extended by setting $u(t,x) = 0$ for t$\leq$0. Then u(t,x) is smooth on $\mathbb{R} \times M \setminus (0,y)$ and satisfies $\frac{\partial u}{\partial t} - \Delta_x u = \delta_{(0,y)}$ for $y \in M$

Heat kernel operator

Fundamental regular solutions of the heat operator are given by the heat kernel operator. See One parameter group in the Banach Spaces section.

Definition 2428 *The* **heat kernel** *of a riemannian manifold (M,g) with metric g is the semi-group of operators* $\left(P(t) = e^{t\Delta}\right)_{t \geq 0}$ *where* Δ *is the laplacian on M acting on functions on M*

The domain of P(t) is well defined as : $\left\{ f \in \overline{C_{\infty c}(M; \mathbb{C})} : \Delta'T(f) \in T\left(L^2(M, \varpi_0, \mathbb{C})\right) \right\}$ where the closure is taken in $L^2(M, \varpi_0, \mathbb{C})$

It can be enlarged :

Theorem 2429 *(Taylor 3 p.38) On a riemannian compact manifold with boundary M, the heat kernel* $\left(P(t) = e^{t\Delta}\right)_{t \geq 0}$ *is a strongly continuous semi-group on the Banach space :* $\{f \in C_1(M; \mathbb{C}), f = 0 \text{ on } \partial M\}$

Theorem 2430 *(Taylor 3 p.35) On a riemannian compact manifold with boundary M :* $e^{z\Delta}$ *is a holomorphic semi-group on* $L^p(M, \varpi_0, \mathbb{C})$ *for* $1 \leq p \leq \infty$

Theorem 2431 *(Gregor'yan p.10) For any function $f \in L^2(M, \varpi_0, \mathbb{C})$ on a riemannian smooth manifold (without boundary) (M,g), the function u(t,x)=P(t,x)f(x)$\in$ $C_\infty(\mathbb{R}_+ \times M; \mathbb{C})$ and satisfies the heat equation : $\frac{\partial u}{\partial t} = \Delta_x u$ with the conditions : $u(t,.) \to f$ in $L^2(M, \varpi_0, \mathbb{C})$ when $t \to 0_+$ and $\inf(f) \leq u(t,x) \leq \sup(f)$*

Heat kernel function

The heat kernel operator, as a fundamental solution of a heat operator, has an associated Green's function.

Theorem 2432 *(Gregor'yan p.12) For any riemannian manifold (M,g) there is a unique function p, called the **heat kernel function**, $p(t,x,y) \in C_\infty(\mathbb{R}_+ \times M \times M; \mathbb{C})$ such that $\forall f \in L^2(M, \varpi_0, \mathbb{C}), \forall t \geq 0 : (P(t) f)(x) = \int_M p(t,x,y) f(y) \varpi_0(y).$*

p(t) is the integral kernel of the heat operator P(t)=$e^{t\Delta_x}$ (whence its name) and $U(t,y) = T(p(t,x,y))$ is a regular fundamental solution on the heat operator at y.

As a regular fundamental solution of the heat operator:

Theorem 2433 $\forall f \in L^2(M, \varpi_0, \mathbb{C}), u(t,x) = \int_M p(t,x,y) f(y) \varpi_0(y)$ *is solution of the PDE :*
$\frac{\partial u}{\partial t} = \Delta_x u$ *for t > 0*
$\lim_{t \to 0_+} u(t,x) = f(x)$
and $u \in C_\infty(\mathbb{R}_+ \times M; \mathbb{R})$

Theorem 2434 *The heat kernel function has the following properties :*
i) p(t,x,y) =p(t,y,x)
ii) $p(t+s,x,y) = \int_M p(t,x,z)p(s,z,y)\varpi_0(z)$
iii) $p(t,x,.) \in L^2(M, \varpi_0, \mathbb{C})$
iv) As $p \geq 0$ and $\int_M p(t,x,y) \varpi_0(x) \leq 1$ the domain of the operator P(t) can be extended to any positive or bounded measurable function f on M by : $(P(t) f)(x) = \int_M p(t,x,y) f(y) \varpi_0(y)$
v) Moreover $G(x,y) = \int_0^\infty p(t,x,y) dt$ is the Green's function of -Δ (see above).

Warning ! because the heat operator is not defined for t<0 we cannot have p(t,x,y) for t<0

The heat kernel function depends on the domain in M

Theorem 2435 *If O, O' are open subsets of M, then $O \subset O' \Rightarrow p_O(t,x,y) \leq p_{O'}(t,x,y)$. If $(O_n)_{n \in \mathbb{N}}$ is a sequence $O_n \to M$ then $p_{O_n}(t,x,y) \to p_M(t,x,y)$*

With this meaning p is the minimal fundamental positive solution of the heat equation at y

Theorem 2436 *If $(M_1, g_1), (M_2, g_2)$ are two riemannian manifolds, $(M_1 \times M_2, g_1 \otimes g_2)$ is a riemannian manifold, the laplacians Δ_1, Δ_2 commute, $\Delta_{M_1 \times M_2} = \Delta_1 + \Delta_2$, and $P_{M_1 \times M_2}(t) = P_1(t) P_2(t), p_{M_1 \times M_2}(t, (x_1, x_2), (y_1, y_2)) = p_1(t, x_1, y_1) p_2(t, x_2, y_2)$*

Theorem 2437 *Spectral resolution : if $dE(\lambda)$ is the spectral resolution of $-\Delta$ then $P(t) = \int_0^\infty e^{-\lambda t} dE(\lambda)$*

If M is some open, *relatively compact*, in a manifold N, then the spectrum of $-\Delta$ is an increasing countable sequence $(\lambda_n)_{n=1}^\infty, \lambda_n > 0$ with eigenvectors $e_n \in L^2(M, \varpi_0, \mathbb{C}) : -\Delta e_n = \lambda_n e_n$ forming an orthonormal basis. Thus for $t \geq 0, x, y \in M$:

$p(t, x, y) = \sum_{n=1}^\infty e^{-\lambda_n t} e_n(x) e_n(y)$

$\lim_{t \to \infty} \frac{\ln p(t,x,y)}{t} = -\lambda_1$

$\int_M p(t, x, x) \varpi_0(x) = \sum_{n=1}^\infty e^{-\lambda_n t}$

Heat kernel, volume and distance

The heat kernel function is linked to the relation between volume and distance in (M,g).

1. In a neighborhood of y we have :

p(t,x,y) $\sim \frac{1}{(4\pi t)^{m/2}} \left(\exp \frac{-d^2(x,y)}{4t} \right) u(x, y)$ when $t \to 0$

where m=dimM, d(x,y) is the geodesic distance between x,y and u(x,y) a smooth function.

2. If $f : M \to M$ is an isometry on (M,g), meaning a diffeomorphism preserving g, then it preserves the heat kernel function :

$p(t, f(x), f(y)) = p(t, x, y)$

3. For two Lebesgue measurable subsets A,B of M :

$\int_A \int_B p(t, x, y) \varpi_0(x) \varpi_0(y) \leq \sqrt{\varpi_0(A) \varpi_0(B)} \left(\exp \frac{-d^2(A,B)}{4t} \right)$

4. The heat kernel function in $\mathbb{R}^m$ is : $p(t, x, y) = \frac{1}{(4\pi t)^{m/2}} \left(\exp \frac{-\|x-y\|^2}{4t} \right)$

30.3.2 Brownian motion

Manifold stochastically complete

Definition 2438 *A riemannian manifold (M,g) is said to be stochastically complete if $\forall y \in M, t \geq 0 : \int_M p(t, x, y) \varpi_0(x) = 1$*

then any fundamental solution of the heat equation at y is equal to the heat kernel and is unique.

Theorem 2439 *(Gregor'yan p.15) The following conditions are equivalent :*
i) (M,g) is stochastically complete
ii) any non negative bounded solution of $\Delta f - f = 0$ on (M,g) is null
iii) the bounded Cauchy problem on (M,g) has a unique solution

A compact manifold is stochastically complete, $\mathbb{R}^m$ is stochastically complete
A geodesic ball centered in y in M is the set of points B(y,r)=$\{p \in M, d(p, y) < r\}$.
A manifold is geodesically complete iff the geodesic balls are relatively compact.
If a manifold M is geodesically complete and $\exists y \in M : \int_0^\infty \frac{r}{\log Vol(B(y,r))} dr = \infty$ then M is stochastically complete

Brownian motion

Beacause it links a time parameter and two points x,y of M, the heat kernel function is the tool of choice to define brownian motions, meaning random paths on a manifold. The principle is the following (see Gregor'yan for more).

1. Let (M,g) be a riemannian manifold, either compact or compactified with an additional point $\{\infty\}$.

Define a path in M as a map : $\varpi : [0, \infty] \to M$ such that if $\exists t_0 : \varpi(t_0) = \infty$ then $\forall t > t_0 : \varpi(t) = \infty$

Denote Ω the set of all such paths and Ω_x the set of paths starting at x..

Define on Ω_x the $\sigma-$algebra comprised of the paths such that : $\exists N, \{t_1, ..., t_N\}, \varpi(0) = x, \varpi(t_i) \in A_i$ where A_i is a mesurable subset of M. Thus we have measurables sets (Ω_x, S_x)

2. Define the random variable : $X_t = \varpi(t)$ and the stochastic process with the transition probabilities :

$P_t(x, A) = \int_A p(t, x, y) \varpi_0(y)$
$P_t(x, \infty) = 1 - P_t(x, M)$
$P_t(\infty, A) = 0$

The heat semi group relates the transition probabilities by : $P_t(x, A) = P(t) 1_A(x)$ whith the characteristic function 1_A of A.

So the probability that a path ends at ∞ is null if M is stochastically complete (which happens if it is compact).

The probability that a path is such that : $\varpi(t_i) \in A_i, \{t_1, ..., t_N\}$ is :

$P_x(\varpi(t_i) \in A_i) = \int_{A_1} ... \int_{A_N} P_{t_1}(x, y_1) P_{t_2 - t_1}(y_1, y_2) ... P_{t_N - t_{N-1}}(y_{N-1}, y_N) \varpi_0(y_1) ... \varpi_0(y_N)$

So : $P_x(\varpi(t) \in A) = P_t(x, A) = P(t) 1_A(x)$

We have a stochastic process which meets the conditions of the Kolmogoroff extension (see Measure).

Notice that this is one stochastic process among many others : it is characterized by transition probabilities linked to the heat kernel function.

3. Then for any bounded function f on M we have :

$\forall t \geq 0, \forall x \in M : (P(t) f)(x) = E_x(f(X_t)) = \int_{\Omega_x} f(\varpi(t)) P_x(\varpi(t))$

30.4   Pseudo differential operators

A linear differential operator D on the space of complex functions over $\mathbb{R}^m$ can be written :

if $f \in S(\mathbb{R}^m) : Df = (2\pi)^{-m/2} \int_{\mathbb{R}^m} P(x, it) \widehat{f}(t) e^{i\langle t, x \rangle} dt$ where P is the symbol of D.

In some way we replace the linear operator D by an integral, as in the spectral theory. The interest of this specification is that all the practical content of D is summarized in the symbol. As seen above this is convenient to find fundamental solutions of partial differential equations. It happens that this approach can be usefully generalized, whence the following definition.

30.4.1 Definition

Definition 2440 *A **pseudo differential operator** is a map, denoted P(x,D), on a space F of complex functions on $\mathbb{R}^m$: $P(x, D) : F \to F$ such that there is a function $P \in C_\infty \left(\mathbb{R}^m \times \mathbb{R}^m; \mathbb{C}\right)$ called the symbol of the operator :*

$\forall \varphi \in F : P(x, D)\varphi = (2\pi)^{-m/2} \int_{y \in \mathbb{R}^m} \int_{t \in \mathbb{R}^m} e^{i\langle t, x-y \rangle} P(x, t)\varphi(y)\, dy dt = \int_{t \in \mathbb{R}^m} e^{i\langle t, x \rangle} P(x, t)\widehat{\varphi}(t)\, dt$

Using the same function P(x,t) we can define similarly a pseudo differential operator acting on distributions.

Definition 2441 *A pseudo differential operator, denoted P(x,D'), acting on a space F' of distributions is a map : $P(x, D\prime) : F' \to F'$ such that there is a function $P \in C_\infty\left(\mathbb{R}^m \times \mathbb{R}^m; \mathbb{C}\right)$ called the symbol of the operator :*

$\forall S \in F', \forall \varphi \in F : P(x, D')(S)\varphi = S_t\left(P(x, t) e^{i\langle t, x \rangle}\widehat{\varphi}(t)\right)$

These definitions assume that the Fourier transform of a test function φ is well defined. This point is seen below.

Hörmander classes

A pseudo differential operator is fully defined through its symbol : the function P(x,t). They are classified according to the space of symbols P. The usual classes are the **Hörmander classes** denoted $D_{\rho b}^r$:

r$\in \mathbb{R}$ is the order of the class, $\rho, b \in [0, 1]$ are parameters (usually $\rho = 1, b = 0$) such that :

$\forall (\alpha) = (\alpha_1, ..\alpha_k), (\beta) = (\beta_1, .., \beta_l) : \exists C_{\alpha\beta} \in \mathbb{R} : |D_{\alpha_1..\alpha_k}(x) D_{\beta_1..\beta_l}(t) P(x, t)| \leq C_{\alpha\beta}\left(\left(1 + \|t\|^2\right)^{1/2}\right)^{r - \rho k + bl}$

and one usually indifferently says that P(x,D)$\in D_{\rho b}^r$ or P(x,t)$\in D_{\rho b}^r$

The concept of principal symbol is the following : if there are smooth functions $P_k(x, t)$ homogeneous of degree k in t, and :

$\forall n \in \mathbb{N} : P(x, t) - \sum_{k=0}^{n} P_k(x, t) \in D_{1,0}^{r-n}$ one says that $P(x, t) \in D^r$ and $P_r(x, t)$ is the principal symbol of P(x,D).

Comments

From the computation of the Fourier transform for linear differential operators we see that the definition of pseudo differential operators is close. But it is important to notice the differences.

i) Linear differential operators of order r with smooth bounded coefficients (acting on functions or distributions) are pseudo-differential operators of order r : their symbols is a map polynomial in t. Indeed the Hörmander classes are equivalently defined with the symbols of linear differential operators. But the converse is not true.

ii) Pseudo differential operators are not local : to compute the value $P(x, D)\varphi$ at x we need to know the whole of φ to compute the Fourier transform, while for a differential operator it requires only the values of the function and their

derivatives at x. So a pseudo differential operator is not a differential operator, with the previous definition (whence the name pseudo).

On the other hand to compute a pseudo differential operator we need only to know the function, while for a differential operator we need to know its r first derivatives. A pseudo differential operator is a map acting on sections of F, a differential operator is a map acting (locally) on sections of $J^r F$.

iii) Pseudo differential operators are linear with respect to functions or distributions, but they are not necessarily linear with respect to sections of the J^r extensions.

iv) and of course pseudo differential operators are scalar : they are defined for functions or distributions, not sections of fiber bundles.

So pseudo differential operators can be seen as a "proxy" for linear scalar differential operators. Their interest lies in the fact that often one can reduce a problem in analysis of pseudo-differential operators to a sequence of algebraic problems involving their symbols, and this is the essence of microlocal analysis.

30.4.2 General properties

Linearity

Pseudo-differential operators are linear endomorphisms on the spaces of functions or distributions (but not their jets extensions) :

Theorem 2442 *(Taylor 2 p.3) The pseudo differential operators in the class $D_{\rho b}^r$ are such that :*
$P(x, D) : S(\mathbb{R}^m) \to S(\mathbb{R}^m)$
$P(x, D') : S(\mathbb{R}^m)' \to S(\mathbb{R}^m)'$ *if b < 1*
and they are linear continuous operators on these vector spaces.

So the same function P(x,t) can indifferently define a pseudo differential operator acting on functions or distributions, with the Fréchet space $S(\mathbb{R}^m)$ and its dual $S(\mathbb{R}^m)'$.

Moreover :

Theorem 2443 *(Taylor 2 p.17) If P(x,D)$\in D_{\rho b}^0$ and $b < \rho, s \in \mathbb{R}$ then :*
$P(x, D) : L^2(\mathbb{R}^m dx, \mathbb{C}) \to L^2(\mathbb{R}^m dx, \mathbb{C})$
$P(x, D) : H^s(\mathbb{R}^m) \to H^{s-r}(\mathbb{R}^m)$
(Taylor 3 p.18) If P(x,D)$\in D_{1b}^0$ and $b \in [0, 1]$, then :
$P(x, D) : L^p(\mathbb{R}^m dx, \mathbb{C}) \to L^p(\mathbb{R}^m dx, \mathbb{C})$ *for 1< p < ∞*

Composition

Pseudo-differential are linear endomorphisms, so they can be composed and we have :

Theorem 2444 *(Taylor 2 p.11) The composition $P_1 \circ P_2$ of two pseudo-differential operators is again a pseudo-differential operator $Q(x, D)$:*

If $P_1 \in D^{r_1}_{\rho_1 b_1}, P_2 \in D^{r_2}_{\rho_2 b_2} : b_2 \leq \min(\rho_1, \rho_2) = \rho, \delta = \max(\delta_1, \delta_2)$

$Q(x, D) \in D^{r_1 + r_2}_{\rho b}$

$Q(x, t) \sim \sum_{\|\alpha\| \geq 0} \frac{i^{\|\alpha\|}}{\alpha!} D_\alpha(t_1) P(x, t_1) D_\alpha(x) P_2(x, t)$ when $x, t \to \infty$

The commutator of two pseudo differential operators is defined as : $[P_1, P_2] = P_1 \circ P_2 - P_2 \circ P_1$ and we have :

$[P_1, P_2] \in D^{r_1 + r_2 - \rho - b}_{\rho b}$

Adjoint of a pseudo differential operator:

$\forall p : 1 \leq p \leq \infty : S(\mathbb{R}^m) \subset L^p(\mathbb{R}^m, dx, \mathbb{C})$ so $S(\mathbb{R}^m) \subset L^2(\mathbb{R}^m, dx, \mathbb{C})$ with the usual inner product : $\langle \varphi, \psi \rangle = \int_{\mathbb{R}^m} \overline{\varphi} \psi dx$

A pseudo differential operator $P(x, D) : S(\mathbb{R}^m) \to S(\mathbb{R}^m)$ is a continuous operator on the Hilbert space $S(\mathbb{R}^m)$ and has an adjoint : $P(x, D)^* : S(\mathbb{R}^m) \to S(\mathbb{R}^m)$ such that : $\langle P(x, D)\varphi, \psi \rangle = \langle \varphi, P(x, D)^* \psi \rangle$

Theorem 2445 *(Taylor 2 p.11)The adjoint of a pseudo differential operator is also a pseudo-differential operator :*

$P(x, D^*) \varphi = \int_{t \in \mathbb{R}^m} e^{i\langle t, x \rangle} P(x, t)^* \widehat{\varphi}(t) dt$

If $P(x, D) \in D^r_{\rho b}$ then $P(x, D)^* \in D^r_{\rho b}$ with : $P(x, t)^* \sim \sum_{k \geq 0} \frac{i^k}{\beta_1! \alpha_2! .. \beta_m!} i^k D_{\alpha_1 .. \alpha_k}(t) D_{\alpha_1 .. \alpha_k}(x) P(x, t)$ when $x, t \to \infty$ where β_p is the number of occurences of the index p in $(\alpha_1, ..., \alpha_k)$

Transpose of a pseudo differential operator:

Theorem 2446 *The transpose of a pseudo differential operator $P(x, D) : S(\mathbb{R}^m) \to S(\mathbb{R}^m)$ is a pseudo operator $P(x, D)^t : S(\mathbb{R}^m)' \to S(\mathbb{R}^m)'$ with symbol $P(x, t)^t = P(t, x)$*

Proof. $P(x, t)^t$ is such that : $\forall S \in S(\mathbb{R}^m)', \forall \varphi \in S(\mathbb{R}^m) : P(x, D)^t(S)(\varphi) = S(P(x, D)\varphi)$

$P(x, D)^t(S)\varphi = S_t\left(P(x, t)^t e^{i\langle t, x \rangle} \widehat{\varphi}(t)\right) = S_t\left((2\pi)^{m/2} \mathcal{F}^*_x\left(P(x, t)^t \widehat{\varphi}(t)\right)\right) = (2\pi)^{m/2} \mathcal{F}^*_x\left(S_t\left(P(x, t)^t \widehat{\varphi}(t)\right)\right)$

$S(P(x, D)\varphi) = S_x\left(\int_{t \in \mathbb{R}^m} e^{i\langle t, x \rangle} P(x, t) \widehat{\varphi}(t) dt\right) = S_x\left((2\pi)^{m/2} \mathcal{F}^*_t(P(x, t) \widehat{\varphi}(t))\right) = (2\pi)^{m/2} \mathcal{F}^*_t(S_x(P(x, t) \widehat{\varphi}(t)))$

with $\mathcal{F}^* : S(\mathbb{R}^m)' \to S(\mathbb{R}^m)' :: \mathcal{F}^*(S)(\varphi) = S(\mathcal{F}^*(\varphi))$ ∎

30.4.3 Schwartz kernel

For $\varphi, \psi \in S(\mathbb{R}^m)$ the bilinear functional :

$K(\varphi \otimes \psi) = \int_{\mathbb{R}^m} \varphi(x)(P(x, D)\psi) dx = \int \int \int_{\mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m} e^{i\langle t, x - y \rangle} \varphi(x) \psi(y) P(x, t) dx dy dt = \int \int_{\mathbb{R}^m \times \mathbb{R}^m} k(x, y) \varphi(x) \psi(y) dx dy$

with $k(x, y) = \int_{\mathbb{R}^m} P(x, t) e^{i\langle t, x - y \rangle} dt$ and $\varphi \otimes \psi(x, y) = \varphi(x) \psi(y)$

can be extended to the space $C_\infty(\mathbb{R}^m \times \mathbb{R}^m; \mathbb{C})$ and we have the following :

Theorem 2447 *(Taylor 2 p.5) For any pseudo differential operator $P(x, D) \in$ $D^r_{\rho b}$ on $S(\mathbb{R}^m)$ there is a distribution $K \in (C_{\infty c}(\mathbb{R}^m \times \mathbb{R}^m; \mathbb{C}))'$ called the* **Schwartz kernel** *of P(x,D) such that : $\forall \varphi, \psi \in S(\mathbb{R}^m) : K(\varphi \otimes \psi) = \int_{\mathbb{R}^m} \varphi(x)(P(x, D)\psi) dx$*

K is induced by the function $k(x, y) = \int_{\mathbb{R}^m} P(x, t) e^{i\langle t, x-y \rangle} dt$

If $\rho > 0$ then $k(x, y)$ belongs to $C_\infty(\mathbb{R}^m \times \mathbb{R}^m; \mathbb{C})$ for $x \neq y$

$\exists C \in \mathbb{R} : \|\beta\| > -m - r : |D_\beta(x, y)k| \leq C \|x - y\|^{-m-r-\|\beta\|}$

The Schwartz kernel is characteristic of P(x,D) and many of the theorems about pseudo differential operators are based upon its properties.

30.4.4 Support of a pseudo-differential operator

Elliptic pseudo differential operators

Definition 2448 *(Taylor 2 p.14) A pseudo differential operator $P(x, D) \in$ $D^r_{\rho b}, \rho < b$ on $S(\mathbb{R}^m)$ is said to be* **elliptic** *if*

$$\exists c \in \mathbb{R} : \forall \|t\| > c : \left| P(x, t)^{-1} \right| \leq c \|t\|^{-r}$$

Then if $H_c : \mathbb{R}^m \to \mathbb{R} :: \|t\| \leq c : H_c(t) = 0, \|t\| > c : H_c(t) = 1$ we have $H_c(t) P(x, t)^{-1} = Q(x, t)$ and the pseudo differential operator Q(x,D)$\in D^{-r}_{\rho b}$ is such that :

$Q(x, D) \circ P(x, D) = Id + R_1(x, D)$

$P(x, D) \circ Q(x, D) = Id - R_2(x, D)$

with $R_1(x, D), R_2(x, D) \in D^{-(\rho-b)}_{\rho b}$

so Q(x,D) is a proxy for a left and right inverse of P(x,D) : this is called a two sided **parametrix**.

Moreover we have for the singular supports :

$\forall S \in S(\mathbb{R}^m)' : SSup(P(x, D)S) = SSup(S)$

which entails that an elliptic pseudo differerential operator does not add any critical point to a distribution (domains where the distribution cannot be identified with a function). Such an operator is said be microlocal. It can be more precise with the following.

Characteristic set

(Taylor 2 p.20):

The **characteristic set** of a pseudo differential operator P(x,D)$\in D^r$ with principal symbol $P_r(x, t)$ is the set :

$Char(P(x, D)) = \{(x, t) \in \mathbb{R}^m \times \mathbb{R}^m, (x, t) \neq (0, 0) : P_r(x, t) \neq 0\}$

The **wave front set** of a distribution $S \in H^{-\infty}(\mathbb{R}^m) = \cup_s H^{-s}(\mathbb{R}^m)$ is the set :

$WF(S) = \cap_P \{Char(P(x, D), P(x, D) \in S^0 : P(x, D)S \in T(C_\infty(\mathbb{R}^m; \mathbb{C}))\}$

then $P_{r1}(WF(S)) = SSup(S)$ with the projection $\text{Pr}_1 : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m ::$ $P_{r1}(x, t) = x$

Essential support

(Taylor 2 p.20)

A pseudo differential operator $P(x,D) \in D^r_{\rho b}$ is said to be of order $-\infty$ at $(x_0, t_0) \in \mathbb{R}^m \times \mathbb{R}^m$ if :

$\forall (\alpha) = (\alpha_1, ..\alpha_k), (\beta) = (\beta_1, .., \beta_l), \forall N \in \mathbb{N}$:

$\exists C \in \mathbb{R}, \forall \xi > 0 : \left| D_{\alpha_1..\alpha_k}(x) D_{\beta_1..\beta_l}(t) P(x,t) \right|_{(x_0, \xi t_0)} \leq C \left\| \xi t_0 \right\|^{-N}$

A pseudo differential operator $P(x,D) \in D^r_{\rho b}$ is said to be of order $-\infty$ in an open subset U$\subset \mathbb{R}^m \times \mathbb{R}^m$ if it is of order $-\infty$ at any point $(x_0, t_0) \in U$

The **essential support** $ESup(P(x,D))$ of a pseudo differential operator $P(x,D) \in D^r_{\rho b}$ is the smallest closed subset of $\mathbb{R}^m \times \mathbb{R}^m$ on the complement of which P(x,D) is of order $-\infty$

For the compose of two pseudo-differential goperators $P_1, P_2 :$ ESup$(P_1 \circ P_2) \subset ESup(P_1) \cap ESup(P_2)$

If $S \in H^{-\infty}(\mathbb{R}^m), P(x,D) \in D^r_{\rho b}, \rho > 0, b < 1$ then $WF(P(x,D)S) \subset WF(S) \cap ESup(P(x,D))$

If $S \in H^{-\infty}(\mathbb{R}^m), P(x,D) \in D^r_{\rho b}, b < \rho$ is elliptic, then P(x,D) has the microlocal regularity property : WF(P(x,D)S)=WF(S).

For any general solution U of the scalar hyperbolic equation : $\frac{\partial S}{\partial t} = iP(x,D)S$ on $S(\mathbb{R}^m)'$ with P(x,D)$\in D^1$ with real principal symbol : WF(U)=C(t)WF(S) where C(t) is a smooth map.

31 DIFFERENTIAL EQUATIONS

Differential equations are usually met as equations whose unknown variable is a map $f : E \to F$ with conditions such as : $\forall x \in E$:L$\left(x, f\left(x\right), f'\left(x\right), ..., f^{(r)}\left(x\right)\right) = 0$ subject to conditions such as $\forall x \in A$:M$\left(x, f\left(x\right), f'\left(x\right), ..., f^{(r)}\left(x\right)\right) = 0$ for some subset A of E.

Differential equations raise several questions : existence, unicity and regularity of a solution, and eventually finding an explicit solution, which is not often possible. The problem is "well posed" when, for a given problem, there is a unique solution, depending continuously on the parameters.

Ordinary differential equations (ODE) are equations for which the map f depends on a unique real variable. For them there are general theorems which answer well to the first two questions, and many ingenious more or less explicit solutions.

Partial differential equations (PDE) are equations for which the map f depends of more than one variable, so x is in some subset of $\mathbb{R}^m, m > 1$. There are no longer such general theorems. For linear PDE there are many results, and some specific equations of paramount importance in physics will be reviewed with more details. Non linear PDE are a much more difficult subject, and we will limit ourselves to some general results.

Anyway the purpose is not here to give a list of solutions : they can be found at some specialized internet sites

(such that : http://eqworld.ipmnet.ru/en/solutions/ode.htm), and in the exhaustive handbooks of A.Polyanin and V.Zaitsev.

31.1 Ordinary Differential equations (ODE)

31.1.1 Definitions

1. As $\mathbb{R}$ is simply connected, any vector bundle over $\mathbb{R}$ is trivial and a r order ordinary differential equation is an evolution equation :

$D : J^r X \to V_2$ is a differential operator

the unknown function $X \in C(I; V_1)$

I is some interval of $\mathbb{R}$, V_1 a m dimensional complex vector space, V_2 is a n dimensional complex vector space.

The Cauchy conditions are : $X(a) = X_0, X'(a) = X_1, ...X^{(r)}(a) = X_r$ for some value x=a$\in I \subset \mathbb{R}$

2. An ODE of order r can always be replaced by an equivalent ODE of order 1 :

Define : $Y_k = X^{(k)}, k = 0...r$,$Y \in W = (V_1)^{r+1}$

Replace $J^r X = J^1 Y \in C(I; W)$

Define the differential operator :

$G : J^1 W \to V_2 :: G\left(x, j_x^1 Y\right) = D\left(x, j_x^r X\right)$

with initial conditions : $Y(a) = (X_0, ..., X_r)$

3. Using the implicit map theorem (see Differential geometry, derivatives). a first order ODE can then be put in the form : $\frac{dX}{dx} = L(x, X(x))$

31.1.2 Fundamental theorems

Existence and unicity

Due to the importance of the Cauchy problem, we give several theorems about existence and unicity.

Theorem 2449 *(Schwartz 2 p.351) The 1st order ODE :* $\frac{dX}{dx} = L(x, X(x))$ *with :*

i) $X : I \to O$, I an interval in $\mathbb{R}$, O an open subset of an affine Banach space E

ii) $L : I \times O \to E$ a continuous map, globally Lipschitz with respect to the second variable :

$\exists k \geq 0 : \forall (x, y_1), (x, y_2) \in I \times O : \|L(x, y_1) - L(x, y_2)\|_E \leq k \|y_1 - y_2\|_E$

iii) the Cauchy condition : $x_0 \in I, y_0 \in O : X(x_0) = y_0$

has a unique solution and : $\|X(x) - y_0\| \leq e^{k|x-x_0|} \int_{|x_0,x|} \|L(\xi, y_0)\| d\xi$

The problem is equivalent to the following : find X such that : $X(x) = y_0 + \int_{x_0}^{x} L(\xi, f(\xi)) d\xi$ and the solution is found by the Picard iteration method : $X_{n+1}(x) = y_0 + \int_{x_0}^{x} L(\xi, X_n(\xi)) d\xi :: X_n \to X$. Moreover the series : $X_0 + \sum_{k=1}^{n} (X_k - X_{k-1}) \to X$ is absolutely convergent on any compact of I.

if L is not Lipshitz :

i) If E is finite dimensional there is still a solution (Cauchy-Peano theorem - Taylor 1 p.110), but it is not necessarily unique.

ii) If E is infinite dimensional then the existence itself is not assured.

2. In the previous theorem L is globally Lipschitz. This condition can be weakened as follows :

Theorem 2450 *(Schwartz 2 p.364) The 1st order ODE : $\frac{dX}{dx} = L\left(x, X\left(x\right)\right)$ with*

i) $X : I \to O$, I an interval in $\mathbb{R}$, O an open subset of a finite dimensional affine Banach space E
ii) $L : I \times O \to \overrightarrow{E}$ a continuous map, locally Lipschitz with respect to the second variable:
$\forall a \in I, \forall y \in O, \exists n\left(a\right) \subset I, n\left(y\right) \subset O : \exists k \geq 0 : \forall\left(x, y_1\right),\left(x, y_2\right) \in n\left(a\right) \times n\left(y\right) : \left\|L\left(x, y_1\right) - L\left(x, y_2\right)\right\|_E \leq k\left\|y_1 - y_2\right\|_E$
iii) the Cauchy condition $x_0 \in I, y_0 \in O : X\left(x_0\right) = y_0$
has a unique solution in a maximal interval $\left]a, b\right[: \quad I' \subset \left]a, b\right[\subset I$, and for any compact C in O, $\exists \varepsilon > 0 : b - x < \varepsilon \Rightarrow X\left(x\right) \notin C$

(meaning that X(x) tends to the border of O, possibly infinite).
The solution is still given by the integrals $X_n\left(x\right) = y_0 + \int_{x_0}^{x} L\left(\xi, X_{n-1}\left(\xi\right)\right) d\xi$
and the series : $X_0 + \sum_{k=1}^{n}\left(X_k - X_{k-1}\right)$

3. The previous theorem gives only the existence and unicity of local solutions. We can have more.

Theorem 2451 *(Schwartz 2 p.370) The 1st order ODE : $\frac{dX}{dx} = L\left(x, X\left(x\right)\right)$ with :*

i) $X : I \to E$, I an interval in $\mathbb{R}$, E an affine Banach space
ii) $L : I \times E \to \overrightarrow{E}$ a continuous map such that :
$\exists \lambda, \mu \geq 0, A \in E : \forall\left(x, y\right) \in I \times E : \left\|L\left(x, y\right)\right\| \leq \lambda\left\|y - A\right\| + \mu$
$\forall \rho > 0$ L is Lipschitz with respect to the second variable on $I \times B\left(0, \rho\right)$
iii) the Cauchy condition $x_0 \in I, y_0 \in O : X\left(x_0\right) = y_0$
has a unique solution defined on I

We have the following if $E = \mathbb{R}^m$:

Theorem 2452 *(Taylor 1 p.111) The 1st order ODE : $\frac{dx}{dx} = L\left(x, X\left(x\right)\right)$ with*
i) $X : I \to O$, I an interval in $\mathbb{R}$, O an open subset of $\mathbb{R}^m$
ii) $L : I \times O \to \mathbb{R}^m$ a continuous map such that :
$\forall\left(x, y_1\right),\left(x, y_2\right) \in I \times O : \left\|L\left(x, y_1\right) - L\left(x, y_2\right)\right\|_E \leq \lambda\left(\left\|y_1 - y_2\right\|\right)$ where : $\lambda \in C_{0b}\left(\mathbb{R}_+; \mathbb{R}_+\right)$ is such that $\int \frac{ds}{\lambda\left(s\right)} = \infty$
iii) the Cauchy condition $x_0 \in I, y_0 \in O : X\left(x_0\right) = y_0$
has a unique solution defined on I

Majoration of solutions

Theorem 2453 *(Schwartz 2 p.370) Any solution g of the scalar first order ODE :* $\frac{dX}{dx} = L(x, X(x))$*, with*

i) $X : [a, b] \to \mathbb{R}_+$,

ii) $L:[a, b] \times \mathbb{R}_+ \to \mathbb{R}_+$ continuous,

iii) $F \in C_1([a, b]; E)$, E affine normed space such that :

$\exists A \in E : \forall x \in [a, b] : \|F'(x)\| < L(x, \|F(x) - A\|)$

$y_0 = \|F(a) - A\|$

iv) the Cauchy condition $X(a) = y_0$

is such that : $\forall x > a : \|F(x) - A\| < g(x)$

Differentiability of solutions

Theorem 2454 *(Schwartz 2 p.377) Any solution of the 1st order ODE :* $\frac{dX}{dx} = L(x, X(x))$ *with*

i) $X : I \to O$, J an interval in $\mathbb{R}$, O an open subset of an affine Banach space E

ii) $L : I \times O \to E$ a class r map,

iii) the Cauchy condition $x_0 \in I, y_0 \in O : X(x_0) = y_0$

on IxO is a class r+1 map on O

If E is a finite m dimensional vector space, and if in the neighborhood $n(x_0)$ of the Cauchy conditions $(x_0, y_0) \in I \times O$ the Cauchy problem has a unique solution, then there are exactly m independant conservations laws, locally defined on $n(x_0)$.

Generally there are no conservations laws globally defined on the whole of O.

Solution depending on a parameter

1. Existence and continuity

Theorem 2455 *(Schwartz 2 p.353) Let be the 1st order ODE :* $\frac{\partial X}{\partial x} = L(x, \lambda, X(x, \lambda))$ *with :*

i) $X : I \times \Lambda \to O$, I an interval in $\mathbb{R}$, O an open subset of an affine Banach space E, Λ a topological space

ii) $L : I \times \Lambda \times O \to E$ a continuous map, globally Lipschitz with respect to the second variable :

$\exists k \geq 0 : \forall x \in I, y_1, y_2 \in O, \lambda \in \Lambda : \|L(x, \lambda, y_1) - L(x, \lambda, y_2)\|_E \leq k \|y_1 - y_2\|_E$

iii) the Cauchy condition : $X(x_0(\lambda), \lambda) = y_0(\lambda)$ where $y_0 : \Lambda \to O, x_0 : \Lambda \to J$ are continuous maps

Then for any $\lambda_0 \in \Lambda$ the Cauchy problem :

$\frac{dX}{dx} = L(x, \lambda_0, X(x, \lambda_0)), f(x_0(\lambda_0), \lambda_0) = y_0(\lambda_0)$ *has a unique solution $X(x, \lambda_0)$*

and $X(x, \lambda_0) \to X(x, \lambda_1)$ uniformly on any compact of I when $\lambda_0 \to \lambda_1$

2. We have a theorem with weaker Lipschitz conditions :

Theorem 2456 *(Schwartz 2 p.374) Let be the 1st order ODE :* $\frac{\partial X}{\partial x} = L(x, \lambda, X(x, \lambda))$
with :

i) $X : I \times \Lambda \to O$, *I an interval in* $\mathbb{R}$ *, O an open subset of an affine Banach space E,* Λ *a topological space*

ii) $L : I \times \Lambda \times O \to \overrightarrow{E}$ *a continuous map, locally Lipschitz with respect to the second variable :*

$\forall(a, y, \mu) \in I \times O \times \Lambda : \exists n(a) \subset I, n(p) \subset O, n(\mu) \in \Lambda, \exists k \geq 0 :$

$\forall(x, y_1, \mu), (x, y_2, \nu) \in n(a) \times n(y) \times n(\mu) : \|L(x, \lambda, y_1) - L(x, \lambda, y_2)\|_E \leq k\|y_1 - y_2\|_E$

iii) the Cauchy condition : $X(x_0(\lambda), \lambda) = y_0(\lambda)$ *where* $y_0 : \Lambda \to O, x_0 : \Lambda \to I$ *are continuous maps*

Then for any $\lambda_0 \in \Lambda$ *, there is an interval* $I_0(\lambda_0) \subset I$ *such that :*

i) for any compact $K \subset I_0(\lambda_0)$, *there are a neighborhood* $n(X(K, \lambda_0))$ *of* $X(K, \lambda_0)$, *a neighborhood* $n(\lambda_0)$ *of* λ_0

such that the Cauchy problem has a unique solution on $K \times n(\lambda_0)$ *valued in* $n(X(K, \lambda_0))$

ii) $X(., \lambda) \to X(., \lambda_0)$ *when* $\lambda \to \lambda_0$ *uniformly on K*

iii) f is continuous in $I_0(\lambda_0) \times \Lambda$

3. Differentiability of the solution with respect to the parameter :

Theorem 2457 *(Schwartz 2 p.401) Let be the 1st order ODE :* $\frac{\partial X}{\partial x} = L(x, \lambda, X(x, \lambda))$
with :

i) $X : I \times \Lambda \to O$, *I an interval in* $\mathbb{R}$ *, O an open subset of an affine Banach space E,* Λ *a topological space*

ii) $L : I \times \Lambda \times O \to \overrightarrow{E}$ *a continuous map, with a continuous partial derivative* $\frac{\partial L}{\partial y} : I \times \Lambda \times O \to \mathcal{L}\left(\overrightarrow{E}; \overrightarrow{E}\right)$

iii) the Cauchy condition : $X(x_0(\lambda), \lambda) = y_0(\lambda)$ *where* $y_0 : \Lambda \to O, x_0 : \Lambda \to I$ *are continuous maps*

If for $\lambda_0 \in \Lambda$ *the ODE has a solution* X_0 *defined on I then:*

i) there is a neighborhood $n(\lambda_0)$ *such that the ODE has a unique solution* $X(x, \lambda)$ *for* $(x, \lambda) \in I \times n(\lambda_0)$

ii) the map $X : n(\lambda) \to C_b\left(I; \overrightarrow{E}\right)$ *is continuous*

iii) if Λ *is an open subset of an affine normed space F,* $L \in C_r\left(I \times \Lambda \times O; \overrightarrow{E}\right), r \geq 1$ *then*

the solution $X \in C_{r+1}\left(I \times \Lambda; \overrightarrow{E}\right)$ *and :*

iv) If $x_0 \in C_1(\Lambda; I), y_0 \in C_1(\Lambda; O)$ *the derivative* $\varphi(x) = \frac{\partial X}{\partial \lambda}(x, \lambda) \in C\left(I; \mathcal{L}\left(\overrightarrow{F}; \overrightarrow{E}\right)\right)$ *is solution of the ODE :*

$\frac{d\varphi}{dx} = \frac{\partial L}{\partial y}(x_0(\lambda), \lambda_0, X_0(x)) \circ \varphi(x) + \frac{\partial L}{\partial \lambda}(x_0(\lambda), \lambda_0, X_0(x))$

with the Cauchy conditions : $\varphi(x_0(\lambda_0)) = \frac{dy_0}{d\lambda}(\lambda_0) - L(x_0(\lambda_0), \lambda_0, X_0(x_0(\lambda_0))) \frac{dx_0}{d\lambda}(\lambda_0)$

6. Differentiability of solutions with respect to the initial conditions

Theorem 2458 *(Taylor 1.p.28) If there is a solution $Y(x, x_0)$ of the 1st order ODE : $\frac{dX}{dx} = L(X(x))$ with:*

i) $X : I \to O$, I an interval in $\mathbb{R}$, O an open convex subset of a Banach vector space E

ii) $L : O \to E$ a class 1 map

iii) the Cauchy condition $x_0 \in I, y_0 \in O : X(x_0) = y_0$

over I, whenever $x_0 \in I$, then

i) Y is continuously differentiable with respect to x_0

ii)the partial derivative : $\varphi(x, x_0) = \frac{\partial Y}{\partial x_0}$ is solution of the ODE :$\frac{\partial \varphi}{\partial x}(x, x_0) = \frac{\partial}{\partial y} L(f(x, x_0)) \varphi(x, x_0), \varphi(x_0, x_0) = y_0$

iii) If $L \in C_r(I; O)$ then $Y(x, x_0)$ is of class r in x_0

iv) If E is finite dimensional and L real analytic, then $Y(x, x_0)$ is real analytic in x_0

ODE on manifolds

So far we require only initial conditions from the solution. One can extend the problem to the case where X is a path on a manifold.

Theorem 2459 *(Schwartz p.380) A class 1 vector field V on a a real class $r > 1$ m dimensional manifold M with charts $(O_i, \varphi_i)_{i \in I}$ is said to be locally Lipschitz if for any p in M there is a neighborhood and an atlas of M such that the maps giving the components of V in a holonomic basis $v_i : \varphi_i(O_i) \to \mathbb{R}^m$ are Lipschitz.*

The problem, find :

$c : I \to M$ where I is some interval of $\mathbb{R}$ which comprises 0

$c'(t) = V(c(t))$

V(0)=p where p is some fixed point in M

defines a system of 1st order ODE, expressed in charts of M and the components $(v^\alpha)_{\alpha=1}^m$ of V :

If V is locally Lipschitz, then for any p in M, there is a maximal interval J such that there is a unique solution of these equations.

31.1.3   Linear ODE

(Schwartz 2 p.387)

General theorems

Definition 2460 *The first order ODE : $\frac{dX}{dx} = L(x)X(x)$ with:*

i) $X : I \to E$, I an interval in $\mathbb{R}$, E a Banach vector space

ii) $\forall x \in I : L(x) \in \mathcal{L}(E; E)$

iii) the Cauchy conditions : $x_0 \in I, y_0 \in E : X(x_0) = y_0$

is a linear 1st order ODE

Theorem 2461 *For any Cauchy conditions a first order linear ODE has a unique solution defined on I*

If $\forall x \in I : \|L(x)\| \leq k$ then $\forall x \in J : \|X(x)\| \leq \|y_0\| e^{k|x|}$

If $L \in C_r(I; \mathcal{L}(E; E))$ then the solution is a class r map : $X \in C_r(I; E)$

The set V of solutions of the Cauchy problem, when (x_0, y_0) varies on IxE, is a vector subspace of $C_r(I; E)$ and the map : $F : V \to E :: F(X) = X(x_0)$ is a linear map.

If E is finite m dimensional then V is m dimensional.

If m=1 the solution is given by : $X(x) = y_0 \exp \int_{x_0}^{x} A(\xi) \, d\xi$

Resolvent

Theorem 2462 *For a first order linear ODE there is a unique map : $R : E \times E \to \mathcal{L}(E; E)$ called evolution operator (or resolvent) characterized by : $\forall X \in V, \forall x_1, x_2 \in E : X(x_2) = R(x_2, x_1) X(x_1)$*

R has the following properties :

$\forall x_1, x_2, x_3 \in E :$

$R(x_3, x_1) = R(x_3, x_2) \circ R(x_2, x_1)$

$R(x_1, x_2) = R(x_2, x_1)^{-1}$

$R(x_1, x_1) = Id$

R is the unique solution of the ODE :

$\forall \lambda \in E : \frac{\partial R}{\partial x}(x, \lambda) = L(x) R(x, \lambda)$

$R(x, x) = Id$

If $L \in C_r(I; \mathcal{L}(E; E))$ then the resolvent is a class r map : $R \in C_r(I \times I; \mathcal{L}(E; E))$

Affine equation

Definition 2463 *An affine 1st order ODE (or inhomogeneous linear ODE) is* $\frac{dX}{dx} = L(x)X(x) + M(x)$ *with :*

i) $X : I \to E$, I an interval in $\mathbb{R}$, E a Banach vector space

ii) $\forall x \in I : L(x) \in \mathcal{L}(E; E)$

iii) $M : I \to E$ is a continuous map

iv) Cauchy conditions : $x_0 \in I, y_0 \in E : X(x_0) = y_0$

The homogeneous linear ODE associated is given by $\frac{dX}{dx} = L(x)X(x)$

If g is a solution of the affine ODE, then any solution of the affine ODE is given by : $X = g + \varphi$ where φ is the general solution of : $\frac{d\varphi}{dx} = L(x)\varphi(x)$

For any Cauchy conditions (x_0, y_0) the affine ODE has a unique solution given by :

$X(x) = R(x, x_0) y_0 + \int_{x_0}^{x} R(x, \xi) M(\xi) \, d\xi$

where R is the resolvent of the associated homogeneous equation.

If $E = \mathbb{R}$ the solution reads :

$X(x) = e^{\int_{x_0}^{x} A(\xi) d\xi} \left(y_0 + \int_{x_0}^{x} M(\xi) e^{-\int_{x_0}^{\xi} A(\eta) d\eta} d\xi \right) = y_0 e^{\int_{x_0}^{x} A(\xi) d\xi} + \int_{x_0}^{x} M(\xi) e^{\int_{\xi}^{x} A(\eta) d\eta} d\xi$

If L is a constant operator $L \in \mathcal{L}(E; E)$ and M a constant vector then the solution of the affine equation is :

$X(x) = e^{(x-x_0)A}\left(y_0 + \int_{x_0}^x Me^{-(\xi-x_0)A}d\xi\right) = e^{(x-x_0)A}y_0 + \int_{x_0}^x Me^{(x-\xi)A}d\xi$

$R(x_1, x_2) = \exp[(x_1 - x_2)A]$

Time ordered integral

For the linear 1st order ODE $\frac{dX}{dx} = L(x)X(x)$ with:

i) $X : I \to \mathbb{R}^m$, I an interval in $\mathbb{R}$

ii) $L \in C_0(J; \mathcal{L}(\mathbb{R}^m; \mathbb{R}^m))$

iii) Cauchy conditions : $x_0 \in I, y_0 \in \mathbb{R}^m : X(x_0) = y_0$

the map L and the resolvent R are mxm matrices.

For $a, b \in I$, $n \in \mathbb{N}$, and a partition of $[a, b] : [a = t_0, t_1, ..., t_n = b], \Delta_k = t_k - t_{k-1}$, if X is Riemann integrable and $X(x) \geq 0$ then :

$R(b, a) = \lim_{n\to\infty} \prod_{k=0}^{k=n} \exp(\Delta_k L(x_k)) = \lim_{n\to\infty} \prod_{k=0}^{k=n} (I + L(x_k)\Delta_k)$

$= \lim_{n\to\infty} \prod_{k=0}^{k=n-1} \int_{x_k}^{x_{k+1}} (\exp L(\xi)) d\xi$

So :

$X(x) = R(x, x_0) X(x_0) = \left(\lim_{n\to\infty} \prod_{k=0}^{k=n} \exp(\Delta_k L(x_k))\right) y_0$

31.2 Partial differential equations (PDE)

31.2.1 General definitions and results

Definition of a PDE

The most general and, let us say, the "modern" way to define a differential equation is through the jet formalism.

Definition 2464 *A differential equation of order r is a closed subbundle F of the r jet extension $J^r E$ of a fibered manifold E.*

If the fibered manifold E is E(N,π) then :

- the base of F is a submanifold M of N and $\pi_F^r = \pi_E^r|_M$.

- the fibers of F over M are themselves defined usually through a family of differential operators with conditions such as: $D_k(X) = 0$.

- a solution of the PDE is a section $X \in \mathfrak{X}_r(E)$ such that $\forall x \in M : J^r X(x) \in F$

- the set of solutions is a the subset $S = (J^r)^{-1}(F) \subset \mathfrak{X}_r(E)$

If S is a vector space we have a homogeneous PDE. Then the "superposition principle" applies : any linear combination (with fixed coefficients) of solutions is still a solution.

If S is an affine space then the underlying vector space gives "general solutions" and any solution of the PDE is obtained as the sum of a particular solution and any general solution.

Differential equations can be differentiated, and give equations of higher order. If the s jet extension $J^s F$ of the subbundle F is a differential equation (meaning a closed subbundle of $J^{r+s}E$) the equation is said to be regular. A necessary condition for a differential equation to have a solution is that the maps : $J^s F \to F$ are onto, and then, if it is regular, there is a bijective correspondance between the solution of the r+s order problem and the r order problem.

Conversely an integral of a differential equation is a section $Y \in \mathfrak{X}_k(E), k < r$ such that $J^{r-k}Y \in F$. In physics it appears quite often as a conservation law : the quantity Y is preserved inside the problem. Indeed if 0 belongs to F then Y=Constant brings a solution. It can be used to lower the order of a differential equation : F is replaced by a a subbundle G of $J^k E$ defined through the introduction of additional parameters related to Y and the problem becomes of order k : $J^k X \in G$.

The most common and studied differential equations are of the kinds :

Dirichlet problems

$D : J^r E_1 \to E_2$ is a r order differential operator between two smooth complex finite dimensional vector bundles $E_1(N, V_1, \pi_1), E_2(N, V_2, \pi_2)$ on the same real manifold N.

M is a manifold with boundary (so M itself is closed) in N, which defines two subbundles in E_1, E_2 with base $\overset{\circ}{M}$, denoted M_1, M_2 and their r jet prolongations.

A solution of the PDE is a section $X \in \mathfrak{X}(E_1)$ such that :

i) for $x \in \overset{\circ}{M} : D(J^r X) = Y_0$ where Y_0 is a given section on M_2 meaning : X is a r differentiable map for $x \in \overset{\circ}{M}$ and $\forall x \in \overset{\circ}{M} : D(x)(j_x^r X) = Y_0(x)$

ii) for $x \in \partial M : X(x) = Y_1$ where Y_1 is a given section on M_1

So if $Y_0, Y_1 = 0$ the problem is homogeneous.

Neumann problems

$D : J^r E \to E$ is a r order scalar differential operator on complex functions over a manifold N : $E = C_r(M; \mathbb{C})$.

M is a manifold with smooth riemannian boundary (∂M is a hypersurface) in N, which defines the subbundle with base $\overset{\circ}{M}$

A solution of the PDE is a section $X \in \mathfrak{X}(E)$ such that :

i) for $x \in \overset{\circ}{M} : D(J^r X) = Y_0$ where Y_0 is a given section on $\overset{\circ}{M}$ meaning : X is a r differentiable map for $x \in \overset{\circ}{M}$ and $\forall x \in \overset{\circ}{M} : D(x)(j_x^r X) = Y_0(x)$

ii) for $x \in \partial M : X'(x) n = 0$ where n is the outward oriented normal to ∂M

So if $Y_0 = 0$ the problem is homogeneous.

Evolution equations

N is a m dimensional real manifold.

D is a r order scalar differential operator acting on complex functions on $\mathbb{R} \times N$ (or $\mathbb{R}_+ \times N$) seen as maps $u \in C(\mathbb{R}; C(N; \mathbb{C}))$

So there is a family of operators D(t) acting on functions u(t,x) for t fixed

There are three kinds of PDE :

1..Cauchy problem :

The problem is to find $u \in C(\mathbb{R}; C(N; \mathbb{C}))$ such that :

i) $\forall t, \forall x \in N : D(t)(x)(J_x^r u) = f(t, x)$ where f is a given function on $\mathbb{R} \times N$

ii) with the initial conditions, called Cauchy conditions :

u(t,x) is continuous for t=0 (or $t \to 0_+$) and $\forall x \in N : \frac{\partial^s}{\partial t^s} u(0, x) = g_s(x), s = 0...r-1$

2. Dirichlet problem :

M is a manifold with boundary (so M itself is closed) in N

The problem is to find $u \in C\left(\mathbb{R}; C\left(\overset{\circ}{M}; \mathbb{C}\right)\right)$ such that :

i) $\forall t, \forall x \in \overset{\circ}{M} : D(t)(x)(J_x^r u) = f(t, x)$ where f is a given function on $\mathbb{R} \times M$

ii) with the initial conditions, called Cauchy conditions :

u(t,x) is continuous for t=0 (or $t \to 0_+$) and $\forall x \in \overset{\circ}{M} : \frac{\partial^s}{\partial t^s} u(0, x) = g_s(x), s = 0...r-1$

iii) and the Dirichlet condition :

$\forall t, \forall x \in \partial M : u(t, x) = h(t, x)$ where h is a given function on ∂M

3. Neumann problem :

M is a manifold with smooth riemannian boundary (∂M is a hypersurface) in N

The problem is to find $u \in C\left(\mathbb{R}; C\left(\overset{\circ}{M}; \mathbb{C}\right)\right)$ such that :

i) $\forall t, \forall x \in \overset{\circ}{M} : D(t)(x)(J_x^r u) = f(t, x)$ where f is a given function on $\mathbb{R} \times \overset{\circ}{M}$

ii) with the initial conditions, called Cauchy conditions :

u(t,x) is continuous for t=0 (or $t \to 0_+$) and $\forall x \in \overset{\circ}{M} : \frac{\partial^s}{\partial t^s} u(0, x) = g_s(x), s = 0...r-1$

iii) and the Neumann condition :

$\forall t, \forall x \in \partial M : \frac{\partial}{\partial x} u(t, x) n = 0$ where n is the outward oriented normal to ∂M

Box boundary

The PDE is to find X such that :

$DX = 0$ in $\overset{\circ}{M}$ where $D : J^r E_1 \to E_2$ is a r order differential operator

$X = Y$ on ∂M

and the domain M is a rectangular box of $\mathbb{R}^m : M = \{0 \le x^\alpha \le a^\alpha, \alpha = 1...m\}, a = \sum_\alpha a^\alpha \varepsilon_\alpha$

X can always be replaced by Z defined in $\mathbb{R}^m$ and periodic : for any translation $\tau_n : x + na :: Z(\tau_n x) = Z(x)$ and the ODE becomes:

Find Z such that :

$DZ = 0$ in $\mathbb{R}^m$

$Z(\tau_n a) = Y$

and use series Fourier on the components of Z.

31.2.2 Linear PDE

They are the only PDE for which there are general - however partial - theorems about the existence of solutions.

General theorems

Theorem 2465 *Cauchy-Kowalesky theorem (Taylor 1 p.433) : The PDE :*
Find u$\in C\left(\mathbb{R}^{m+1};\mathbb{C}\right)$ such that :
$Du = f$
$u\left(t_0, x\right) = g_0\left(x\right), ... \frac{\partial^s}{\partial t^s} u\left(t_0, x\right) = g_s\left(x\right), s = 0..r-1$
where D is a scalar linear differential operator on $C\left(\mathbb{R}^{m+1};\mathbb{C}\right)$:
$D\left(\varphi\right) = \frac{\partial^r}{\partial t^r}\varphi + \sum_{s=0}^{r-1} A^{\alpha_1...\alpha_s}\left(t, x\right)\frac{\partial^s}{\partial \xi^{\alpha_1}..\partial \xi^{\alpha_s}}\varphi$
If $A^{\alpha_1...\alpha_s}$ are real analytic in a neighborhood n_0 of $\left(t_0, x_0\right)$ and g_s are real analytic in a neighborhood of $\left(x_0\right)$, then in n_0 there is a unique solution u(t,x)

Theorem 2466 *(Taylor 1 p.248) The PDE : find u$\in C\left(\mathbb{R}^m;\mathbb{C}\right) : Du = f$ in $B_R = \{\|x\| < R\}$*
where D is a scalar linear r order differential operator D on $\mathbb{R}^m$ with constant coefficients, and $f \in C_{\infty c}\left(\mathbb{R}^m;\mathbb{C}\right)'$,
has always a solution.

Fundamental solution
(see Differential operator)
1. A fundamental solution at a point y of M of a linear differential operator D acting on a space of complex functions F on a manifold M is a distribution U(y)$\in C_{\infty c}\left(M;\mathbb{C}\right)'$ such that : $D'U\left(y\right) = \delta_y$ where D' on $C_{\infty c}\left(M;\mathbb{C}\right)'$ is the operator associated to D. A Green's function G(x,y) is a function : G$\in C\left(M \times M;\mathbb{C}\right)$ such that : $\forall y \in M : U(y) = T\left(G\left(., y\right)\right)$ with a map $T : F \to C_{\infty c}\left(M;\mathbb{C}\right)'$.

2. If M=an open O of $\mathbb{R}^m$ and S a compactly supported distribution in $C_{\infty c}\left(M;\mathbb{C}\right)'$ then $U\left(0\right) * S$ is a solution of D'V=S
For any compactly supported function f, $u = U\left(y\right)_t\left(f\left(x + y - t\right)\right)$ is a solution of Du = f

3. Notice that a fundamental solution is related to a scalar linear differential operator on distributions, and not to a PDE. Thus a fundamental solution or a Green function may or may not meet coercive conditions.

4. If D' is a linear scalar differential operator on the space of tempered distributions S($\mathbb{R}^m$)' then the Fourier transform $\widehat{U}$ of a fundamental solution is such that : $P(x, it)\widehat{U} = \left(2\pi\right)^{-m/2} T\left(1\right)$ where P is the symbol of D'.

5. If D is a scalar linear differential operator on functions on $\mathbb{R}^m$ with constant coefficients then this equation reads: $\left(\sum_{s=0}^{r}\sum_{\alpha_1...\alpha_s=1}^{m} A^{\alpha_1...\alpha_s}(-i)^s t^\alpha\right)\widehat{U} = \left(2\pi\right)^{-m/2}$ from where $\widehat{U}$ and U can be computed by inverse Fourier transform.

General linear elliptic boundary problems

If the possible solutions of a linear PDE must meet some inequalities, then there is a solution.

(Taylor 1 p.380-395) The PDE is : find $X \in \mathfrak{X}_r(E_1)$ such that :

$DX = Y_0$ on $\overset{\circ}{M}$

$D_j X = Y_j, j = 1...N$ on ∂M

where :

E_1, E_2 are vector bundles on the same smooth compact manifold M with boundary

F_j are vector bundles on ∂M

$D : \mathfrak{X}_r(E_1) \to \mathfrak{X}(E_2)$ is a weakly elliptic r order linear differential operator

$D_j : \mathfrak{X}_r(E_1) \to \mathfrak{X}(F_j), j = 1..N$ are r_j order linear differential operators

The problem is said to be regular if for any solution :

$\exists s \in \mathbb{R}, \exists C \in \mathbb{R} : \|X\|^2_{H^{s+r}(E_1)} \le C \left(\|DX\|^2_{H^r(E_2)} + \sum_j \|D_j X\|^2_{H^{r+s-r_j-1/2}(F_j)} + \|X\|^2_{H^{r+s-1}(E_1)} \right)$

If the problem is regular elliptic, then for k$\in \mathbb{N}$ the map :

$\phi : H^{r+k}(E_1) \to H^k(E_1) \oplus \oplus_{j=1}^N H^{r+k-r_j-1/2}(F_j)$

defined by :

$\phi(X) = DX$ on $\overset{\circ}{M}$ and $\phi(X) = D_j X = Y_j, j = 1...N$ on ∂M

is Fredholm. ϕ has a finite dimensional kernel, closed range, and its range has finite codimension. It has a right Fredholm inverse.

So the problem has a solution, which is given by the inverse of ϕ.

As a special case we have the following

Theorem 2467 *(Taylor 1 p.393) M is a riemannian manifold with boundary, n is the unitary normal outward oriented to ∂M.*

The PDE find $u \in \Lambda_r(M; \mathbb{C})$ such that $\Delta u = f$ on $\overset{0}{M}$ and
Problem 1 : $n \wedge u = g_0, n \wedge \delta u = g_1$ on ∂M
Problem 2 : $i_n u = g_0, i_n du = g_1$ on ∂M
are both regular and have a solution.

Hyperbolic PDE

A PDE is said to be hyperbolic at a point u if it is an evolution equation with Cauchy conditions such that in a neighborhood of p, there is a unique solution for any initial conditions.

Theorem 2468 *(Taylor 1 p.435) The PDE is to find $u \in C(M; \mathbb{C})$ such that :*
$Du = f$ on M
$u(x)=g_0$ on S_0,
$Yu = g_t$ on S_t
where
M is a m+1 dimensional manifold endowed with a Lorentz metric of signature (m-,1+), folliated by compact space like hypersurfaces S_t.
D is the scalar operator $D = \square + P$ with a first order differential operator P
Y is a vector field transverse to the S_t

If $f \in H^{k-1}(M), g_0 \in H^k(S_0), g_t \in H^{k-1}(S_t), k \in \mathbb{N}, k > 0$ then there is a unique solution $u \in H^k(M)$ which belongs to $H^1(\Omega), \Omega$ being the region swept by $S_t, t \in [0,T]$

31.2.3 Poisson like equations

They are PDE with the scalar laplacian on a riemannian manifold as operator. Fundamental solutions for the laplacian are given through the Green's function denoted here G, itself built from eigen vectors of $-\Delta$ (see Differential operators)

Poisson equation

The problem is to find a function u on a manifold M such that : $-\Delta u = f$ where f is a given function. If f=0 then it is called the Laplace equation (and the problem is then to find harmonic functions).

This equation is used in physics whenever a force field is defined by a potential depending on the distance from the sources, such that the electric or the gravitational field (if the charges move or change then the wave operator is required).

1. Existence of solutions :

Theorem 2469 *(Gregor'yan p.45) If O is a relatively compact open in a riemannian manifold M such that $M \backslash \overline{O} \neq \varnothing$ then the Green function G on M is finite and : $\forall f \in L^2(M, \varpi_0, \mathbb{R}), u(x) = \int_M G(x,y) f(y) \varpi_0(y)$ is the unique solution of $-\Delta u = f$*

The solutions are not necessarily continuous or differentiable (in the usual sense of functions). Several cases arise (see Lieb p.262). However :

Theorem 2470 *(Lieb p.159) If O is an open subset of $\mathbb{R}^m$ and $f \in L^1_{loc}(O, dx, \mathbb{C})$ then $u(x) = \int_O G(x,y) f(y) dy$ is such that $-\Delta u = f$ and $u \in L^1_{loc}(O, dx, \mathbb{C})$. Moreover for almost every x : $\partial_\alpha u = \int_O \partial_\alpha G(x,y) f(y) dy$ where the derivative is in the sense of distribution if needed. If $f \in L^p_c(O, dx, \mathbb{C})$ with $p > m$ then u is differentiable. The solution is given up to a harmonic function : $\Delta u = 0$*

Theorem 2471 *(Taylor 1 p.210) If $S \in S(\mathbb{R}^m)'$ is such that $\Delta S = 0$ then S=T(f) with f a polynomial in $\mathbb{R}^m$*

2. Newtons's theorem : in short it states that a spherically symmetric distribution of charges can be replaced by a single charge at its center.

Theorem 2472 *If μ_+, μ_- are positive Borel measure on $\mathbb{R}^m, \mu = \mu_+ - \mu_-, \nu = \mu_+ + \mu_-$ such that $\int_{\mathbb{R}^m} \phi_m(y) \nu(y) < \infty$ then $V(x) = \int_{\mathbb{R}^m} G(x,y) \mu(y) \in L^1_{loc}(\mathbb{R}^m, dx, \mathbb{R})$*
If μ is spherically symmetric (meaning that $\mu(A) = \mu(\rho(A))$ for any rotation ρ) then : $|V(x)| \leq |G(0,x)| \int_{\mathbb{R}^m} \nu(y)$
If for a closed ball B(0,r) centered in 0 and with radius r $\forall A \subset \mathbb{R}^m$: $A \cap B(0,r) = \varnothing \Rightarrow \mu(A) = 0$ then : $\forall x : \|x\| > r : V(x) = G(0,x) \int_{\mathbb{R}^m} \nu(y)$

The functions ϕ_m are :

m>2 : $\phi_m(y) = (1 + \|y\|)^{2-m}$

m=2 : $\phi_m(y) = \ln(1 + \|y\|)$

m=1 : $\phi_m(y) = \|y\|$

Dirichlet problem

Theorem 2473 *(Taylor 1 p.308) The PDE : find $u \in C(M; \mathbb{C})$ such that :*

$\Delta u = 0$ on $\overset{\circ}{M}$

$u = f$ on ∂M

where

M is a riemannian compact manifold with boundary.

$f \in C_\infty(\partial M; \mathbb{C})$

has a unique solution u=PI(f) and the map PI has a unique extension : PI

: $H^s(\partial M) \to H^{s+\frac{1}{2}}\left(\overset{\circ}{M}\right)$.

This problem is equivalent to the following :

find v : $\Delta v = -\Delta F$ on $\overset{\circ}{M}, v = 0$ on ∂M where $F \in C_\infty(M; \mathbb{C})$ is any function such that : $F = f$ on ∂M

Then there is a unique solution v with compact support in $\overset{\circ}{M}$ (and null on the boundary) given by the inverse of Δ and u=F + v.

If M is the unit ball in $\mathbb{R}^m$ with boundary S^{m-1} the map $PI : H^s(S^{m-1}) \to H^{s+\frac{1}{2}}(B)$ for s $\geq 1/2$ is :

$PI(f)(x) = u(x) = \frac{1-\|x\|^2}{A(S^{m-1})} \int_{S^{m-1}} \frac{f(y)}{\|x-y\|^m} d\sigma(y)$ with $A(S^{m-1}) = \frac{2\pi^{m/2}}{\Gamma\left(\frac{m}{2}\right)}$

Neumann problem

Theorem 2474 *(Taylor 1 p.350) The PDE : find $u \in C(M; \mathbb{C})$ such that*

$\Delta u = f$ on $\overset{\circ}{M}$,

$\frac{\partial u}{\partial n} = 0$ on ∂M

where

M is a riemannian compact manifold with smooth boundary,

$f \in L^2(M, \varpi_0, \mathbb{C})$

has a solution $u \in H^2\left(\overset{\circ}{M}\right)$ iff $\int_M f\varpi_0 = 0$. Then the solution is unique up

to an additive constant and belongs to $H^{r+2}\left(\overset{\circ}{M}\right)$ if $f \in H^r\left(\overset{\circ}{M}\right), r \geq 0$

31.2.4 Equations with the operator $-\Delta + P$

They are equations with the scalar operator $D = -\Delta + P$ where P is a first order differential operator, which can be a constant scalar, on a riemannian manifold.

The PDE : find $u : -\Delta u = \lambda u$ in $\overset{\circ}{M}$, $u = g$ on ∂M where λ is a constant scalar comes to find eigenvectors e such that $e = g$ on ∂M. There are solutions only if λ is one of the eigenvalues (which depend only on M). Then the eigenvectors are continuous on M and we have the condition : $e_n(x) = g_0(x)$ on ∂M.

Theorem 2475 *(Taylor 1 p.304) The differential operator : $D = -\Delta + P$ ::*
$$H^1_c\left(\overset{\circ}{M}\right) \to H^{-1}\left(\overset{\circ}{M}\right)$$
on a smooth compact manifold M with boundary in a riemannian manifold N,

with a smooth first order differential operator P with smooth coefficients is Fredholm of index 0. It is surjective iff it is injective.

A solution of the PDE : find $u \in H^1_c\left(\overset{\circ}{M}\right)$ such that Du=f on $\overset{\circ}{M}$ with

$f \in H^{k-1}\left(\overset{\circ}{M}\right), k \in \mathbb{N}$, belongs to $H^{k+1}\left(\overset{\circ}{M}\right)$

Theorem 2476 *(Zuily p.93) The differential operator : $D = -\Delta + \lambda$ where $\lambda \geq 0$ is a fixed scalar, is an isomorphism $H^1_c(O) \to H^{-1}(O)$ on an open subset O of $\mathbb{R}^m$, it is bounded if $\lambda = 0$.*

Theorem 2477 *(Zuily p.149) The differential operator : $D = -\Delta + \lambda$ where $\lambda \geq 0$ is a fixed scalar,*

is, $\forall k \in \mathbb{N}$, an isomorphism $H^{k+2}(\overset{\circ}{M}) \cap H^1_c(\overset{\circ}{M}) \to H^k(\overset{\circ}{M})$,

and an isomorphism $\left(\cap_{k \in \mathbb{N}} H^k(\overset{\circ}{M})\right) \cap H^1_c(\overset{\circ}{M}) \to \cap_{k \in \mathbb{N}} H^k(\overset{\circ}{M})$

where M is a smooth manifold with boundary of $\mathbb{R}^m$, compact if $\lambda = 0$

31.2.5 Helmoltz equation

Also called "scattering problem". The differential operator is $\left(-\Delta + k^2\right)$ where k is a real scalar

Green's function

Theorem 2478 *(Lieb p.166) In $\mathbb{R}^m$ the fundamental solution of $\left(-\Delta + k^2\right) U(y) = \delta_y$ is $U(y) = T(G(x,y))$ where the Green's function G is given for $m \geq 1, k > 0$ by: $G(x,y) = \int_0^\infty (4\pi\zeta)^{-m/2} \exp\left(-\frac{\|x-y\|^2}{4\zeta} - k^2\zeta\right) d\zeta$*

G is called the "Yukawa potential"

G is symmetric decreasing, > 0 for $x \neq y$

$\int_{\mathbb{R}^m} \int_0^\infty (4\pi\zeta)^{-m/2} \exp\left(-\frac{\|\xi\|^2}{4\zeta} - k^2\zeta\right) d\zeta d\xi = k^{-2}$

when $\xi \to 0$: $\int_0^\infty (4\pi\zeta)^{-m/2} \exp\left(-\frac{\|\xi\|^2}{4\zeta} - k^2\zeta\right) d\zeta \to 1/2k$ for m=1, and

$\sim \frac{1}{(2-m)A(S_{m-1})} \|\xi\|^{2-m}$ for m>1

when $\xi \to \infty$: $-\ln\left(\int_0^\infty (4\pi\zeta)^{-m/2} \exp\left(-\frac{\|\xi\|^2}{4\zeta} - k^2\zeta\right) d\zeta\right) \sim k\|\xi\|$

The Fourier transform of $\int_0^\infty (4\pi\zeta)^{-m/2} \exp\left(-\frac{\|\xi\|^2}{4\zeta} - k^2\zeta\right) d\zeta$ is : $(2\pi)^{-m/2}\left(\|t\|^2 + k^2\right)^{-1}$

General problem

Theorem 2479 *If* $f \in L^p(\mathbb{R}^m; dx, \mathbb{C}), 1 \le p \le \infty$, *then* $u(x) = \int_{\mathbb{R}^m} G(x,y) f(y) dy$ *is the unique solution of* $\left(-\Delta + k^2\right) u = f$ *such that* $u(x) \in L^r(\mathbb{R}^m; dx, \mathbb{C})$ *for some r.*

Theorem 2480 *(Lieb p.257) If* $f \in L^p(\mathbb{R}^m; dx, \mathbb{C}), 1 \le p \le \infty$ *is such that :* $\left(-\Delta + k^2\right) T(f) = 0$ *then f=0*

Dirichlet problem

This is the "scattering problem" proper.

Theorem 2481 *(Taylor 2 p.147) The PDE is to find a function* $u \in C\left(\mathbb{R}^3; \mathbb{C}\right)$ *such that :*

$-\left(\Delta + k^2\right) u = 0$ *in O with a scalar $k > 0$*

$u = f$ *on ∂K*

$\|ru(x)\| < C, r\left(\frac{\partial u}{\partial r} - iku\right) \to 0$ *when $r = \|x\| \to \infty$*

where K is a compact connected smooth manifold with boundary in $\mathbb{R}^3$ with complement the open O

i) if f=0 then the only solution is u=0

ii) if $f \in H^s(\partial K)$ there is a unique solution u in $H_{loc}^{s+\frac{1}{2}}(O)$

31.2.6 Wave equation

In physics, the mathematical model for a force field depending on the distance to the sources is no longer the Poisson equation when the sources move or the charges change, but the wave equation, to account for the propagation of the field.

Wave operator

On a manifold endowed with a non degenerate metric g of signature $(+1,-p)$ and a folliation in space like hypersurfaces S_t, p dimensional manifolds endowed with a riemannian metric, the **wave operator** is the d'Alembertian : $\Box u = \frac{\partial^2 u}{\partial t^2} - \Delta_x$ acting on families of functions $u \in C(\mathbb{R}; C(S_t; \mathbb{C}))$. So $\Box$ is split in

Δ_x and a "time component" which can be treated as $\frac{\partial^2}{\partial t^2}$, and the functions are then $\varphi(t,x) \in C(\mathbb{R}; C(S_t; \mathbb{C}))$. The operator is the same for functions or distributions : $\square' = \square$

Fundamental solution of the wave operator in $\mathbb{R}^m$

The **wave operator** is the d'Alembertian : $\square u = \frac{\partial^2 u}{\partial t^2} - \sum_{\alpha=1}^{m} \frac{\partial^2 u}{\partial x_\alpha^2}$ acting on families of functions $u \in C(\mathbb{R}; C(\mathbb{R}^m; \mathbb{C}))$ or distributions $u \in C(\mathbb{R}; C(\mathbb{R}^m; \mathbb{C})')$. The operator is the same for functions or distributions : $\square' = \square$. It is symmetric with respect to the inversion of t : $t \to -t$.

Theorem 2482 *The fundamental solution of the wave operator $\square u = \frac{\partial^2 u}{\partial t^2} - \sum_{\alpha=1}^{m} \frac{\partial^2 u}{\partial x_\alpha^2}$ acting on families of functions $u \in C(\mathbb{R}; C(\mathbb{R}^m; \mathbb{C}))$ is the distribution : $U \in C(\mathbb{R}; S'(\mathbb{R}^m))$:*

$$U(\varphi(t,x)) = (2\pi)^{-m/2} \int_0^\infty \left(\int_{R^m} e^{i\xi x} \frac{\sin(t\|\xi\|)}{\|\xi\|} \varphi(t,x)\, d\xi \right) dt$$

Proof. It is obtained from a family of distribution through Fourier transform

$$\mathcal{F}_x \square U(t) = \mathcal{F}_x \frac{\partial^2}{\partial t^2} U - \sum_k \mathcal{F}_x \left(\frac{\partial}{\partial x_k} \right)^2 U = \left(\frac{\partial^2}{\partial t^2} \mathcal{F}_x U \right) - (-i)^2 \sum_k (\xi_k)^2 \mathcal{F}_x(U) =$$

$$\left(\frac{\partial^2}{\partial t^2} + \sum_k (\xi_k)^2 \right)(\mathcal{F}_x U)$$

If $\square U = \delta_0(t,x) = \delta_0(t) \otimes \delta_0(x)$

then $\mathcal{F}_x U = \mathcal{F}_x(\delta_0(t) \otimes \delta_0(x)) = \mathcal{F}_x(\delta_0(t)) \otimes \mathcal{F}_x(\delta_0(x)) = \delta_0(t) \otimes [1]_\xi$

and $\mathcal{F}_x(\delta_0(t)) = \delta_0(t)$ because : $\mathcal{F}_x(\delta_0(t))(\varphi(t,x)) = \delta_0(t)(\mathcal{F}_x\varphi(t,x)) = \mathcal{F}_x(\varphi(0,x))$

Thus we have the equation : $\left(\frac{\partial^2}{\partial t^2} + \|\xi\|^2 \right)(\mathcal{F}_x U) = \delta_0(t) \otimes [1]_\xi$

$\mathcal{F}_x U = u(t,\xi)$. For $t \neq 0$ the solutions of the ODE $\left(\frac{\partial^2}{\partial t^2} + \|\xi\|^2 \right) u(t,\xi) = 0$ are :

$u(t,\xi) = a(\|\xi\|) \cos t\|\xi\| + b(\|\xi\|) \sin t\|\xi\|$

So for $t \in \mathbb{R}$ we take :

$u(t,\xi) = H(t)(a(\|\xi\|) \cos t\|\xi\| + b(\|\xi\|) \sin t\|\xi\|)$ with the Heavyside function H(t)=1 for t$\geq$ 0

and we get :

$\left(\frac{\partial^2}{\partial t^2} + \|\xi\|^2 \right) u(t,\xi) = \delta_0(t) \otimes \|\xi\| b(\|\xi\|) + \frac{d}{dt}\delta_0(t) \otimes a(\|\xi\|) = \delta_0(t) \otimes [1]_\xi$

$\Rightarrow \mathcal{F}_x U = H(t) \frac{\sin(t\|\xi\|)}{\|\xi\|}$

$U(t) = \mathcal{F}_x^* \left(H(t) \frac{\sin(t\|\xi\|)}{\|\xi\|} \right) = (2\pi)^{-m/2} \int_{R^m} e^{i\xi x} H(t) \frac{\sin(t\|\xi\|)}{\|\xi\|} d\xi$ ∎

The fundamental solution has the following properties (Zuily) :

U(t,x) = 0 for $\|x\| > t$ which is interpreted as propagation at speed 1

If $m > 2$ and m odd then $Supp U(t,.) \subset \{\|x\| = t\}$ so we have a "ligth cone"

If m=3 then U can be expressed through the the Lebesgue measure σ on the unique sphere S^2

$U : U(\varphi) = \int_0^\infty \left(\frac{t}{4\pi} \int_{S^2} \psi_t(s)\sigma\right) dt$ where $\psi_t(s) = \varphi(tx)\big|_{\|x\|=1}$

$V(t)(\varphi) = \frac{t}{4\pi} \int_{S^2} \psi_t(s)\sigma, V \in C_\infty\left(\mathbb{R}; C_\infty\left(\mathbb{R}^3; \mathbb{C}\right)'\right), V(0) = 0, \frac{dV}{dt} = \delta_0, \frac{d^2V}{dt^2} = 0,$

(Taylor 1 p.222) As $-\Delta$ is a positive operator in the Hilbert space $H^1(\mathbb{R}^m)$ it has a square root $\sqrt{-\Delta}$ with the following properties :

$U(t, x) = \left(\sqrt{-\Delta}\right)^{-1} \circ \sin\sqrt{-\Delta} \circ \delta(x)$

$\frac{\partial U}{\partial t} = \cos t\sqrt{-\Delta} \circ \delta(x)$

If $f \in S(\mathbb{R}^m)$: with $r = \|x\|$

$f\left(\sqrt{-\Delta}\right)\delta(x) = \frac{1}{\sqrt{2\pi}} \left[-\frac{1}{2\pi r}\frac{\partial}{\partial r}\right]^k \widehat{f}(t)$ if $m = 2k + 1$

$f\left(\sqrt{-\Delta}\right)\delta(x) = \frac{1}{\sqrt{\pi}} \int_t^\infty \left[-\frac{1}{2\pi s}\frac{\partial}{\partial s}\widehat{f}(s)\right]^k \frac{s}{\sqrt{s^- - r^2}} ds$ if $m = 2k$

Cauchy problem on a manifold

Theorem 2483 *(Taylor 1 p.423) The PDE : find $u \in C(\mathbb{R}; C(M; \mathbb{C}))$ such that :*

$\Box u = \frac{\partial^2 u}{\partial t^2} - \Delta_x u = 0$

$u(0,x)=f(x)$

$\frac{\partial u}{\partial t}(0, x) = g(x)$

where M is a geodesically complete riemannian manifold (without boundary).
$f \in H_c^1(M), g \in L_c^2(M, \varpi_0, \mathbb{C})$ have non disjointed support
has a unique solution u, $u \in C\left(\mathbb{R}; H^1(M)\right) \cap C\left(\mathbb{R}; L^2(M, \varpi_0, \mathbb{C})\right)$ and has a compact support in M for all t.

Cauchy problem in $\mathbb{R}^m$

Theorem 2484 *(Zuily p.170) The PDE : find $u \in C(I; C(\mathbb{R}^m; \mathbb{C}))$:*

$\Box u = \frac{\partial^2 u}{\partial t^2} - \sum_{\alpha=1}^m \frac{\partial^2 u}{\partial x_\alpha^2} = 0$ *in $I \times O$*

$u(t_0, x) = f(x),$

$\frac{\partial u}{\partial t}(t_0, x) = g(x),$

where
O is a bounded open subset of $\mathbb{R}^m$,
I an interval in $\mathbb{R}$ with $t_0 \in I$
$f \in H_c^1(O), g \in L^2(O, dx, \mathbb{C})$
has a unique solution : $u(t,x) = U(t) * g(x) + \frac{dU}{dt} * f(x)$ and it belongs to $C\left(I; H_c^1(O)\right)$.

$u(t,x) = \sum_{k=1}^\infty \{\langle e_k, f\rangle \cos(\sqrt{\lambda_k}(t - t_0)) + \frac{\langle e_k, g\rangle}{\sqrt{\lambda_k}} \sin(\sqrt{\lambda_k}(t - t_0))\} e_k(x)$

where $(\lambda_n, e_n)_{n\in\mathbb{N}}$ are eigen values and eigen vectors of the -laplacian $-\Delta$, the e_n being chosen to be a Hilbertian basis of $L^2(O, dx, \mathbb{C})$

Theorem 2485 *(Taylor 1 p.220) The PDE : find $u \in C\left(\mathbb{R}; C\left(\mathbb{R}^m; \mathbb{C}\right)'\right)$:*

$\Box u = \dfrac{\partial^2 u}{\partial t^2} - \sum_{\alpha=1}^m \frac{\partial^2 u}{\partial x_\alpha^2} = 0$ *in* $\mathbb{R} \times \mathbb{R}^m$

$u\left(0, x\right) = f\left(x\right)$

$\frac{\partial u}{\partial t}\left(0, x\right) = g\left(x\right)$

where $f, g \in S\left(\mathbb{R}^m\right)'$

has a unique solution $u \in C_\infty\left(\mathbb{R}; S\left(\mathbb{R}^m\right)'\right)$. *It reads :* $u(t, x) = U\left(t\right) * g\left(x\right) + \frac{dU}{dt} * f\left(x\right)$

Cauchy problem in $\mathbb{R}^4$

(Zuily)

1. Homogeneous problem:

Theorem 2486 *The PDE: find* $u \in C\left(\mathbb{R}; C\left(\mathbb{R}^m; \mathbb{C}\right)'\right)$ *such that :*

$\Box u = \dfrac{\partial^2 u}{\partial t^2} - \sum_{\alpha=1}^3 \frac{\partial^2 u}{\partial x_\alpha^2} = 0$ *in* $\mathbb{R}_+ \times \mathbb{R}^3$

$u\left(0, x\right) = f\left(x\right)$

$\frac{\partial u}{\partial t}\left(0, x\right) = g\left(x\right)$

has a unique solution : $u(t, x) = U\left(t\right) * g\left(x\right) + \frac{dU}{dt} * f\left(x\right)$ *which reads :*

$u\left(t, x\right) = \frac{1}{4\pi t} \int_{\|y\|=t} g\left(x - y\right) \sigma_{ty} + \frac{\partial}{\partial t}\left(\frac{1}{4\pi t} \int_{\|y\|=t} f\left(x - y\right) \sigma_{ty}\right)$

2. Properties of the solution :
If $f, g \in C_\infty\left(\mathbb{R}^3; \mathbb{C}\right)$ then $u(t,x) \in C_\infty\left(\mathbb{R}^4; \mathbb{C}\right)$

When $t \to \infty$ $u(t,x) \to 0$ and $\exists M > 0 : t \geq 1 : |u(t, x)| \leq \dfrac{M}{t}$

The quantity (the "energy") is constant : $W(t) = \int \{\left(\dfrac{\partial u}{\partial t}\right)^2 + \sum_{i=1}^3 \left(\dfrac{\partial u}{\partial x_i}\right)^2\} dx = Ct$

Propagation at the speed 1 :
If $\|x\| > R \Rightarrow f\left(x\right) = g\left(x\right) = 0$ then : $u(t, x) = 0$ for $\|x\| \leq R + t$ or $\{t > R, \|x\| \leq t - R\}$

The value of u(t,x) in (t_0, x_0) depends only on the values of f and g on the hypersuface $\|x - x^\circ\| = t^\circ$

Plane waves :
if $f(x) = -k.x = -\sum_{l=1}^3 k_l x^l$ with k fixed, $g\left(x\right) = 1$ then : $u\left(t, x\right) = t - k.x$

3. Inhomogeneous problem :

Theorem 2487 *The PDE : find* $u \in C\left(\mathbb{R}_+; C\left(\mathbb{R}^3; \mathbb{C}\right)\right)$ *such that :*

$\Box u = \dfrac{\partial^2 u}{\partial t^2} - \sum_{\alpha=1}^3 \frac{\partial^2 u}{\partial x_\alpha^2} = F$ *in* $\mathbb{R}_+ \times \mathbb{R}^3$

$u\left(0, x\right) = f\left(x\right)$

$\frac{\partial u}{\partial t}\left(0, x\right) = g\left(x\right)$

where $F \in C_\infty\left(\mathbb{R}_+; C\left(\mathbb{R}^3; \mathbb{C}\right)\right)$

has a unique solution : $u(t, x) = v = u + U * F$ *,where U is the solution of the homogeneous problem, which reads :*

$t \geq 0 : u(t, x) = \frac{1}{4\pi t} \int_{\|y\|=t} g(x - y) \sigma_{ty} + \frac{\partial}{\partial t} \left(\frac{1}{4\pi t} \int_{\|y\|=t} f(x - y) \sigma_{ty} \right) +$
$\int_{t=0}^{\infty} \frac{1}{4\pi s} \left(\int_{\|x\|=s} F(t - s, x - y) \sigma_{yt} \right) ds$

31.2.7 Schrödinger operator

It is of course linked to the celebrated Schrödinger equation of quantum mechanics.

Definition

This is the scalar operator $D = \frac{\partial}{\partial t} - i\Delta_x$ acting on functions $f \in V \subset C(J \times M; \mathbb{C})$ where J is some interval in $\mathbb{R}$ and M a manifold upon which the laplacian is defined. Usually $M = \mathbb{R}^m$ and $V = S(\mathbb{R} \times \mathbb{R}^m)$.

If $V = C(J; F)$, where F is some Fréchet space of functions on M, f can be seen as a map : $f : J \to F$. If there is a family of distributions : $S : J \to F'$ then $\widetilde{S}(f) = \int_J S(t)_x (f(t, x)) dt$ defines a distribution $\widetilde{S} \in V'$

This is the basis of the search for fundamental solutions.

Cauchy problem in $\mathbb{R}^m$

From Zuily p.152

Theorem 2488 *If $S \in S(\mathbb{R}^m)'$ there is a unique family of distributions $u \in C_\infty \left(\mathbb{R}; S(\mathbb{R}^m)' \right)$ such that :*

$D\widetilde{u} = 0$

$u(0) = S$

where $\widetilde{u} : \forall \varphi \in S(\mathbb{R} \times \mathbb{R}^m) : \widetilde{u}(\varphi) = \int_{\mathbb{R}} u(t)(\varphi(t, .)) dt$

which is given by : $\widetilde{u}(\varphi) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}^m} \mathcal{F}_\xi^* \left(e^{-it\|\xi\|^2} \widehat{S}(\varphi) \right) dx \right) dt = \int_{\mathbb{R}} \int_{\mathbb{R}^m} \phi(t, x) \varphi(t, x) dx dt$

Which needs some explanations...

Let $\varphi \in S(\mathbb{R} \times \mathbb{R}^m)$ .The Fourier transform $\widehat{S}$ of S is a distribution $\widehat{S} \in S(\mathbb{R}^m)'$ such that :

$\widehat{S}(\varphi) = S(\mathcal{F}_y(\varphi)) = S_\zeta \left((2\pi)^{-m/2} \int_{\mathbb{R}^m} e^{-i\langle y, \zeta \rangle} \varphi(t, y) dy \right)$ (S acts on the ζ function)

This is a function of t, so $e^{-it\|\xi\|^2} \widehat{S}(\varphi)$ is a function of t and $\xi \in \mathbb{R}^m$

$e^{-it\|\xi\|^2} \widehat{S}(\varphi) = e^{-it\|\xi\|^2} S_\zeta \left((2\pi)^{-m/2} \int_{\mathbb{R}^m} e^{-i\langle y, \zeta \rangle} \varphi(t, y) dy \right)$

$= S_\zeta \left((2\pi)^{-m/2} \int_{\mathbb{R}^m} e^{-it\|\xi^2\| - i\langle y., \zeta \rangle} \varphi(t, y) dy \right)$

Its inverse Fourier transform $\mathcal{F}_\xi^* \left(e^{-it\|\xi\|^2} \widehat{S}(\varphi) \right)$ is a function of t and $x \in \mathbb{R}^m$ (through the interchange of ξ, x)

$\mathcal{F}_\xi^* \left(e^{-it\|\xi\|^2} \widehat{S}(\varphi) \right) = (2\pi)^{-m/2} \int_{\mathbb{R}^m} e^{i\langle \xi, x \rangle} S_\zeta \left((2\pi)^{-m/2} \int_{\mathbb{R}^m} e^{-it\|\xi^2\| - i\langle y., \zeta \rangle} \varphi(t, y) dy \right) d\xi$

$= S_\zeta \left((2\pi)^{-m} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} e^{-it\|\xi^2\| + i\langle x, \xi \rangle - i\langle y., \zeta \rangle} \varphi(t, y) dy d\xi \right)$

and the integral with respect both to x and t gives :

$\widetilde{u}(\varphi) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}^m} S_\zeta \left((2\pi)^{-m} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} e^{-it\|\xi^2\| + i\langle x, \xi \rangle - i\langle y., \zeta \rangle} \varphi(t, y) dy d\xi \right) dx \right) dt$

By exchange of x and y :

$\widetilde{u}(\varphi) = \int_{\mathbb{R}} \int_{\mathbb{R}^m} S_\zeta \left((2\pi)^{-m} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} e^{-it\|\xi^2\|+i\langle y,\xi\rangle - i\langle x.,\zeta\rangle} dy d\xi \right) \varphi(t,x) \, dx dt$

$\widetilde{u}(\varphi) = \int_{\mathbb{R}} \int_{\mathbb{R}^m} \phi(t,x) \varphi(t,x) \, dx dt$

where $\phi(t,x) = (2\pi)^{-m} S_\zeta \left(e^{-i\langle x.,\zeta\rangle} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} e^{-it\|\xi^2\|+i\langle y,\xi\rangle} d\xi dy \right)$

If S=T(g) with g$\in S(\mathbb{R}^m)$ then : $\phi(t,x) = (2\pi)^{-m} \int_{\mathbb{R}^m} e^{-it\|\xi^2\|+i\langle y,\xi\rangle+i\langle x.,\xi\rangle} \widehat{g}(\xi) \, d\xi \in C_\infty(\mathbb{R}; S(\mathbb{R}^m))$

If $S \in H^s(\mathbb{R}^m), s \in \mathbb{R}$ then $\phi(t,x) \in C_0(\mathbb{R}; H^s(\mathbb{R}^m))$

If $S \in H^k(\mathbb{R}^m)$ k$\in \mathbb{N}$ then $\|\phi(t,.)\|_{H^k} = \|g\|_{H^k}$; $\frac{\partial^r \phi}{\partial x_{\alpha_1}..\partial x_{\alpha_r}} \in C_0(\mathbb{R}; H^{s-2k}(\mathbb{R}^m))$;

If S=T(g),g$\in L^2(\mathbb{R}^m, dx, \mathbb{C}), \forall r > 0 : \|x\|^r g \in L^2(\mathbb{R}^m, dx, \mathbb{C})$ then $\forall t \neq 0 :$ $\phi(t,.) \in C_\infty(\mathbb{R}^m; \mathbb{C})$

If S=T(g),g$\in L^1(\mathbb{R}^m, dx, \mathbb{C})$, then $\forall t \neq 0 : \|\phi(t,.)\|_\infty \leq (4\pi|t|)^{-m/2} \|g\|_1$

31.2.8   Heat equation

The operator is the heat operator : $D = \frac{\partial}{\partial t} - \Delta_x$ on $C(\mathbb{R}_+; C(M;\mathbb{C}))$

Cauchy problem

M is a riemannian manifold, the problem is to find $u \in C(\mathbb{R}_+; C(M;\mathbb{C}))$ such that $\frac{\partial u}{\partial t} - \Delta_x u = 0$ on M and u(0,x)=g(x) where g is a given function.

Theorem 2489 *(Gregor'yan p.10) The PDE : find $u \in C_\infty(\mathbb{R}_+ \times M; \mathbb{C})$ such that :*

$\frac{\partial u}{\partial t} = \Delta_x u$

$u(t,.) \to g$ *in* $L^2(M, \varpi_0, \mathbb{C})$ *when* $t \to 0_+$

where

M is a riemannian smooth manifold (without boundary)

$g \in L^2(M, \varpi_0, \mathbb{C})$

has a unique solution

If M is an open, relatively compact, in a manifold N, then $u \in C_0(\mathbb{R}_+; H^1(M))$ and is given by $u(t,x) = \sum_{n=1}^\infty \langle e_n, \widehat{g} \rangle e^{-\lambda_n t} e_n(x)$ where (e_n, λ_n) are the eigen vectors and eigen values of $-\Delta$, the e_n being chosen to be a Hilbertian basis of $L^2(M, dx, \mathbb{C})$.

Dirichlet problem

Theorem 2490 *(Taylor 1 p.416) The PDE : find $u \in C(\mathbb{R}_+; C(N;\mathbb{C}))$ such that :*

$\frac{\partial u}{\partial t} - \Delta_x u = 0$ *on* $\overset{\circ}{N}$

$u(0,x) = g(x)$

$u(t,x) = 0$ *if* $x \in \partial N$

where

M is a compact smooth manifold with boundary in a riemannian manifold N

$g \in L^2(N, dx, \mathbb{C})$

has a unique solution, and it belongs to $C\left(\mathbb{R}_+; H^{-s}\left(\overset{\circ}{M}\right)\right) \cap C_1\left(\mathbb{R}_+; H^{-s+2}\left(\overset{\circ}{M}\right)\right).$

It reads : $u(t, x) = \sum_{n=1}^{\infty} \langle e_n, \widehat{g} \rangle e^{-\lambda_n t} e_n(x)$ *where* $(\lambda_n, e_n)_{n \in \mathbb{N}}$ *are eigen values and eigen vectors of* $-\Delta$ *,- the laplacian, on* $\overset{\circ}{N}$*, the* e_n *being chosen to be a Hilbertian basis of* $L^2\left(\overset{\circ}{M}, dx, \mathbb{C}\right)$

31.2.9 Non linear partial differential equations

There are few general results, and the diversity of non linear PDE is such that it would be impossible to give even a hint of the subject. So we will limit ourselves to some basic definitions.

Cauchy-Kowalesky theorem

There is an extension of the theorem to non linear PDE.

Theorem 2491 *(Taylor 3 p.445) The scalar PDE : find* $u \in C(\mathbb{R} \times O; \mathbb{C})$
$\frac{\partial^r u}{\partial t^r} = D\left(t, x, u, \frac{\partial^s u}{\partial x_{\alpha_1} \dots \partial x_{\alpha_s}}, \frac{\partial^{s+k} u}{\partial x_{\alpha_1} \dots \partial x_{\alpha_s} \partial t^k}\right)$ *where s=1...r,k=1..r-1,*$\alpha_j = 1...m$
$\frac{\partial^k u}{\partial t^k}(0, x) = g_k(x), k = 0, ... r - 1$
where O is an open in $\mathbb{R}^m$:

If D, g_k *are real analytic for* $x_0 \in O$*, then there is a unique real analytic solution in a neighborhood of* $(0, x_0)$

Linearization of a PDE

The main piece of a PDE is a usually differential operator : $D : J^r E_1 \to E_2$. If D is regular, meaning at least differentiable, in a neighborhood $n(Z_0)$ of a point $Z_0 \in J^r E_1(x_0), x_0$ fixed in M, it can be Taylor expanded with respect to the r-jet Z. The resulting linear operator has a principal symbol, and if it is elliptic D is said to be locally elliptic.

Locally elliptic PDE

Elleiptic PDE usually give smooth solutions.

If the scalar PDE : find $u \in C(O; \mathbb{C})$ such that :$D(u) = g$ in O, where O is an open in $\mathbb{R}^m$ has a solution $u_0 \in C_\infty(O; \mathbb{C})$ at x_0 and if the scalar r order differential operator is elliptic in u_0 then, for any s, there are functions $u \in C_s(O; \mathbb{C})$ which are solutions in a neighborhood n(x_0). Moreover is g is smooth then $u(x) - u_0(x) = o\left(\|x - x_0\|^{r+1}\right)$

Quasi linear symmetric hyperbolic PDE

Hyperbolic PDE are the paradigm of "well posed" problem : they give unique solution, continuously depending on the initial values. One of their characteristic is that a variation of the initial value propagated as a wave.

Theorem 2492 *(Taylor 3 p.414) A quasi linear symmetric PDE is a PDE :
find u∈ $C(J \times M; \mathbb{C})$ such that :*

$B(t, x, u) \frac{\partial u}{\partial t} = \sum_{\alpha=1}^{m} A^{\alpha}(t, x, u) \frac{\partial u}{\partial x^{\alpha}} + g(t, x, u)$

$u(0, x) = g(x)$

where

$E(M, V, \pi)$ *is a vector bundle on a m dimensional manifold M,*

$J = [0, T] \subset \mathbb{R}$

$B(t, x, u), A^{\alpha}(t, x, u) \in \mathcal{L}(V; V)$ *and the A^{α} are represented by self adjoint
matrices $[A^{\alpha}] = [A^{\alpha}]^{*}$*

*Then if $g \in H^{n}(E)$ with $n > 1 + \frac{m}{2}$ the PDE has a unique solution u in a
neighborhood of t=0 and $u \in C(J; H^{n}(E))$*

The theorem can be extended to a larger class of semi linear non symmetric
equations. There are similar results for quasi linear second order PDE.

Method of the characteristics

This method is usually, painfully and confusely, explained in many books
about PDE. In fact it is quite simple with the r jet formalism. A PDE of order r
is a closed subbundle F of the r jet extension $J^{r}E$ of a fibered manifold E. This
seems a bit abstract, but is not too difficult to imagine when one remembers that
a subbundle has, for base, a submanifold of M. Clearly a solution $X : M \to \mathfrak{X}(E)$
is such that there is some relationship between the coordinates ξ_{α} of a point x
in M and the value of $X^{i}(x)$. This is all the more true when one adds the value
of the derivatives, linked in the PDE, with possibly other coercives conditions.
When one looks at the solutions of classical PDE one can see they are some
deformed maps of these constraints (notably the shape of a boundary). So one
can guess that solutions are more or less organized around some subbundles
of E, whose base is itself a submanifold of M. In each of these subbundles
the value of $X^{i}(x)$, or of some quantity computed from it, are preserved.
So the charcateristics" method sums up to find "primitives", meaning sections
Y∈ $\mathfrak{X}_{k}(E), k < r$ such that $J^{r-k}Y \in F$.

One postulates that the solutions X belong to some subbundle, defined
through a limited number of parameters (thus these parameters define a sub-
manifold of M). So each derivative can be differentiated with respect to these
parameters. By reintroducing these relations in the PDE one gets a larger sys-
tem (there are more equations) but usually at a lower order. From there one
can reduce the problem to ODE.

The method does not give any guarantee about the existence or unicity of
solutions, but usually gives a good insight of particular solutions. Indeed they
give primitives, and quantities which are preserved along the transformation of
the system. For instance in hyperbolic PDE they can describe the front waves
or the potential discontinuities (when two waves collide).

The classical example is for first order PDE.

Let be the scalar PDE in $\mathbb{R}^{m}$: find $u \in C_{1}(\mathbb{R}^{m}; \mathbb{C})$ such that $D(x, u, p) = 0$
where p=$\left(\frac{\partial u}{\partial x_{\alpha}}\right)_{\alpha=1}^{m}$. In the r jet formalism the 3 variables x,u,p can take any

value at any point x. $J^r E$ is here the 1 jet extension which can be assimilated with $\mathbb{R}^m \otimes \mathbb{R} \otimes \mathbb{R}^{m*}$

One looks for a curve in $J^1 E$, that is a map : $\mathbb{R} \to J^1 E :: Z(s) = (x(s), u(s), p(s))$ with the scalar parameter s, which is representative of a solution. It must meet some conditions. So one adds all the available relations coming from the PDE and the defining identities : $\frac{du}{ds} = p\frac{dx}{ds}$ and $d\left(\sum_\alpha p_\alpha dx^\alpha\right) = 0$. Put all together one gets the Lagrange-Charpit equations for the characteristics :

$$\frac{\dot{x}_\alpha}{D'_{p_\alpha}} = \frac{\dot{p}_\alpha}{D'_{x_\alpha} + D'_u p_\alpha} = \frac{\dot{u}}{\sum_\alpha p_\alpha D'_{p_\alpha}} \text{ with } \dot{x}_\alpha = \frac{dx_\alpha}{ds}, \dot{p}_\alpha = \frac{dp_\alpha}{ds}, \dot{u} = \frac{du}{ds}$$

which are ODE in s.

A solution u(x(s)) is usually specific, but if one has some quantities which are preserved when s varies, and initial conditions, these solutions are of physical significance.

32 VARIATIONAL CALCULUS

The general purpose of variational calculus is to find functions such that the value of some functional is extremum. This problem could take many different forms. We start with functional derivatives, which is an extension of the classical method to find the extremum of a function. The general and traditional approach of variational calculus is through lagrangians. It gives the classical Euler-Lagrange equations, and is the starting point for more elaborate studies on the "variational complex", which is part of topological algebra.

We start by a reminder of definitions and notations which will be used all over this section.

32.0.10 Notations

1. Let M be a m dimensional real manifold with coordinates in a chart $\psi(x) = (\xi^\alpha)_{\alpha=1}^m$. A holonomic basis with this chart is $\partial\xi_\alpha$ and its dual $d\xi^\alpha$. The space of antisymmetric p covariant tensors on TM^* is a vector bundle, and the set of its sections : $\mu : M \to \Lambda_p TM^* :: \mu = \sum_{\{\alpha_1...\alpha_p\}} \mu_{\alpha_1...\alpha_p}(x) d\xi^{\alpha_1} \wedge .. \wedge d\xi^{\alpha_p}$ is denoted as usual $\Lambda_p(M;\mathbb{R}) \equiv \mathfrak{X}(\Lambda_p TM^*)$. We will use more often the second notation to emphasize the fact that this is a map from M to $\Lambda_p TM^*$.

2. Let $E(M,V,\pi)$ be a fiber bundle on M, φ be some trivialization of E, so an element p of E is : $p = \varphi(x,u)$ for $u \in V$. A section $X \in \mathfrak{X}(E)$ reads : X(x)=$\varphi(x,\sigma(x))$

Let ϕ be some chart of V, so a point u in V has the coordinates $\phi(u) = (\eta^i)_{i=1}^n$

The tangent vector space $T_p E$ at p has the basis denoted $\partial x_\alpha = \varphi'_x(x,u)\partial\xi_\alpha, \partial u_i = \varphi'_u(x,u)\partial\eta_i$ for $\alpha = 1...m, i = 1...n$ with dual (dx^α, du^i)

3. $J^r E$ is a fiber bundle $J^r E(M, J_0^r(\mathbb{R}^m, V)_0, \pi^r)$. A point Z in $J^r E$ reads : $Z = (z, z_{\alpha_1...\alpha_s} : 1 \leq \alpha_k \leq m, s = 1..r)$. It is projected on a point x in M and has for coordinates :

$\Phi(Z) = \zeta = (\xi^\alpha, \eta^i, \eta^i_{\alpha_1...\alpha_s} : 1 \leq \alpha_1 \leq \alpha_2... \leq \alpha_s \leq m, i = 1...n, s = 1..r)$

A section $Z \in \mathfrak{X}(J^r E)$ reads : $Z = (z, z_{\alpha_1...\alpha_s} : 1 \leq \alpha_k \leq m, s = 1..r)$ where each component $z_{\alpha_1...\alpha_s}$ can be seen as independant section of $\mathfrak{X}(E)$.

A section X of E induces a section of $J^r E$: $J^r X(x)$ has for coordinates :
$\Phi(J^r X) = \left(\xi^\alpha, \sigma^i(x), \frac{\partial^s \sigma^i}{\partial\xi^{\alpha_1}..\partial\xi^{\alpha_s}}|_x\right)$

The projections are denoted:
$\pi^r : J^r E \to M : \pi^r(j_x^r X) = x$
$\pi_0^r : J^r E \to E : \pi_0(j_x^r X) = X(x)$
$\pi_s^r : J^r E \to J^s E : \pi_s^r(j_x^r X) = j_x^s X$

4. As a manifold $J^r E$ has a tangent bundle with holonomic basis :
$(\partial x_\alpha, \partial u_i^{\alpha_1...\alpha_s}, 1 \leq \alpha_1 \leq \alpha_2... \leq \alpha_s \leq m, i = 1...n, s = 0...r)$ and a cotangent bundle with basis $\left(dx^\alpha, du^i_{\alpha_1...\alpha_s}, 1 \leq \alpha_1 \leq \alpha_2... \leq \alpha_s \leq m, i = 1...n, s = 0...r\right)$.

The vertical bundle is generated by the vectors $\partial u_i^{\alpha_1...\alpha_s}$.

A vector on the tangent space $T_Z J^r E$ of $J^r E$ reads :

$W_Z = w_p + \sum_{s=1}^r \sum_{\alpha_1 \leq .. \leq \alpha_s} w^i_{\alpha_1...\alpha_{\alpha_s}} \partial u_i^{\alpha_1...\alpha_{\alpha_s}}$ with $w_p = w^\alpha \partial x_\alpha + w^i \partial u_i \in T_p E$

The projections give :

$\pi^{r\prime}(Z) : T_Z J^r E \to T_{\pi^r(Z)} M : \pi^{r\prime}(Z) W_Z = w^\alpha \partial \xi_\alpha$

$\pi_0^{r\prime}(Z) : T_Z J^r E \to T_{\pi_0^r(Z)} E : \pi^{r\prime}(Z) W_Z = w_p$

The space of antisymmetric p covariant tensors on $TJ^r E^*$ is a vector bundle, and the set of its sections : $\varpi : J^r E \to \Lambda_p T J^r E^*$ is denoted $\mathfrak{X}(\Lambda_p T J^r E^*) \equiv \Lambda_p(J^r E)$.

A form on $J^r E$ is π^r horizontal if it is null for any vertical vertical vector. It reads :

$\varpi = \sum_{\{\alpha_1...\alpha_p\}} \varpi_{\alpha_1...\alpha_p}(Z) dx^{\alpha_1} \wedge .. \wedge dx^{\alpha_p}$ where $Z \in J^r E$

The set of p horizontal forms on $J^r E$ is denoted $\mathfrak{X}^H(\Lambda_p T J^r E^*)$

5. A projectable vector field W on E is such that :

$\exists Y \in \mathfrak{X}(TM) : \forall p \in E : \pi'(p) W(p) = Y(\pi(p))$

Its components read : $W = W^\alpha \partial x_\alpha + W^i \partial u_i$ where W^α does not depend on u. A vertical vector field on E $(W^\alpha \equiv 0)$ is projectable.

It defines, at least locally, with any section X on E a one parameter group of fiber preserving morphisms on E through its flow : $U(t) : \mathfrak{X}(E) \to \mathfrak{X}(E) ::$ $U(t) X(x) = \Phi_W(X(\Phi_Y(x, -t)), t)$

This one parameter group of morphism on E has a prolongation as a one parameter group of fiber preserving morphisms on $J^r E$:.

$J^r U(t) : \mathfrak{X}(J^r E) \to \mathfrak{X}(J^r E) :: J^r U(t) Z(x)$

$= \Phi_W(j^r X(x), t) = j^r_{\Phi_Y(x,t)}(\Phi_W(X(\Phi_Y(x, -t)), t))$ with any section X on E such that $Z(x) = j^r X(x)$.

A projectable vector field W on E has a prolongation as a projectable vector field $J^r W$ on $J^r E$ defined through the derivative of the one parameter group : $J^r W(j^r X(x)) = \frac{\partial}{\partial t} J^r \Phi_W(j^r X(x), t)|_{t=0}$. Its expression is the following

$J^r W = W^\alpha \partial x_\alpha + W^i \partial u_i + \sum_{s=1}^r \sum_{\alpha_1 \leq .. \leq \alpha_s} W^i_{\alpha_1...\alpha_s} \partial u_i^{\alpha_1...\alpha_s}$ with $W^i_{\beta\alpha_1...\alpha_s} = d_\beta W^i_{\alpha_1...\alpha_s} - \sum_\gamma \eta^i_{\gamma\alpha_1...\alpha_s} \partial u_i^\gamma$ and the total differential d_β

By construction the r jet prolongation of the one parameter group of morphism induced by W is the one parameter group of morphism induced by the r jet prolongation of W: $J^r U(t) Z(x) = \Phi_{J^r W}(Z(x), t)$. So the Lie derivative $\pounds_{J^r W} Z$ of a section Z of $J^r E$ is a section of the vertical bundle : $V J^r E$. And if W is vertical the map : $\pounds_{J^r W} Z : \mathfrak{X}(J^r E) \to V J^r E$ is fiber preserving and $\pounds_{J^r W} Z = J^r W(Z)$

32.1 Functional derivatives

At first functional derivatives are the implementation of the general theory of derivatives to functions whose variables are themselves functions. The theory of distributions gives a rigorous framework to a method which is enticing because it is simple and intuitive.

The theory of derivatives holds whenever the maps are defined over affine normed spaces, even if it is infinite dimensional. So for any map : $\ell : E \to \mathbb{R}$ where E is some normed space of functions, one can define derivatives, and if

E is a product $E_1 \times E_2.. \times E_p$ partial derivatives, of first and higher order. A derivative such that $\frac{d\ell}{dX}$, where X is a function, is a linear map : $\frac{d\ell}{dX} : E \to \mathbb{R}$ meaning a distribution, if it is continuous. So, in many ways, one can see distributions as the "linearisation" of functions of functions. All the classic results apply to this case and there is no need to tell more on the subject.

It is a bit more complicated when the map : $\ell : \mathfrak{X}(J^r E) \to \mathbb{R}$ is over the space $\mathfrak{X}(J^r E)$ of sections of the r jet prolongation $J^r E$ of a vector bundle E. $\mathfrak{X}(J^r E)$ is an infinite dimensional complex vector space.This is not a normed space but only a Fréchet space. So we must look for a normed vector subspace F_r of $\mathfrak{X}(J^r E)$. Its choice depends upon the problem.

32.1.1 Definition of the functional derivative

Definition 2493 *A functional $L = \ell \circ J^r$ with $\ell : J^r E \to \mathbb{R}$ is defined on a vector bundle $E(M, V, \pi)$ and X belongs to some normed vector subspace $F \subset \mathfrak{X}(E)$ has a* **functional derivative** *$\frac{\delta L}{\delta X}$ at X_0 if there is a distribution $\frac{\delta L}{\delta X} \in F'$ such that :*

$\forall \delta X \in F : \lim_{\|\delta X\|_F \to 0} \left| L(X_0 + \delta X) - L(X_0) - \frac{\delta L}{\delta X}(\delta X) \right| = 0$

The key point in the definition is that the distribution $\frac{\delta L}{\delta X}$ acts on δX : we have a unique distribution acting on the section δX.

Any variation δX of the section entails a variation :

$J^r \delta X = (\delta X, D_{\alpha_1...\alpha_s} \delta X : 1 \leq \alpha_k \leq m, s = 1..r)$ and $D_{\alpha_1...\alpha_s} \delta X$ can be Taylor expanded in a neighborhood of δX_0, so if L is differentiable it has a functional derivative but the interest is to give a more practical version of the usual derivative.

If ℓ is a continuous linear map (it belongs to F'_r) then it has a functional derivative : $\ell(Z_0 + \delta Z) - \ell(Z_0) = \ell(\delta Z)$ and $\frac{d\ell}{dZ} = \ell$

Let us assume that there is a normed vector subspace $F_r \subset \mathfrak{X}(J^r E)$ such that $J^r X \in F_r$ when $X \in F$.

The functional $\ell(Z) : \mathfrak{X}(J^r E) \to \mathbb{R}$ is differentiable, in the usual meaning, in F_r at Z_0 if :

$\exists \frac{d\ell}{dZ}(Z_0) \in \mathcal{L}(F_r; \mathbb{R}) :: \forall \delta Z \in F_r : \lim_{\|\delta Z\|_{F_r} \to 0} \left| \ell(Z_0 + \delta Z) - \ell(Z_0) - \frac{d\ell}{dZ} \delta Z \right| = 0$

with the variation of $Z : \delta Z = (\delta z, \delta z_{\alpha_1...\alpha_s} : 1 \leq \alpha_k \leq m, s = 1..r) \in F_r$ that is a map : $\delta Z : M \to J^r E$

F_r is a complex vector space but ℓ is real valued, so we need $\frac{\delta \ell}{\delta Z}$ to be only real continuous.

By definition $\frac{d\ell}{dZ}$ is a linear continuous functional on F_r, so this is a distribution : $\frac{d\ell}{dZ} \in F'_r$ which depends on Z_0.

And if ℓ is continuously differentiable the map : $\frac{d\ell}{dZ} : F_r \to F'_r :: \frac{d\ell}{dZ}(Z_0)$ is continuous.

ℓ can be seen as a multivariable map, and the partial derivatives, computed by keeping all but one $z_{\alpha_1...\alpha_s}$ constant, are $\frac{\partial \ell}{\partial z_{\alpha_1...\alpha_s}}$ and : $\frac{d\ell}{dZ} \delta Z = \sum_{s=0}^{r} \sum_{\alpha_1...\alpha_s=1}^{m} \frac{\partial \ell}{\partial z_{\alpha_1...\alpha_s}} \delta z_{\alpha_1...\alpha_s}$

Each of the components of δZ can be seen as a section of E, which are not related to each other (except if δZ is the prolongation of a section on E). If ℓ is continuously differentiable its partial derivatives are continuous, and each $\frac{\delta \ell}{\delta z_{\alpha_1 \ldots \alpha_s}}$ is a distribution in F'.

For a section $X \in F$ and the map $\ell(J^r X) = \ell \circ J^r(X) = L(X)$, if ℓ has a derivative at $Z_0 = J^r X_0$ we have :

$\forall \delta X \in F : \lim_{\|\delta J^r X\|_{F_r} \to 0} \left| \ell(J^r X_0 + J^r \delta X) - \ell(J^r X_0) - \frac{d\ell}{dZ} J^r \delta X \right| = 0$
with $\delta Z = J^r \delta X$

which reads :

$\forall \delta X \in F : \lim_{\|\delta J^r X\|_{F_r} \to 0} \left| L(X_0 + \delta X) - L(X_0) - \frac{d\ell}{dZ} J^r \delta X \right| = 0$
with $\frac{d\ell}{dZ} J^r \delta X = \sum_{s=0}^{r} \sum_{\alpha_1 \ldots \alpha_s = 1}^{m} \frac{\partial \ell}{\partial z_{\alpha_1 \ldots \alpha_s}} (D_{\alpha_1 \ldots \alpha_s} \delta X)$

δX s assumed to be r differentiable (to compute $J^r \delta X$) so by expanding each component of $J^r \delta X$ we have a, complicated, linear function of δX.

Notice that the order r is involved in the definition of ℓ and we do not consider the derivative of the map : $\frac{\delta L}{\delta X} : F \to F'$

A special interest of functional derivatives is that they can be implemented with composite maps.

In the conditions above, assume that Z depends on some map $f : O \to M$ were O is an open in $\mathbb{R}^k$.

Then if L has a functional derivative with respect to X, $\delta X = X(f) - X(f + \delta f) \sim \frac{dX}{df} \delta f$ and L has the functional derivative with respect to f : $\frac{\delta L}{\delta f} = \frac{\delta L}{\delta X} \frac{dX}{df}$. This expression makes sense because it is the product of a distribution by a function.

32.1.2 Functional defined as an integral

ℓ is defined through the integral of a m form on the manifold M : $\ell(Z) = \int_M \lambda(Z) d\xi^1 \wedge \ldots \wedge d\xi^m$ . . The quantity $\lambda(Z) d\xi^1 \wedge \ldots \wedge d\xi^m$ is a m form, so its component λ is a function which changes accordingly in a change of charts of M.

We precise the domain of each map as follows, with $1 \leq p < \infty$. We assume that there is a Radon measure μ on M.

i) The sections of $J^r E : Z \in L^p(M, \mu, J^r E) \subset \mathfrak{X}(J^r E)$. This is a Banach vector space with the usual norm $\|Z\|_p = \left(\sum_{s=0}^{r} \int_M \|z_{\alpha_1 \ldots \alpha_s}(x)\|^p \mu \right)^{1/p}$ and $\|z_{\alpha_1 \ldots \alpha_s}(x)\|$ the norm in V^s. And each of the component $z_{\alpha_1 \ldots \alpha_s} \in L^p(M, \mu, E)$.

ii) The sections of E : $X \in W^{r,p}(E)$, they are r continuously differentiable and their r prolongation belong to $L^p(M, \mu, J^r E)$. It is a Banach vector space with the norm : $\|X\|_{r,p} = \left(\sum_{s=0}^{r} \sum_{\alpha_1 \ldots \alpha_s} \int_M \|D_{\alpha_1 \ldots \alpha_s} X\|^p \mu \right)^{1/p}$

iii) λ belongs to the space of 1 continuously differentiable functions on $L^p(M, \mu, J^r E)$. So their partial derivatives $\frac{\partial \lambda}{\partial z_{\alpha_1 \ldots \alpha_s}}(Z_0)$ belong to the topological dual $L^q(M, \mu, E)$ of $L^p(M, \mu, E)$ with $\frac{1}{p} + \frac{1}{q} = 1$ which is a Banach vector space and

$\int_M \left\| \frac{\partial \lambda}{\partial z_{\alpha_1 \ldots \alpha_s}}(Z_0) \right\|^q \mu < \infty$

Theorem 2494 *If $E(M,V,\pi)$ is a class r vector bundle, λ a r+1 differentiable map in $L^p\left(M,\mu,J^{2r}E\right)$ with $1\leq p<\infty$, then the map : $L\left(X\right)=\int_M \lambda\left(J^r X\right)\mu$ is differentiable for any section $X_0\in W^{r,p}\left(E\right)$, and its derivative is given by :*

$$\frac{d\ell}{dZ}J^r\delta X = \sum_{s=0}^r \sum_{\alpha_1...\alpha_s=1}^m T\left(\frac{\partial\lambda}{\partial z_{\alpha_1...\alpha_s}}\left(J^r X_0\right)\right)D_{\alpha_1...\alpha_s}\delta X$$

Proof. In each fiber : $\lambda\left(Z\left(x\right)\right)=\lambda\left(z\left(x\right),z_\alpha\left(x\right),..z_{\alpha_1...\alpha_r}\left(x\right)\right)$. is a map : $V\times V..\times V=V^N\to\mathbb{R}$. It is continuously differentiable then we have continuous partial derivatives and :

$\forall Z_0\in F_r,\delta Z\in F_r, x\in M : \lambda\left(Z_0\left(x\right)+\delta Z\left(x\right)\right)=\lambda\left(Z_0\left(x\right)\right)+\sum_{s=0}^r\sum_{\alpha_1...\alpha_s=1}^m\frac{\partial\lambda}{\partial z_{\alpha_1...\alpha_s}}\left(Z_0\left(x\right)\right)\delta z_{\alpha_1...\alpha_s}\left(x\right)\cdot$
$\varepsilon\left(\delta Z\left(x\right)\right)\left\|\delta Z\left(x\right)\right\|_{V^N}$

with $\varepsilon\left(\delta Z\left(x\right)\right)\to 0$ when $\delta Z\left(x\right)\to 0$

In this formula $\left\|\delta Z\left(x\right)\right\|_{V^N}=\left(\sum_{s=0}^r\sum_{\alpha_1...\alpha_s=1}^m\left\|\delta z_{\alpha_1...\alpha_r}\left(x\right)\right\|^p\right)^{1/p}$ (a fiber is a finite dimensional vector space so all norms are equivalent).

$\ell\left(Z\right)=\ell\left(Z_0\right)+\sum_{s=1}^r\sum_{\alpha_1...\alpha_s=0}^m\int_M\frac{\partial\lambda}{\partial z_{\alpha_1...\alpha_s}}\left(Z_0\left(x\right)\right)\delta z_{\alpha_1...\alpha_s}\left(x\right)d\xi+\int_M\varepsilon\left(\delta Z\left(x\right)\right)\left\|\delta Z\left(x\right)\right\|_{V^N}d\xi$
with $d\xi=d\xi^1\wedge...\wedge d\xi^m$

From the Hölder's inequality: $\frac{\partial\lambda}{\partial z_{\alpha_1...\alpha_s}}\left(Z_0\right)\in L^q\left(M,\mu,E\right),\delta z_{\alpha_1...\alpha_s}\in L^p\left(M,\mu,E\right)$
so $\frac{\partial\lambda}{\partial z_{\alpha_1...\alpha_s}}\left(Z_0\right)\delta z_{\alpha_1...\alpha_s}\in L^1\left(M,\mu,E\right)$

As $\varepsilon\left(\delta Z\left(x\right)\right)\to 0$ there is some $C\in\mathbb{R}:\left|\varepsilon\left(\delta Z\left(x\right)\right)\right|<C$ and $\int_M\left|\varepsilon\left(\delta Z\left(x\right)\right)\right|\left\|\delta Z\left(x\right)\right\|_{V^N}d\xi\leq C\left\|\delta Z\right\|_{L^p}^p$

$\frac{d\ell}{dZ}\delta Z=\sum_{s=1}^r\sum_{\alpha_1...\alpha_s=0}^m\int_M\frac{\partial\lambda}{\partial z_{\alpha_1...\alpha_s}}\left(Z_0\left(x\right)\right)\delta z_{\alpha_1...\alpha_s}\left(x\right)d\xi$

With the map : $T:L^q\left(M,\mu,E\right)\to L^p\left(M,\mu,E\right)'::T(f)=\int_M fzd\xi$

$\frac{d\ell}{dZ}\delta Z=\sum_{s=0}^r\sum_{\alpha_1...\alpha_s=1}^m T\left(\frac{\partial\lambda}{\partial z_{\alpha_1...\alpha_s}}\left(Z_0\right)\right)\delta z_{\alpha_1...\alpha_s}$ and : $\frac{\partial\ell}{\partial z_{\alpha_1...\alpha_s}}=T\left(\frac{\partial\lambda}{\partial z_{\alpha_1...\alpha_s}}\left(Z_0\right)\right)$

For a section X :
$\frac{d\ell}{dZ}J^r\delta X=\sum_{s=0}^r\sum_{\alpha_1...\alpha_s=1}^m T\left(\frac{\partial\lambda}{\partial z_{\alpha_1...\alpha_s}}\left(J^r X_0\right)\right)D_{\alpha_1...\alpha_s}\delta X$ ∎

Theorem 2495 *If $E(M,V,\pi)$ is a class 2r vector bundle, λ a r+1 differentiable compactly supported map in $L^p\left(M,\mu,J^{2r}E\right)$ with $1\leq p<\infty$,$\mu=d\xi^1\wedge..\wedge d\xi^m$, then the map : $L\left(X\right)=\int_M \lambda\left(J^r X\right)\mu$ has a functional derivative in any point $X_0\in W^{2r,p}\left(E\right)$, given by :*

$$\frac{\delta L}{\delta X}\left(X_0\right)=T\left(\sum_{s=0}^r\sum_{\alpha_1...\alpha_s=1}^m\left(-1\right)^s d_{\alpha_1}d_{\alpha_2}...d_{\alpha_s}\frac{\partial\lambda}{\partial z_{\alpha_1...\alpha_s}}\left(J^{2r}X_0\right)\right)$$

Now we use the additional conditions.

Proof. If λ is r+1 differentiable and compactly supported (see the Derivatives of distributions) : $T\left(f\left(J^r X_0\right)\right)D_{\alpha_1...\alpha_s}\delta X=\left(-1\right)^s T\left(D_{\alpha_1...\alpha_s}f\left(J^r X_0\right)\right)\left(\delta X\right)$

thus (we must consider $J^{2r}X_0$ to account for the double differentiation):

$T\left(\frac{\partial\lambda}{\partial z_{\alpha_1...\alpha_s}}\left(J^{2r}X_0\right)\right)D_{\alpha_1...\alpha_s}\delta X=\left(-1\right)^s T\left(D_{\alpha_1...\alpha_s}\frac{\partial\lambda}{\partial z_{\alpha_1...\alpha_s}}\left(J^{2r}X_0\right)\right)\left(\delta X\right)$

So L has a functional derivative :
$\frac{\delta L}{\delta X}=\sum_{s=0}^r\sum_{\alpha_1...\alpha_s=1}^m\left(-1\right)^s T\left(D_{\alpha_1...\alpha_s}\frac{\partial\lambda}{\partial z_{\alpha_1...\alpha_s}}\left(J^{2r}X_0\right)\right)$
$=T\left(\sum_{s=0}^r\sum_{\alpha_1...\alpha_s=1}^m\left(-1\right)^s D_{\alpha_1...\alpha_s}\frac{\partial\lambda}{\partial z_{\alpha_1...\alpha_s}}\left(J^{2r}X_0\right)\right)$ ∎

The first derivation is with respect to x in $J^r X_0$. This is called the total differential and denoted : $D_{\alpha_1...\alpha_s}=d_{\alpha_1}d_{\alpha_2}...d_{\alpha_s}$

with $d_\alpha f = \frac{\partial f}{\partial \xi^\alpha} + \sum_{s=1}^r \sum_{\beta_1...\beta_s} \frac{\partial f}{\partial z^i_{\alpha\beta_1...\beta_s}} z^i_{\alpha\beta_1...\beta_s}$

Notice that here we have a distribution, acting on functions, and defined through a m form, thus we need to account for $(-1)^s$.

32.2   Lagrangian

This is the general framework of variational calculus. It can be implemented on any fibered manifold E.

The general problem is to find a section of a E for which the integral on the base manifold M: $\ell(Z) = \int_M \mathcal{L}(Z)$ is extremum. L is a m form depending on a section of the r jet prolongation of E called the lagrangian.

32.2.1   Lagrangian form

Definition 2496 *A r order lagrangian is a base preserving map* $\mathcal{L} : J^r E \to \Lambda_m(M;\mathbb{R})$ *where* $E(M, V, \pi)$ *is a fiber bundle on a real m dimensional manifold.*

Comments :

i) Base preserving means that the operator is local : if $Z \in J^r E, \pi^r(Z) = x \in M$ and $L(x) : J^r E(x) \to \Lambda_m T_x M^*$

ii) L is a differential operator as defined previously and all the results and definitions hold. The results which are presented here are still valid when E is a fibered manifold.

iii) Notice that L is a m form, the same dimension as M, and it is real valued.

iv) This definition is consistent with the one given by geometers (see Kolar p.388). Some authors define the lagrangian as a horizontal form on $J^r E$, which is more convenient for studies on the variational complex. As we need the integral $\ell(Z)$ and one uses commonly the lagrangian function defined on M, we would have to go from a horizontal form to a m form on M. So it is better to keep it simple. Moreover it is useful to keep in mind that the lagrangian is, for variational calculus, a differential operator which goes from a section on E to a m form. With the present definition we preserve, in the usual case of vector bundle and in all cases with the lagrangian function, all the apparatus of differential operators already available.

v) $J^r : E \mapsto J^r E$ and $\Lambda_m : M \mapsto \Lambda_m(M;\mathbb{R})$ are two bundle functors and $\mathcal{L} : J^r \hookrightarrow \Lambda_m$ is a natural operator.

With the bases and coordinates above, L reads : $\lambda\left(\xi^\alpha, \eta^i, \eta^i_{\alpha_1...\alpha_s}\right) d\xi^1 \wedge .. \wedge d\xi^m$ that we denote $\mathcal{L}(Z) = \lambda(z) d\xi^1 \wedge ... \wedge d\xi^m$

For a section $X \in \mathfrak{X}(E)$ one can also write : $\mathcal{L}(J^r X(x)) = \mathcal{L} \circ J^r(X(x)) = \widehat{\mathcal{L}}(X(x)) = J^r X^* \mathcal{L}(x)$ with the pull back of the map $\mathcal{L}$ by the map : $J^r X : M \to J^r E$. But keep in mind that this is not the pull back of a m form defined on $J^r E$. So we have : $\mathcal{L}(J^r X(x)) = J^r X^* \lambda(x) d\xi^1 \wedge ... \wedge d\xi^m$

32.2.2   Scalar lagrangian

The definition above is a geometric, intrinsic definition. In a change of charts on M the value of $\lambda(Z)$ changes as any other m form. In a change of holonomic

basis :

$d\xi^\alpha \to d\widetilde{\xi}^\alpha : \mathcal{L}(Z) = \lambda(Z(x)) d\xi^1 \wedge ... \wedge d\xi^m = \widetilde{\lambda}(Z(x)) d\widetilde{\xi}^1 \wedge ... \wedge d\widetilde{\xi}^m$ with
$\widetilde{\lambda} = \lambda \det\left[J^{-1}\right]$ where J is the jacobian $J = [F'(x)] \simeq \left[\frac{\partial\widetilde{\xi}^\alpha}{\partial\xi^\beta}\right]$

If M is endowed with a volume form ϖ_0 then $\mathcal{L}(Z)$ can be written $L(Z)\varpi_0$. Notice that any volume form (meaning that is never null) is suitable, and we do not need a riemannian structure for this purpose. For any pseudo-riemannian metric we have $\varpi_0 = \sqrt{|\det g|}d\xi^1 \wedge ... \wedge d\xi^m$ and $L(z)=\lambda(Z(x))/\sqrt{|\det g|}$

In a change of chart on M the component of ϖ_0 changes according to the rule above, thus L(Z) does not change : this is a function.

Thus $L : J^r E \to C(M;\mathbb{R})$ is a r order scalar differential operator that we will call the **scalar lagrangian** associated to $\mathcal{L}$.

32.2.3   Covariance

1. The scalar lagrangian is a function, so it shall be invariant by a change of chart on the manifold M.

If we proceed to the change of charts :

$\xi^\alpha \to \widetilde{\xi}^\alpha : \widetilde{\xi} = F(\xi)$ with the jacobian : $\left[\frac{\partial\widetilde{\xi}^\alpha}{\partial\xi^\beta}\right] = K$ and its inverse J.

the coordinate η^i on the fiber bundle will not change, if E is not some tensorial bundle, that we assume.

The coordinates $\eta^i_{\alpha_1...\alpha_s}$ become $\widetilde{\eta}^i_{\alpha_1...\alpha_s}$ with $\widetilde{\eta}^i_{\beta\alpha_1...\alpha_s} = \sum_\gamma \frac{\partial\xi^\gamma}{\partial\widetilde{\xi}^\beta} d_\gamma \eta^i_{\alpha_1...\alpha_s}$ where $d_\gamma = \frac{\partial}{\partial\xi^\gamma} + \sum_{s=1}^r \sum_{\beta_1 \leq ... \leq \beta_s} y^i_{\gamma\beta_1...\beta_s} \frac{\partial}{\partial y^i_{\beta_1...\beta_s}}$ is the total differential.

2. The first consequence is that the coordinates ξ of points of M cannot appear explicitly in the Lagrangian. Indeed the map F can be chosen almost arbitrarily.

3. The second consequence is that there shall be some relations between the partial derivatives of L. They can be found as follows with the simple case r=1 as example :

With obvious notations : $L\left(\eta^i, \eta^i_\alpha\right) = \widetilde{L}\left(\widetilde{\eta}^i, \widetilde{\eta}^i_\alpha\right) = \widetilde{L}\left(\eta^i, J^\beta_\alpha \eta^i_\beta\right)$

The derivative with respect to $J^\lambda_\mu, \lambda, \mu$ fixed is :
$\sum_{\alpha\beta i} \frac{\partial\widetilde{L}}{\partial\widetilde{\eta}^i_\alpha} \eta^i_\beta \delta^\beta_\lambda \delta^\mu_\alpha = 0 = \sum_i \frac{\partial\widetilde{L}}{\partial\widetilde{\eta}^i_\mu} \eta^i_\lambda$

The derivative with respect to $\eta^j_\lambda, \lambda, j$ fixed is :
$\sum_{\alpha\beta i} \frac{\partial\widetilde{L}}{\partial\widetilde{\eta}^i_\alpha} J^\beta_\alpha \delta^\beta_\lambda \delta^i_j = \frac{\partial L}{\partial\eta^j_\lambda} = \sum_\alpha \frac{\partial\widetilde{L}}{\partial\widetilde{\eta}^j_\alpha} J^\lambda_\alpha$

For : $J^\lambda_\mu = \delta^\lambda_\mu$ it comes :
$\forall\lambda : \sum_i \frac{\partial\widetilde{L}}{\partial\widetilde{\eta}^i_\lambda} \eta^i_\lambda = 0$
$\forall\lambda : \frac{\partial L}{\partial\eta^j_\lambda} = \frac{\partial\widetilde{L}}{\partial\widetilde{\eta}^j_\lambda}$
So there is the identity: $\forall\alpha : \sum_i \frac{\partial L}{\partial\eta^i_\alpha} \eta^i_\alpha = 0$

4. The third consequence is that some partial derivatives can be considered as components of tensors :

In the example above :
$\frac{\partial L}{\partial\eta^i}$ do not change, and are functions on M

755

$\sum \frac{\partial L}{\partial \eta_\lambda^j} \partial \xi^\alpha = \sum \frac{\partial \widetilde{L}}{\partial \widetilde{\eta}_\alpha^j} J_\alpha^\lambda \partial \xi^\alpha = \sum \frac{\partial \widetilde{L}}{\partial \widetilde{\eta}_\alpha^j} \partial \widetilde{\xi}^\alpha$ so $\forall i : \left(\frac{\partial L}{\partial \eta_\lambda^j}\right)_{\alpha=1}^m$ are the components of a vector field.

This remark comes handy in many calculations on Lagrangians.

5. The Lagrangian can also be equivariant in some gauge transformations. This is studied below with Noether currents.

32.3 Euler-Lagrange equations

Given a Lagrangian $\mathcal{L} : J^r E \to \Lambda_m(M; \mathbb{R})$ the problem is to find a section $X \in \mathfrak{X}(E)$ such that the integral $\ell(Z) = \int_M \mathcal{L}(Z)$ is extremum.

32.3.1 Vectorial bundle

General result

If E is a vectorial bundle we can use the functional derivative : indeed the usual way to find an extremum of a function is to look where its derivative is null. We use the precise conditions stated in the previous subsection.

1. If $E(M, V, \pi)$ is a class 2r vector bundle, λ a r+1 differentiable compactly supported map in $L^p\left(M, \mu, J^{2r}E\right)$ with $1 \leq p < \infty$ then the map : $\ell(J^r X) = \int_M \mathcal{L}(J^r X)$ has a functional derivative in any point $X_0 \in W^{2r,p}(E)$, given by :

$\frac{\widehat{\delta\ell}}{\delta X}(X_0) = T\left(\sum_{s=0}^r \sum_{\alpha_1...\alpha_s=1}^m (-1)^s d_{\alpha_1} d_{\alpha_2}...d_{\alpha_s} \frac{\partial \lambda}{\partial z_{\alpha_1...\alpha_s}} \left(J^{2r}X_0\right)\right)$

So :

$\forall \delta X \in W^{r,p}(E) : \lim_{\|\delta X\|_{W^{r,p}(E)} \to 0} \left| L(X_0 + \delta X) - L(X_0) - \frac{\widehat{\delta\ell}}{\delta X}(X_0)\delta X \right| = 0$

2. The condition for a local extremum is clearly : $\frac{\widehat{\delta\ell}}{\delta X}(X_0) = 0$ which gives the Euler-Lagrange equations :

i=1...n : $\sum_{s=0}^r \sum_{\alpha_1...\alpha_s=1}^m (-1)^s d_{\alpha_1} d_{\alpha_2}...d_{\alpha_s} \frac{\partial \lambda}{\partial \eta_{\alpha_1...\alpha_s}^i} \left(J^{2r}X_0\right) = 0$

We have a system of n partial differential equations of order 2r

For r=1 : $\frac{\partial L}{\partial \eta^i} - \sum_{\alpha=1}^n \frac{d}{d\xi^\alpha} \frac{\partial L}{\partial \eta_\alpha^i} = 0; i = 1...n$

For r=2 : $\frac{\partial L}{\partial \eta^i} - \frac{d}{d\xi^\alpha} \left(\frac{\partial L}{\partial \eta_\alpha^i} - \sum_\beta \frac{d}{d\xi^\beta} \frac{\partial L}{\partial \eta_{\alpha\beta}^i}\right) = 0; i = 1...n, \alpha = 1...m$

The derivatives are evaluated for $X(\xi)$: they are total derivatives $\frac{d}{d\xi^\alpha}$

$\alpha = 1...m : d_\alpha \lambda\left(\xi^\alpha, \eta^i, \eta_{\alpha_1...\alpha_s}^i\right) = \frac{\partial \lambda}{\partial \xi^\alpha} + \sum_{s=1}^r \sum_{\beta_1...\beta_s} \frac{\partial \lambda}{\partial \eta_{\alpha\beta_1...\beta_s}^i} \eta_{\alpha\beta_1...\beta_s}^i$

3. The condition λ compactly supported can be replaced by M is a compact manifold with boundary, continuous on the boundary.

First order Lagrangian

The **stress energy tensor** (there are many definitions) is the quantity :

$\sum_{\alpha\beta} T_\beta^\alpha \partial_\alpha x \otimes dx^\beta = \sum_{\alpha\beta} \left(\sum_{i=1}^n \frac{\partial \lambda}{\partial \eta_\alpha^i} \eta_\beta^i - \delta_\beta^\alpha \lambda\right) \partial_\alpha x \otimes dx^\beta \in TM \otimes TM^*$

It is easy to check the identity :

$\sum_{\alpha=1}^m \frac{d}{d\xi^\alpha} T_\beta^\alpha = -\frac{\partial \lambda}{\partial \xi^\beta} + \sum_{i=1}^p \mathfrak{E}_i(\lambda) \eta_\beta^i$

where $\mathfrak{E}_i(\lambda) = \left(\frac{\partial L}{\partial Z^i} - \sum_{\alpha=1}^m \frac{d}{d\xi^\alpha} \left(\frac{\partial L}{\partial Z_\alpha^i}\right)\right) d\xi^1 \wedge ... \wedge d\xi^m \in \Lambda_m TM$

and the Euler-Lagrange form is : $\mathfrak{E}(\lambda) = \sum_i \mathfrak{E}_i(\lambda) du^i \otimes d\xi^1 \wedge ... \wedge d\xi^m \in E^* \otimes \Lambda_m TM$

So if λ does not depend on ξ (as it should because of the covariance) there is a primitive : $\sum_{\alpha=1}^m \frac{d}{d\xi^\alpha} T_\beta^\alpha = 0$

32.3.2   First order lagrangian on $\mathbb{R}$

This is indeed the simplest but one of the most usual case. The theorems below use a method similar to the functional derivative. We give them because they use more precise, and sometimes more general conditions.

Euler Lagrange equations
   1.(Schwartz 2 p.303)
   $I = [a,b]$ is a closed interval in $\mathbb{R}$, F is a normed affine space, U is an open subset of FxF
   $\lambda : I \times U \to \mathbb{R}$ is a continuous function
   The problem is to find a map $X : I \to F$ such that : $\ell(X) = \int_a^b \lambda(t, X(t), X'(t)) dt$ is extremum
   The space of maps X used is the Fréchet space $C_1([a,b]; F)$
   2. We have the following results :
   i) the set of solutions is an open subset O of $C_1(I; F)$
   ii) the function : $\ell : C_1(I; E) \to \mathbb{R}$ is continuously differentiable in O, and its derivative in X_0 is :
   $\frac{\delta \ell}{\delta X}|_{X=X_0} \delta X = \int_a^b \left(\lambda'(t, X_0(t), X_0'(t))(0, \delta X(t), (\delta X)'(t))\right) dt = \int_a^b [\frac{\partial \lambda}{\partial X} \delta X + \frac{\partial \lambda}{\partial X'} (\delta X)'] dt$
   iii) If λ is a class 2 map, X_0 a class 2 map then :

   $\frac{\delta \ell}{\delta X}|_{X=X_0} \delta X = [\frac{\partial \lambda}{\partial X'}(t, X_0(t), X_0'(t)) \delta X(t)]_a^b + \int_a^b [\frac{\partial \lambda}{\partial X} - \frac{d}{dt} \frac{\partial L}{\partial X'}] \delta X dt$

   and if X is an extremum of ℓ with the conditions $X(a) = \alpha, X(b) = \beta$ then it is a solution of the ODE :
   $$\frac{\partial \lambda}{\partial X} = \frac{d}{dt}(\frac{\partial \lambda}{\partial X'}); X(a) = \alpha, X(b) = \beta$$

   If λ does not depend on t then for any solution the quantity : $\lambda - \frac{\partial \lambda}{\partial X'} X' = Ct$

   iv) If X is a solution, it is a solution of the variational problem on any interval in I:
   $\forall [a_1, b_1] \subset [a,b], \int_{a_1}^{b_1} \lambda(t, X, X') dt$ is extremum for X

Variational problem with constraints
   With the previous conditions.
   1. Additional constraints :
   (Schwartz 2 p.323) The problem is to find a map X:I$\to F$ such that :
   $\ell(X) = \int_a^b \lambda(t, X(t), X'(t)) dt$ is extremum
   and for $k = 1..n$:$J_k = \int_a^b M_k(t, X(t), X'(t)) dt$ where J_k is a given scalar
   Then a solution must satisfy the ODE :

$$\frac{\partial \lambda}{\partial X} - \frac{d}{dt}\left(\frac{\partial \lambda}{\partial X'}\right) = \sum_{k=1}^{n} c_k \left[\frac{\partial M_k}{\partial X} - \frac{d}{dt}\frac{\partial M_k}{\partial X'}\right] \text{ with some scalars } c_k$$

2. Variable interval :

(Schwartz 2 p.330)

$I = [a, b]$ is a closed interval in $\mathbb{R}$, F is a normed affine space, U is an open subset of FxF

$\lambda : I \times U \to \mathbb{R}$ is a class 2 function

The problem is to find the maps X:$I \to F$, and the 2 scalars a,b such that :

$\ell(X) = \int_a^b \lambda(t, X(t), X'(t))\, dt$ is extremum

Then the derivative of ℓ with respect to X,a,b at (X_0, a_0, b_0) is :

$$\frac{\delta \ell}{\delta X}(\delta X, \delta\alpha, \delta\beta) = \int_{a_0}^{b_0}\left[\frac{\partial \lambda}{\partial X}\delta X - \frac{d}{dt}\frac{\partial \lambda}{\partial X'}\right](t, X_0(t), X_0'(t))\,\delta X\, dt$$

$$+ \left(\frac{\partial \lambda}{\partial X'}(b, X_0(b), X_0'(b))\,\delta X(b) + \lambda(b, X_0(b), X_0'(b))\right)$$

$$- \left(\frac{\partial \lambda}{\partial X'}(a, X_0(a), X_0'(b_0))\,\delta X(a) + \lambda(a, X_0(a), X_0'(a))\right)$$

With the notation $X(a) = \alpha, X(b) = \beta$ this formula reads :

$\delta X(b) = \delta\beta - X'(b_0)\,\delta b, \delta X(a) = \delta\alpha - X'(a_0)\,\delta a$

Hamiltoninan formulation

This is the classical formulation of the variational problem in physics, when one variable (the time t) is singled out.

(Schwartz 2 p.337)

$I = [a, b]$ is a closed interval in $\mathbb{R}$, F is a finite n dimensional vector space, U is an open subset of FxF

$\lambda : I \times U \to \mathbb{R}$ is a continuous function which is denoted here (with the usual 1 jet notation) : $L\left(t, y^1, ...y^n, z^1, ...z^m\right)$ where for a section X(t) : $z^i = \frac{dy^i}{dt}$

The problem is to find a map X:$I \to F$ such that : $\ell(X) = \int_a^b \lambda(t, X(t), X'(t))\, dt$ is extremum

The space of maps X is the Fréchet space $C_1([a, b]; F)$

By the change of variables :

y^i replaced by q_i

z^i replaced by $p_i = \frac{\partial L}{\partial z^1}, i = 1...n$ which is called the momentum conjugate to q_i

one gets the function, called Hamiltonian :

$H : I \times F \times F^* \to \mathbb{R} :: H(t, q_1, ..., q_n, p_1, ...p_n) = \sum_{i=1}^n p_i z^i - L$

Then :

$dH = \sum_{i=1}^n z^i dp_i - \sum_{i=1}^n \frac{\partial \lambda}{\partial y^i} dq_i - \frac{\partial \lambda}{\partial t}$

$\frac{\partial H}{\partial t} = -\frac{\partial \lambda}{\partial t}, \frac{\partial H}{\partial q_i} = -\frac{\partial \lambda}{\partial y^i}, \frac{\partial H}{\partial p_i} = z^i$

and the Euler-Lagrange equations read :

$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}$

which is the Hamiltonian formulation.

If λ does not depend on t then the quantity $H = Ct$ for the solutions.

32.3.3 Variational problem on a fibered manifold

Given a Lagrangian $\mathcal{L} : J^r E \to \Lambda_m (M; \mathbb{R})$ on any fibered manifold the problem is to find a section $X \in \mathfrak{X}(E)$ such that the integral $\ell(Z) = \int_M \mathcal{L}(Z)$ is extremum.

Stationary solutions

The first step is to define the extremum of a function.

If E is not a vector bundle the general method to define a derivative is by using the Lie derivative.

The - clever - idea is to "deform" a section X according to a some rules that can be parametrized, and to look for a stationary state of deformation, such that in a neighborhood of some X_0 any other transformation does not change the value of $\ell(J^r X)$. The focus is on one parameter groups of diffeomorphisms, generated by a projectable vector field W on E.

It defines a base preserving map : $\mathfrak{X}(J^r E) \to \mathfrak{X}(J^r E) :: J^r U(t) Z(x) = \Phi_{J^r W}(Z(x), t)$. One considers the changes in the value of $\lambda(Z)$ when Z is replaced by $J^r U(t) Z : \lambda(Z) \to \lambda(J^r U(t) Z)$ and the Lie derivative of λ along W is : $\pounds_W \ell = \frac{d}{dt} \ell(J^r U(t) Z)|_{t=0}$. Assuming that λ is differentiable with respect to Z at Z_0 : $\pounds_W \lambda = \lambda'(Z_0) \frac{d}{dt} J^r U(t) Z|_{t=0} = \lambda'(Z_0) \pounds_{J^r W} Z$. The derivative $\lambda'(Z_0)$ is a continuous linear map from the tangent bundle $T J^r E$ to $\mathbb{R}$ and $\pounds_{J^r W} Z$ is a section of the vertical bundle $V J^r E$. So $\pounds_W \lambda$ is well defined and is a scalar.

$\pounds_W \mathcal{L}(X) = \frac{\partial}{\partial t} \mathcal{L}(J^r U(t) J^r X)|_{t=0} = \pounds_W \lambda d\xi^1 \wedge ... \wedge d\xi^m$ is called the **variational derivative** of the Lagrangian $\mathcal{L}$. This is the Lie derivative of the natural operator (Kolar p.387) : $\mathcal{L} : J^r \hookrightarrow \Lambda_m$:

$\pounds_W \mathcal{L}(p) = \pounds_{(J^r W, \Lambda_m T M^* Y)} \mathcal{L}(p) = \frac{\partial}{\partial t} \Phi_{T M^* Y}(\mathcal{L}(\Phi_{J^r W}(p, -t)), t)|_{t=0}$: $J^r Y \to \Lambda_m T M^*$

A section $X_0 \in \mathfrak{X}(E)$ is a **stationary solution** of the variational problem if $\int_M \pounds_W \mathcal{L}(X) = 0$ for any projectable vector field.

The problem is that there is no simple formulation of $\pounds_{J^r W} Z$.

Indeed : $\pounds_{J^r W} Z = -\frac{\partial \Phi_{J^r W}}{\partial Z}(Z, 0) \frac{dZ}{dx} Y + J^r W(Z) \in V_p E$ where Y is the projection of W on TM (see Fiber bundles).

For $Z = J^r X, X \in \mathfrak{X}(E)$

$\Phi_W(J^r X(x), t) = j^r_{\Phi_Y(x,t)}(\Phi_W(X(\Phi_Y(x, -t)), t))$

$= (D_{\alpha_1..\alpha_s} \Phi_W(X(\Phi_Y(x, -t)), t), s = 0..r, 1 \leq \alpha_k \leq m)$

$\frac{\partial \Phi_{J^r W}}{\partial Z}(J^r X, 0) = (D_{\alpha_1..\alpha_s} \Phi_W(X, 0), s = 0..r, 1 \leq \alpha_k \leq m)$

$\pounds_{J^r W} J^r X = J^r W(J^r X) - \sum_{s=0}^r \sum_{\alpha_1...\alpha_s=1}^m (D_{\alpha_1..\alpha_s} \Phi_W(X, 0)) \left(\sum_\beta Y^\beta D_{\beta \alpha_1..\alpha_s} X \right)$

The solution, which is technical, is to replace this expression by another one, which is "equivalent", and gives someting nicer when one passes to the integral.

The Euler-Lagrange form

Theorem 2497 *(Kolar p.388) For every r order lagrangian $\mathcal{L} : J^r E \to \Lambda_m(M; \mathbb{R})$ there is a morphism $K(\mathcal{L}) : J^{2r-1} E \to V J^{r-1} E^* \otimes \Lambda_{m-1} T M^*$ and a unique*

morphism : $\mathfrak{E}(\mathcal{L}) : J^{2r}E \to VE^* \otimes \Lambda_m TM^*$ such that for any vertical vector field W on E and section X on E: $\pounds_W \mathcal{L} = \mathfrak{D}\left(K(\mathcal{L})\left(J^{r-1}W\right)\right) + \mathfrak{E}(\mathcal{L})(W)$

$\mathfrak{D}$ is the total differential (see below).

The morphism $K(\mathcal{L})$, called a **Lepage equivalent** to $\mathcal{L}$, reads:

$K(\mathcal{L}) = \sum_{s=0}^{r-1} \sum_{\beta=1}^{m} \sum_{i=1}^{n} \sum_{\alpha_1 \leq .. \leq \alpha_s} K_i^{\beta \alpha_1 .. \alpha_s} d\eta_{\alpha_1 .. \alpha_s}^i \otimes d\xi^1 \wedge .\widehat{\wedge d\xi^\beta}. \wedge d\xi^m$

It is not uniquely determined with respect to $\mathcal{L}$. The mostly used is the **Poincaré-Cartan equivalent** $\Theta(\lambda)$ defined by the relations :

$\Theta(\lambda) = \left(\lambda + \sum_{\alpha=1}^{m} \sum_{s=1}^{r-1} K_i^{\alpha\beta_1 ... \beta_s} y_{\beta_1 .. \beta_s}^i\right) \wedge d\xi^1 \wedge .. \wedge d\xi^m$

$K_i^{\beta_1 ... \beta r+1} = 0$

$K_i^{\beta_1 ... \beta s} = \frac{\partial\lambda}{\partial\eta_{\beta_1 .. \beta_s}^i} - \sum_\gamma d_\gamma K_i^{\gamma\beta_1 ... \beta s}, s = 1..r$

$y_{\beta_1 .. \beta_s}^i = d\eta_{\beta_1 .. \beta_s}^i - \sum_\gamma \eta_{\beta_1 .. \beta_s \gamma}^i d\xi^\gamma$

$y_\beta = d\xi^1 \wedge .\widehat{\wedge d\xi^\beta}. \wedge d\xi^m$

It has the property that : $\lambda = h(\Theta(\lambda))$ where h is the horizontalization (see below) and $h\left(\Theta\left(J^{r+1}X(x)\right)\right) = \Theta(J^r X(x))$

The **Euler-Lagrange form** $\mathfrak{E}(\mathcal{L})$ is :

$\mathfrak{E}(\mathcal{L}) = \sum_{i=1}^{n} \mathfrak{E}(\mathcal{L})_i \, du^i \wedge d\xi^1 \wedge .. \wedge d\xi^m$

$\mathfrak{E}(\mathcal{L})_i = \sum_{s=0}^{r} (-1)^s \sum_{\alpha_1 \leq .. \leq \alpha_s} d_{\alpha_1} d_{\alpha_2} .. d_{\alpha_s} \frac{\partial\lambda}{\partial\eta_{\alpha_1 .. \alpha_s}^i}$ where d_α is the total differentiation

$d_\alpha f = \frac{\partial f}{\partial\xi^\alpha} + \sum_{s=1}^{r} \sum_{\beta_1 \leq ... \leq \beta_s} \frac{\partial f}{\partial\eta_{\alpha\beta_1 ... \beta_s}^i} \eta_{\alpha\beta_1 ... \beta_s}^i$

$\mathfrak{E}(\mathcal{L})$ is a linear natural operator which commutes with the Lie derivative :

Theorem 2498 *(Kolar p.390) For any projectable vector field W on E : $\pounds_W \mathfrak{E}(\mathcal{L}) = \mathfrak{E}(\pounds_W \mathcal{L})$*

Solutions

The solutions are deduced from the previous theorem.

The formula above reads :

$J^{r-1}W$ is a vertical vector field in $VJ^{r-1}E$:

$J^{r-1}W = W^i \partial u_i + \sum_{s=1}^{r-1} \sum_{\alpha_1 \leq ... \leq \alpha_s} \left(d_{\alpha_1} d_{\alpha_2} ... d_{\alpha_s} W^i\right) \partial u_i^{\alpha_1 ... \alpha_s}$

so $F = K(\mathcal{L})\left(J^{r-1}W\right) = \sum_{\alpha=1}^{m} F_\alpha d\xi^1 \wedge .\widehat{\wedge d\xi^\beta}.. \wedge d\xi^m \in \Lambda_{m-1} TM^*$

$\mathfrak{D}$ is the total differential :

$\mathfrak{D}F = \sum_{\alpha,\beta=1}^{m} (d_\alpha F_\beta) d\xi^\alpha \wedge d\xi^1 \wedge .\widehat{\wedge d\xi^\beta}. \wedge d\xi^m = \sum_{\alpha,\beta=1}^{m} (-1)^{\alpha-1} (d_\alpha F_\alpha) d\xi^1 \wedge ... \wedge d\xi^m$

with : $d_\alpha F = \frac{\partial F}{\partial\xi^\alpha} + \sum_{s=1}^{r} \sum_{\beta_1 \leq .. \leq \beta_s} \frac{\partial F}{\partial\eta_{\beta_1 ... \beta_s}^i} \eta_{\alpha\beta_1 ... \beta_s}^i$

$\mathfrak{E}(\mathcal{L})(W) = \sum_{i=1}^{n} \mathfrak{E}(\mathcal{L})_i W^i d\xi^1 \wedge ... \wedge d\xi^m$

$= \sum_{i=1}^{n} \sum_{s=0}^{r} (-1)^s \sum_{\alpha_1 \leq .. \leq \alpha_s} W^i \left(d_{\alpha_1} d_{\alpha_2} .. d_{\alpha_s} \frac{\partial\lambda}{\partial\eta_{\alpha_1 .. \alpha_s}^i}\right) d\xi^1 \wedge ... \wedge d\xi^m$

The first term $\mathfrak{D}\left(K(\mathcal{L})\left(J^{r-1}W\right)\right)$ is the differential of a m-1 form : $\mathfrak{D}\left(K(\mathcal{L})\left(J^{r-1}W\right)\right) = d\mu$

The second reads : $\mathfrak{E}(\mathcal{L})(W) = i_W \mathfrak{E}(\mathcal{L})$

So :

$\ell(X) = \int_M (d\mu + i_W \mathfrak{E}(\mathcal{L}))$

Any open subset O of M, relatively compact, gives a manifold with boundary and with the Stokes theorem :

$\ell(X) = \int_{\partial O} \mu + \int_M i_W \mathfrak{E}(\mathcal{L})$

It shall hold for any projectable vector field W. For W with compact support in O the first integral vanishes. We are left with : $\ell(X) = \int_M i_W \mathfrak{E}(\mathcal{L})$ which is linearly dependant of W^i. So we must have : $J^{2r} X^* \mathfrak{E}(\mathcal{L}) = 0$

Theorem 2499 *A section $X \in \mathfrak{X}_{2r}(E)$ is a stationary solution of $\ell(J^r X)$ only if $J^{2r} X^* \mathfrak{E}(\mathcal{L}) = 0$*

Notice that there is no guarantee that the solution is a maximum or a minimum, and it may exist "better solutions" which do not come from a one parameter group of morphisms.

The Euler Lagrange equations read :

$\sum_{k=0}^{r} (-1)^k d_{\alpha_1} d_{\alpha_2} ... d_{\alpha_k} \frac{\partial L}{\partial \eta^i_{\alpha_1 .. \alpha_k .}} (j^{2r} X) = 0; i = 1...n$

So they are exactly the same as the equations that we have found for a vectorial bundle with the functional derivatives.

32.3.4 Noether currents

Principle The symmetries in the model provided by a lagrangian are of physical great importance. They can resort to different categories : they can be physical (ex: spherical or cylindrical symmetry) so depending on the problem (and the model should account for them), or be gauge symmetries. They can be modelled by an automorphism on E : $G : E \rightarrow E$ such that : $\forall X \in \mathfrak{X}(E) : \mathcal{L}(J^r(G(X))) = \mathcal{L}(J^r X)$. Similarly they keep inchanged the scalar lagrangian. Usually they are studied by one parameter groups of automorphisms, parametrized by their infinitesimal generator, which is a vector field, but not necessarily a vector field on E.

Classical symmetries

Definition 2500 *A vector field $W_r \in \mathfrak{X}(TJ^r E)$ is said to be the generator of a (classical, exact) symmetry of $\mathcal{L}$ if : $\forall Z \in J^r E : \pounds_{W_r} \pi^{r*} \mathcal{L}(Z) = 0$*

Some explanations...

i) $\pi^{r*} \mathcal{L}(Z)$ is the pull back of a lagrangian $\mathcal{L}$ by the projection : $\pi^r : J^r E \rightarrow M$. This is a horizontal m form on $J^r E$, such that :

$\forall W^k \in \mathfrak{X}(TJ^r E), k = 1...m : \pi^{r*} \mathcal{L}(Z)(W^1, .., W^m)$

$= \mathcal{L}(\pi^r(Z))(\pi^r(Z)' W^1, ...\pi^r(Z)' W^m)$

ii) $\pounds_{W_r}$ is the Lie derivative (in the usual meaning on the manifold $J^r E$) defined by the flow Φ_{W_r} of a vector field $W_r \in \mathfrak{X}(TJ^r E)$ defines, by its flow $\Phi_{W_r} : \frac{\partial}{\partial t} \pi^{r*} \mathcal{L}(Z) \Phi_{W_r}(X(\Phi_{W_r}(Z, -t)), t) = \pounds_{W_r} \pi^{r*} \mathcal{L}(Z)$

W_r is not necessarily projectable on TM.

So the set of such vectors W_r provides a large range of symmetries of the model. It has the structure of a vector space and of Lie algebra with the commutator (Giachetta p.70).

Theorem 2501 *First Noether theorem : If W_r is a classical symmetry for $\mathcal{L}$ then $\mathfrak{D}\left(K\left(\mathcal{L}\right)\left(J^{r-1}W_r\right)\right) = 0$ for the solutions*

It is common to say that a property is satisfied "on the shell" when it is satisfied for the solutions of the Euler-Lagrange equations. As a consequence the quantity called a Noether current $\mathfrak{I} = K\left(\mathcal{L}\right)\left(J^{r-1}W_r\right) = \sum_\alpha \mathfrak{I}_\alpha d\xi^1 \wedge .\wedge \widehat{d\xi^\alpha}. \wedge d\xi^m$ is conserved on the shell.

Theorem 2502 *Noether-Bessel-Hagen theorem (Kolar p.389): A projectable vector field W on E is a generalized infinitesimal automorphism of the r order lagrangian $\mathcal{L} : J^r E \to \Lambda_m\left(M;\mathbb{R}\right)$ iff $\mathfrak{E}\left(\pounds_W \mathcal{L}\right) = 0$*

Gauge symmetries

1. Gauge symmetries arise if there is a principal bundle $P\left(M,G,\pi_P\right)$ over the same manifold as E, and the group G is the generator of local symmetries on E : $\Psi\left(g\left(x\right)\right) : E\left(x\right) \to E\left(x\right)$

Examples : a change of gauge on P defined by a map : $g : M \to G$ entails the following symmetries :

on the vector bundle $E\left(M, T_1 G \times TM^*, \pi\right)$ of potentials $\grave{A} = \grave{A}^i_\alpha \kappa_i \otimes dx^\alpha$ of a connection :
$$\Psi\left(g\left(x\right)\right)\left(\grave{A}\left(x\right)\right) = Ad_{g(x)}\left(\grave{A}\left(x\right) - L'_{g(x)^{-1}}(g\left(x\right))g'\left(x\right)1\right)$$

on the vector bundle $E\left(M, T_1 G \times \Lambda_2 TM^*, \pi\right)$ of the strength $\mathcal{F} = \mathcal{F}^i_{\alpha\beta}\kappa_i \otimes dx^\alpha \wedge dx^\beta$ of a connection :
$$\Psi\left(g\left(x\right)\right)\left(\grave{A}\left(x\right)\right) = Ad_{g(x)^{-1}}\left(\mathcal{F}\left(x\right)\right)$$

2. The adjoint bundle $P\left[T_1 G, Ad\right]$ is a vector bundle, and the generators of one parameter groups of local symmetries can be defined by a vector field $\kappa : M \to P\left[T_1 G, Ad\right] :: g\left(x,t\right) = \exp t\kappa\left(x\right)$

As the example of the potential shows, a symmetry Ψ can be a differential operator : $\Psi\left(J^k g\left(t\right)\right) = \Psi\left(\frac{\partial g(x,t)}{\partial \xi^{\alpha_1}..\partial \xi^{\alpha_p}}, p = 0,..k\right)$. However the derivative at t=0 gives a more practical formula :
$$\frac{\partial}{\partial t}\Psi\left(J^k g\left(t\right)\right)\left(X\right)|_{t=0} = \sum_{p=0}^k \sum_{\alpha_1..\alpha_p} \frac{\partial \Psi}{\partial \gamma_{\alpha_1...\alpha_p}} \frac{\partial \gamma_{\alpha_1...\alpha_p}}{\partial t}|_{t=0}\left(X\right)$$
$$= \sum_{p=0}^k \sum_{\alpha_1..\alpha_p} \frac{\partial \Psi}{\partial \gamma_{\alpha_1...\alpha_p}} \kappa^a_{\alpha_1..\alpha_r}\left(X\right)$$
where
$\gamma_{\alpha_1...\alpha_p} = \frac{\partial g(x,t)}{\partial \xi^{\alpha_1}..\partial \xi^{\alpha_p}}$
$\left(\kappa^a_{\alpha_1..\alpha_r}\right) = J^k \kappa \in \mathfrak{X}\left(J^k P\left[T_1 G, Ad\right]\right)$.
Moreover the derivatives $\frac{\partial \Psi}{\partial \gamma_{\alpha_1...\alpha_p}}\left(X\right)$ are evaluated at t=0, then g=1.

This an operator : $\widetilde{\Psi}\left(J^k \kappa\right) : E \to E$ wich is linear in $J^k \kappa$

Proceeding similarly for each of the components of $J^r X$ on gets an operaor : $\widetilde{\Psi}\left(J^k \kappa\right) : J^r E \to J^r E$

3. The lagrangian must be such that :
$$\forall J^k \kappa \in \mathfrak{X}\left(J^k P\left[T_1 G, Ad\right]\right) : \lambda\left(\widetilde{\Psi}\left(J^k \kappa\right)\left(J^r X\right)\right) = \lambda\left(J^r X\right)$$

By differentiation with respect to each component of $J^k \kappa$ one gets a set of identies that the lagangian must meet.

An example of this method can be found in (Dutailly).

32.4   The variationnal bicomplex

The spaces of p forms on r prolongation of fibered manifolds can be endowed with algebraic structures which are actively studied, in the frame work of variational calculus as well as algebraic topology. We give just a brief introduction on the subject. We use the same notations as above.

32.4.1   Infinite jet prolongation

If M, V and the trivializations are smooth a r jet can be extended to $r = \infty$. The infinite jet prolongation $J^\infty E$ is a Fréchet space and not a true manifold.

The projections yield a direct sequence of forms on the jet prolongations :

$\mathfrak{X}(T_p M^*) \overset{\pi^*}{\to} \mathfrak{X}(\Lambda_p TE^*) \overset{\pi_0^{1*}}{\to} \mathfrak{X}(\Lambda_p TJ^1 E^*) \overset{\pi_1^{2*}}{\to} \mathfrak{X}(\Lambda_p TJ^1 E^*) \ldots$

with the pull back of forms :

$\varpi_p \in \mathfrak{X}(\Lambda_p TJ^{r-1} E^*), W_r \in \mathfrak{X}(TJ^r E) : \pi^{r*}_{r-1} \varpi_p(Z_r) W_r$
$= \varpi_p(\pi^r_{r-1}(Z_r)) \pi^r_{r-1}(Z_r)' W_r$

In particular a p form λ on M defines a p horizontal form $\pi^{r*}\varpi_p$ on $J^r E$

It exists a direct limit, and p formscan be defined on $J^\infty E$ with the same operations (exterior differential and exterior product).

Any closed form ϖ on $J^\infty E$ can be decomposed as : $\varpi = \mu + d\lambda$ where μ is closed on E.

32.4.2   Contact forms

A key point, in order to understand what follows, is to keep in mind that, as usual in jet formalism, a point of $J^r E$ does not come necessarily from a section on E, which assumes that the components are related. Indeed the prolongation $J^r X$ of a section X of E imposes the relations between the coordinates : $z^i_{\alpha_1..\alpha_s} = \frac{\partial^s \sigma^i}{\partial \xi^{\alpha_1}..\partial \xi^{\alpha_s}}|_x$ and the image of M by the map $J^r X$ is a subset of $J^r E$. So a p form on $J^r E$ takes specific values $\varpi(Z)$ when it is evaluated at the r jet prolongation $Z = J^r X(x)$ of a section of E and there is a map : $M \to \Lambda_p TJ^r E^*::\varpi(J^r X(x)) = J^r X^* \varpi(x)$ with the usual notation for the pull back of a map. Notice that this is not the pullback to M of the p form on $J^r E$.

Contact forms

To fully appreciate this remark there are non zero forms on $J^r E$ which are identically null when evaluated by the prolongation of a section. This can be understood : in accounting for the relations between coordinates in $du^i_{\alpha_1...\alpha_s}$ the result can be 0. This leads to the following definitions :

Definition 2503 *(Vitolo) A p form on $J^r E : \varpi \in \Lambda_p J^r E$ is said to be contact if $\forall X \in \mathfrak{X}_r(E) : J^r X^* \varpi = 0$. It is said to be a k-contact form if it generated by k powers (with the exterior product) of contact forms.*

Any p form with p > m is contact. If ϖ is contact then $d\varpi$ is contact. A 0 contact form is an ordinary form.

The space of 1-contact p-forms is generated by the forms :

$\varpi^i_{\beta_1..\beta_s} = d\eta^i_{\beta_1..\beta_s} - \sum_\gamma \eta^i_{\beta_1..\beta_s\gamma} dx^\gamma, s = 0..r-1$ and $d\varpi^i_{\beta_1..\beta_{r-1}}$

The value of a contact form vanishes in any computation using sections on E. So they can be discarded, or conversely, added if it is useful.

Horizontalization

Theorem 2504 *(Vitolo) There is a map, called (p,q) horizontalisation, such that:* $h_{(k,p-k)} : \mathfrak{X}\left(T_p J^r E^*\right)^k \to \mathfrak{X}\left(T_k J^{r+1} E^*\right)^k \wedge \mathfrak{X}^H\left(T_{p-k} J^{r+1} E^*\right)$ *where* $\mathfrak{X}\left(T_p J^{r+1} E^*\right)^k$ *is the set of k contacts q forms.*

The most important is the map with k=0 usually denoted simply h. It is a morphism of exterior algebras (Krupka 2000).

So any p form can be projected on a horizontal p form : the result is : $h_{(0,p)}(\varpi) = \Omega_{\alpha_1...\alpha_p} dx^{\alpha_1} \wedge ... \wedge dx^{\alpha_p}$ where Ω is a combination of the components of ϖ. Horizontal forms give : $h_{(0,p)}(\varpi) = \varpi$

The most important properties of horizontalization are :

$\forall \varpi \in \mathfrak{X}\left(T_p J^r E^*\right), X \in \mathfrak{X}(E) : J^r X^* \varpi = J^{r+1} X^* h_{(0,p)}(\varpi)$

$\ker h_{(k,p-k)} = \mathfrak{X}\left(T_p J^r E^*\right)^{k+1}$

32.4.3 Variational bicomplex

(Vitolo)

1. The de Rahm complex (see cohomologie) is defined by the sequence $0 \to \Lambda_0 TM^* \xrightarrow{d} \Lambda_1 TM^* \xrightarrow{d} \Lambda_2 TM^* \xrightarrow{d} ....$ There is something similar, but with two dimensions (so it is called a bicomplex) and two differential operators.

The variational bicomplex uses the maps :

$i_H : \Lambda_p J^r E \to \Lambda_p J^{r+1} E :: i_H = i_{D^{r+1}} \circ \left(\pi_r^{r+1}\right)^*$

$i_V : \Lambda_p J^r E \to \Lambda_p J^{r+1} E :: i_V = i_{\varpi^{r+1}} \circ \left(\pi_r^{r+1}\right)^*$

$d_H : \Lambda_p J^r E \to \Lambda_p J^{r+1} E :: d_H = i_H \circ d - d \circ i_H$

$d_V : \Lambda_p J^r E \to \Lambda_p J^{r+1} E :: d_V = i_V \circ d - d \circ i_V$

with :

$D^{r+1} = \sum_\gamma dx^\gamma \otimes \left(\partial \xi_\gamma + \sum_{s=1}^r \sum_{\alpha_1 \leq ... \leq \alpha_s} \eta^i_{\gamma \alpha_1...\alpha_s} \partial u^i_{\alpha_1...\alpha_s}\right)$

$\varpi^{r+1} = \sum_\gamma \left(\partial u^i_{\alpha_1...\alpha_s} - \eta^i_{\gamma \alpha_1...\alpha_s} dx^\gamma\right) \otimes \partial u^i_{\alpha_1...\alpha_s}$

They are fully defined through their action on functions :

$d_H f = \sum_\gamma \left(\frac{\partial f}{\partial \xi^\gamma} + \sum_{s=1}^r \sum_{\alpha_1 \leq ... \leq \alpha_s} \eta^i_{\gamma \alpha_1...\alpha_s} \frac{\partial f}{\partial \eta^i_{\alpha_1...\alpha_s}}\right) dx^\gamma$

$d_H dx^\alpha = 0, d_H\left(du^i_{\alpha_1...\alpha_s}\right) = -\sum_\beta du^i_{\beta \alpha_1...\alpha_s} \wedge dx^\beta,$

$d_V f = \sum_{\alpha_1..\alpha_s} \frac{\partial f}{\partial \eta^i_{\alpha_1...\alpha_s}} \varpi^i_{\alpha_1...\alpha_s}$

$d_V dx^\alpha = 0, d_V\left(du^i_{\alpha_1...\alpha_s}\right) = \sum_\beta du^i_{\beta \alpha_1...\alpha_s} \wedge dx^\beta$

with the properties :

$d_H^2 = d_V^2 = 0$

$d_H \circ d_V + d_V \circ d_H = 0$

$d_H + d_V = \left(\pi_r^{r+1}\right)^* d$

$\left(J^{r+1}X\right)^* \circ d_V = 0$

$d \circ (J^r X)^* = \left(J^{r+1}X\right)^* \circ d_H$

$d_H = \mathfrak{D}$ the total external differentiation used previously.

On the algebra $\Lambda J^\infty E$ these relations simplify a bit : $d_H + d_V = d$ and the Lie derivatives : $\pounds_{d_\alpha} = d_\alpha \circ d + d \circ d_\alpha$ are such that : $\pounds_{d_\alpha} (\varpi \wedge \mu) = (\pounds_{d_\alpha} \varpi) \wedge \mu + \varpi \wedge \pounds_{d_\alpha} \mu$

2. On this algebra the space of p forms can be decomposed as follows :

$\mathfrak{X}(T_p J^r E^*) = \oplus_k \mathfrak{X}(T_k J^r E^*)^k \wedge h_{(0,p-k)} \left(\mathfrak{X}(T_{p-k} J^r E^*)^0\right)$

The first part are contact forms (they will vanish with sections on E), the second part are horizontal forms.

Then with the spaces :

$F_0^{0q} = h_{(0,q)} \left(\mathfrak{X}(\Lambda_q T J^r E^*)^0\right), F_0^{pq}$

$= \mathfrak{X}(T_p J^r E^*)^p \wedge h_{(0,q)} \left(\mathfrak{X}(T_q J^r E^*)^0\right), F_1^{pn} = F_0^{pn}/d_H \left(F_0^{pn-1}\right)$

one can define a bidirectional sequence of spaces of forms, similar to the de Rahm cohomology, called the variational bicomplex, through which one navigates with the maps d_H, d_V.

The variational sequence is :

$0 \to \mathbb{R} \to F_0^{00} \overset{d_H}{\to} ...F_0^{0n-1} \overset{d_H}{\to} F0+_0^{0n} \to ...$

3. The variational bicomplex is extensively used to study symmetries. In particular :

Theorem 2505 *(Vitolo p.42) A lagrnagian $\mathcal{L} : J^r E \to \wedge_m TM^*$ defines by pull back a horizontal m form $\varpi = (\pi^r)^* \mathcal{L}$ on $J^r E$. If $d_V \varpi = 0$ then $\exists \mu \in h^{0m-1}(\nu), \nu \in T_{m-1} J^{r-1} E^*$ such that : $\varpi = d_H \mu$.*

Theorem 2506 *First variational formula (Vitolo p.42) For $W \in \mathfrak{X}(TE), \varpi \in F_1^{pm}$ the identity :$\pounds_W \varpi = i_W d_V \varpi + d_V (i_W \varpi)$ holds*

It can be seen as a generalization of the classic identity : $\pounds_W \varpi = i_W d\varpi + d(i_W \varpi)$

Part VIII
BIBLIOGRAPHY

A.Axelsson, A.McIntosh *Hodge decomposition on weakly Lipschtiz domain* Internet paper (1991)

C.Berge *Espaces topologiques* Dunond (1965)

N.Bourbaki *General topology* Addison-Wesley (1968)

O.Bratteli, D.W.Robinson *Operators algebras and quantum mechanics 1* 2nd edition Springer (1987)

J.L.Doob *Measure theory* Springer (1994)

J.J.Duistermaat, J.A.C.Kolk *Lie groups* Springer (1999) sep

J.C.Dutailly *Yang Mills model of interactive particles in the classical field theory* arXiv:1109.3897v1 [math-ph] 18 sep 2011

W.Fulton, J.Harris *Representation theory* Springer (1991)

Th.W.Gamelin, R.E.Greene *Introduction to topology* 2nd edition Dover (1999)

G.Giachetta, L.Mangiarotti,,G.Sardanashvily *Advanced classical field theory* World Scientific (2009)

A.Gregor'yan *Heat kernel on weighted manifolds and applications* (2006)

H.Hofer,E.Zehnder *Symplectic invariants and hamiltonian dynamics* Birkhaüser (1994)

D.Husemoller *Fiber bundle* 3d Edition Springer (1994)

M.Kashiwara, P.Shapira *Categories and sheaves* Springer (2006)

S.C.Kleene *Mathematical logic* Wiley (1967)

A.W.Knapp *Lie groups : beyond an introduction (2nd edition)* Birkhäuser (2005)

A.W.Knapp *Representation theory of semi-simple Lie groups* Princeton landmarks (1986)

E.H.Lieb, M.Loss *Analysis* 2nd edition American Mathematical Society (2001)

I.Kolàr,P.Michor,J.Slovak *Natural operations in differential geometry* Springer (1993)

Y.Kosmann-Schwarzbach *Groupes et symétries* Editions de l'Ecole Polytechnique (2006)

D.Krupka *On the local structure of the Euler-Lagrange mapping of the calculus of variations* Mathematical Institute Opava (2002)

D.Krupka *Some geometric aspects of variational problems in fibered manifolds* Mathematical Institute Opava (2001)

D.Krupka, M.Krupka *Jets and contacts elements* Proceedings of the seminar on differential geometry Opava (2000)

O.Krupkova, G.E.Prince *Second order differential equations in jets bundles and the inverse problem of the calcul of variations* Handbook of global analysis Elsevier (2007)

J.Lafontaine *Introduction aux variétés différentiables* EDP (1996)

S.M.Lane *Categories for the working mathematician 2nd edition* Springer (1969)

A.D.Lewis *Affine connections and distributions* Reports on mathematical physics 42 (1/2) 125-164 (1998)

D.Lovelock,H.Rund *Tensors, differential forms and variational principles* Dover (1989)

P.Malliavin *Géométrie différentielle intrinsèque* Hermann (1972)

J.Munkres *Topology* (2nd ed) Pearson (2003)

M.Nakahara *Geometry, topology and physics* 2nd edition IoP (2003)

K.H.Neeb *An introduction to unitary representations of Lie groups* Internet Paper (june 2 2010)

E.Nigsch *Colombeau generalized functions on Manifolds* Technische Universität Wien

A.Polyanin, V.Zaitsev *Handbook of exact solutions for ordinary differential equations* Chapman & Hall (2003)

A.Polyanin, V.Zaitsev *Handbook of linear partial differential equations* for engineers Chapman & Hall (2003)

A.Polyanin, V.Zaitsev *Handbook of non linear partial differential equations* Chapman & Hall (2004)

L.Schwartz *Analyse* 4 tomes Hermann (1997)

G.Svetlichny *Preparation for gauge theories* arXiv:math-ph 9902027 v3 12 march 1999

M.E.Taylor *Partial differential equations* 3 volumes Springer (1999)

M.Thill *Introduction to normed *-algebras and their representations 7th edition* arXiv 1011-1558v1 [math.0A] 6 nov.2010

M.Thill *Representation of hermitian commutative *-algebras by unbounded operators* arXiv 0908-3267v2 [math.0A] 20 nov.2009

R.Vitolo *Variational sequences* (http:\\poincare.unile.it) (2007)

R.M.Wald *General relativity* University of Chicago Press (1984)

A.Wilansky *Topology for analysis* Dover (2008)

C.Zuily *Eléments de distributions et d'équations aux dérivées partielles* Dunod (2002)