MATHEMATICS IN PHYSICS

Jean Claude Dutailly

MATHEMATICS IN PHYSICS

Jean Claude Dutailly

2 February 2016

Contents

ABSTRACT

This book proposes a review and, on some important points, a new interpretation of the main concepts of Theoretical Physics. Rather than offering an interpretation based on exotic physical assumptions (additional dimension, new particle, cosmological phenomenon,...) or a brand new abstract mathematical formalism, it proceeds to a systematic review of the main concepts of Physics, as Physicists have always understood them : space, time, material body, force fields, momentum, energy... and propose the right mathematical tools to deal with them, chosen among well known mathematical theories.

After a short introduction about the place of Mathematics in Physics, a new interpretation of the main axioms of Quantum Mechanics is proposed. It is proven that these axioms come actually from the way mathematical models are expressed, and this leads to theorems which validate most of the usual computations and provide safe and clear conditions for their use, as it is shown in the rest of the book.

Relativity is introduced through the construct of the Geometry of General Relativity, based on 5 propositions and the use of tetrads and fiber bundles, which provide tools to deal with practical problems, such as deformable solids.

A review of the concept of momenta leads to the introduction of spinors in the framework of Clifford algebras. It gives a clear understanding of spin and antiparticles.

The force fields are introduced through connections, in the, now well known, framework of gauge theories, which is here extended to the gravitational field. It shows that this field has actually a rotational and a transversal component, which are masked under the usual treatment by the metric and the Levy-Civita connection. A thorough attention is given to the topic of the propagation of fields with interesting results, notably to explore gravitation.

The general theory of lagrangians in the application of the Principle of Least Action is reviewed, and two general models, incorporating all particles and fields are explored, and used for the introduction of the concepts of currents and energy-momentum tensor. Precise guidelines are given to find operational solutions of the equations of the gravitational field in the most general case.

The last chapter shows that bosons can be understood as discontinuities in the fields.

In this 4th version of this book, changes have been made :

- in Relativist Geometry : the ideas are the same, but the chapter has been rewritten, notably to introduce the causal structure and explain the link with the practical measures of time and space;
- in Spinors : the relation with momenta has been introduced explicitly
- in Force fields : the section dedicated to the propagation of fields is new, and is an important addition.
- in Continuous Models : the section about currents and energy-momentum tensor are new.
- in Discontinuous Processes : the section about bosons has been rewritten and the model improved.

Introduction

A century after the introduction of Relativity and Quantum Mechanics these theories are still the topic of hot controversies. If this is not a serious concern for most physicists, there is no doubt that this is the symptom of a weakness in the tools that they use daily, and is at the core of the lack of progress in Theoretical Physics. The attempts to solve the conundrum have been directed in two directions. Most of them are based on assuming new properties to the physical world, either additional dimensions, special behavior of particles of fields, new particles, alternate universe,... Others on the introduction of new formal systems, involving highly abstract concepts, to base the computational practices on firmer ground. In this book I propose another way : I keep the objects of conventional physics (particles, fields, space, time,...) and their properties, as almost all physicists understand them (so there are 4 dimensions, no special particle or field, no GTU,...), but I focus on their representation by mathematical objects. It is clear that we cannot do the Physics of the XXI° century with the Mathematics of the XIX° century - when Einstein introduced Relativity, Differential Geometry was in infancy - but conversely the race to the introduction of new mathematical tools is void if it does not go with a deep understanding of their adequation with the physical concepts that they are supposed to represent. It is then possible to find from where comes Quantum Mechanics, to show that its main axioms are actually theorems which can be proven, to introduce in a simple and elegant way General Relativity, spinors, the like wave representation of particles, and conversely the particles like representation of force fields, and to work with them and the usual principles of Physics, in a simple and clear way. With the concepts firmly grounded and the adequate tools it is then possible to go further. The main new results presented here are related to gravitation, which appears to have a rotational and a transversal component, with different interpretation, and whose pull can vary with the speed of material bodies, opening a more rational explanation to the anomalies of the motion of stars than the introduction of dark matter.

So the first purpose of this book is to show how mathematical tools can be used to represent rigorously and efficiently some of the key concepts of modern Physics, in Relativity and Quantum Mechanics, and through this representation, show how these concepts, and physical laws which seem often paradoxical, can be clearly understood. The book addresses both Mathematicians who want to use their knowledge to get an overview of some of the main issues of modern physics, and Physicists who want to improve their grasp of mathematical tools in their work. Its purpose is to help graduated and undergraduate students to understand the place and the role of mathematics in Physics. Why do we have Hilbert spaces in Quantum Mechanics ? What is the meaning of probabilist representations ? How physical concepts such as motion, moment, matter or force fields, are related to mathematical objects ? How physical principles, such as the Principle of Relativity or the Principle of Least Action, can be translated into Mathematics and be efficiently used in Physics ? How can we understand the entanglement of particles ? What is a spinor ?

In this book the reader will see how to deal with manifolds, fiber bundles, connections, Clifford algebras, group representations, generalized functions or Lagrange equations. There are many books which deal with these topics, usually for physicists, with the purpose to make understandable in a nut shell what are, after all, some of the most abstract parts of Mathematics. We will not choose this path, not by some pedantic pretense, but because for a scientist the most general approach, which requires few but key concepts, is easier than a pragmatic one based upon the acceptance of many computational rules. So we will, from the beginning, introduce the mathematical tools, usually in their most general definition, into the representation of physical phenomena and show how their properties fit with what we can understand of these phenomena, and how they help to solve some classical problems. This will be illustrated by the building, step by step, of a formal model which incorporates all the bricks to show how they work. We will use many mathematical definitions or theorems. The most important will be recalled, and for the proofs and a more comprehensive understanding I refer to a previous book (”Mathematics for Theoretical Physics”) which is freely available.

But Mathematics can offer only tools : to use them efficiently it is necessary to understand what they mean, and why and how they can be adapted to represent physical phenomena. To do so one needs a new, candid, look at these phenomena, just as Einstein did in his celebrated 1905 article : space and time are not necessarily how we are used to see them, and more than often one needs to pause before jumping to Mathematics. Actually the indiscriminate use of formal systems can be hazardous, and we will see in this books some examples of common mathematical representations, seen as granted because they are used over and over, which are not grounded in any legitimate conceptual frame, or even do not reflect the actual practices of workers in natural sciences. Many physicists are indeed disturbed by the way Mathematics invade Theoretical Physics, with the continuous flow of more and more sophisticated mathematical theories, difficult to understand but for a few specialists, and which are estranged from experimental work. But most of them also comply with the mandatory, at least in the academic world, use of formal practices which seem to be against any logical justification, or are even non sensible, such as minimal substitution rules or divergent integral. Computation in Physics, as it is commonly done, has quite often the appearance of magic, justified only by the usual “experiments have proven that it works”, or more blatantly “the great masters have said that it is right”. Perhaps, but as a consequence it leaves the students with the feeling that Physics is more an exercise in rote knowledge than in rational understanding, and it opens the way to interpretations : if the formal system is strange, it is perhaps that nature itself is strange. A quick Google search for “quantum mechanics interpretations” provides more than 5 millions links, and there are more than 50 elaborate theories, the multiuniverse having the largest support in the scientific community. So one cannot say that modern Physics answer clearly our questions about nature. And it is not true that experiments have proven the rightfulness of the common practices. The discrepancy between what the theories predict and what is observed is patched with the introduction of new concepts, whose physical realization is more and more difficult to check : collapse of the wave function, Higgs boson, dark matter, brown energy,...

This is not the purpose of this book to add another interpretation to the existing long list. There will be few assumptions about the physical world, clearly stated[1], and they are well in line with what Physicists know and most Scientists agree upon. There will be no extra-dimensions, string theory, branes, supersymmetry,...Not that such theories should be discarded, or will be refuted, but only because they are not necessary to get a solid picture of the basic concepts in Physics. And indeed we do not answer to all questions in this book, some issues are still

[1] To be precise : assumptions are labeled “propositions”, and the results which can be proven from these propositions are labeled “theorems”.

open, but I hope that their meaning will be clearer, leading the way to a better and stronger understanding of the real world.

The first chapter is devoted to a bit of philosophy. From many discussions with scientists I felt that it is appropriate. Because the book is centered on the relation between Mathematics and Physics, it is necessary to have a good understanding of what is meant by physical laws, theories, validation by experiments, models, representations,...Philosophy has a large scope, so it deals also with knowledge : epistemology helps us to sort out the different meanings of what we call knowledge, the status of Science and Mathematics, how the Sciences improve and theories are replaced by new ones. This chapter will not introduce any new Philosophy, just provide a summary of what scientists should know from the works of professional philosophers.

The second chapter is dedicated to Quantum Mechanics (QM). This is mandatory, because QM has dominated theoretical Physics for almost a century, with many disturbing and confusing issues. It is at the beginning of the book because, as we will see, actually QM is not a physical theory per se, it does not require any assumption about how Nature works. QM is a theory which deals with the way one represents the world : its axioms, which appear as physical laws, are actually mathematical theorems, which are the consequences of the use by Physicists of mathematical models to make their computations and collect their data from experiments. This is not surprising that measure has such a prominent place in QM : it is all about the measures, that is the image of the world that physicists build, and not about the world itself. And this is the first, and newest, example of how the use of Mathematics can be misleading.

The third chapter is dedicated to the Geometry of the Universe. By this we do not mean how the whole universe is, which is the topic of Cosmology. Cosmology is a branch of Physics of its own, which raises issues of an epistemological nature, and is, from my point of view, speculative, even if it is grounded in Astrophysics. We will only evoke some points of Cosmology in passing in this book. By Geometry of the Universe I mean here the way we represent locations of points, components of vectors and tensors, and the consequences which follow for the rules in a change of representation. This will be done in the relativist framework, and more precisely in the framework of General Relativity. It is less known, seen usually as a difficult topic, but, as we will see, some of the basic concepts of Relativity are easier to understand when we quit the usual, and misleading, representations, and are not very complicated when one uses the right mathematical tools. We show that the concept of deformable solid can be transposed in GR and can be used practically in elaborate models. such as those necessary in Astrophysics.

The fourth chapter addresses Kinematics, which, by the concept of moment, is the gate between forces and geometry. Relativity requires a brand new vision of these concepts, which has been engaged, but neither fully or consistently. Rotation in particular has a different meaning in the 4 dimensional space than in the usual euclidean space, and a revision of rotational moment requires the introduction of a new framework. Spinors are not new in Physics, we will see what they mean, in Physics and in Mathematics, with Clifford algebras. This leads naturally to the introduction of the spin, which has a clear and simple interpretation, and to the representation of particles by fields of spinors, which incorporates in a single quantity the motion, translational and rotational, and the kinematics characteristics of material objects, including deformable solids.

The fifth chapter addresses Force Fields. After a short reminder of the Standard Model we will see how charges of particles and force fields can be represented, with the concept of connections on fiber bundles. We will not deal with all the intricacies of the Standard Model, but focus on the principles and main mechanisms. The integration of Gravity, not in a Great Unification Theory, but with tools similar to the other forces and in parallel with them, opens a fresh vision on important issues in General Relativity. In particular it appears that the common and exclusive use of the Levi-Civita connection and scalar curvature introduces useless complications but, more importantly, misses important features of the gravitational field. One of the basic properties of fields is that they propagate. This phenomenon is more subtle than it is commonly accepted. In a realist view of fields, that is the acceptance that a field is a physical entity which occupies a definite area in the universe, and experimentally checked assumptions, we deduce fundamental equations which can be used to explore the fields which are less well known, and notably gravitation.

The sixth chapter is dedicated to lagrangians. They are the work horses of Theoretical Physics, and we will review the problems, physical and mathematical, that they involve, and how to deal with them. We will see why a lagrangian cannot incorporate explicitly some variables, and build a simple lagrangian with 6 variables, which can be used in most of the problems.

The seventh chapter is dedicated to continuous models. Continuous processes are not the rule in the physical world, but are the simplest to represent and understand. We will see how the material introduced in the previous chapters can be used, how the methods of Variational Calculus, and its extension to functional derivatives, can be used in solving two models, for a field of particles and for a single particle. In this chapter we introduce the concept of currents and Energy-Momentum tensor and prove some important theorems. We give guidelines which can solve the equations for the gravitational field in the vacuum in the most general concept.

The eighth chapter is dedicated to discontinuous processes. They are common in the real world but their study is difficult. From the concept of propagation of fields, we shall accept that this is not always a continuous process. Discontinuities of fields then appear as particles, which can be assimilated to bosons. We show how their known properties can be deduced from this representation.

Chapter 1

WHAT IS SCIENCE ?

Science has acquired a unique status in our societies. It is seen by the laymen as the premier gate to the truth in this world, both feared and respected. Who could not be amazed by its technical prowess ? How many engineers, technicians, daily put their faith in its laws ? For many scientists their work has a distinctive quality, which puts them in another class than novelists, theologians, or artists. Even when dealing with some topics as government, traditions, religion,... they mark their territory by claiming the existence of Social Sciences, such as Economics, Sociology or Political Sciences, endowed with methods and procedures which stand them apart, and lest us say, above the others who engage in narratives on the same topics. But what are the bases for such pretense ? After all, many scientific assertions are controversial, when they impact our daily lives (from the climate warming to almost any drug), but not least in the scientific community itself. The latter is natural and even sound - controversy is consubstantial to science - however it has attained a more bitter tone in the last years, fueled by the fierce competition between its servants, but also by the frustrations of many scientists, mostly in Physics, at a scientifically correct corpus with too many loopholes. A common answer to the discontents is to refer them to the all powerful experimental proofs, but these are more and more difficult to reach and to interpret : how many people could sensibly discuss the discovery of the Higgs boson ?

To put some light on these issues, the natural way is to look towards Philosophy, and more precisely Epistemology, which is its branch that deals with knowledge. After all, for thousands of years philosophers have been the architects of knowledge. It started with the Greeks, mainly Aristotle who provided the foundations, was frozen with the scholastic interpretation, was revitalized by Descartes who brought in experimental knowledge, was challenged by the British empiricists Hume, Locke, Berkeley, achieved its full rigor with Kant, and the American pragmatists (Peirce, James, Putnam) added the concept of revision of knowledge. Poincaré made precise the role of formalism in scientific theory, and Popper introduced, with the concept of falsifiability, a key element in the relation between experiment and formal theories. But since the middle of the XX° century epistemology seems to have drifted away from science, and philosophers tend to think that actually, philosophy and science have little to share. This feeling is shared by many scientists (Stephen Weinberg in “Dreams of a Final Theory”). This is a pity as modern sciences need more than ever a demanding investigation of their foundations.

Without pretending to create a new epistemology and using all the basic work done by philosophers, I will try to draw a schematic view of epistemology, using a format and words which may be more familiar to the scientific reader. The purpose is here to set the ground, starting from questions such as What is knowledge ? How does it appear, is formatted, transformed, challenged ? What are the relations between experimentation and intuition ? We will see what are the specificities of scientific knowledge, how scientific theories are built and improved, what is the role of measures and facts, what is the meaning of the mathematical formalism in our theories. These are the topics of this first chapter.

1.1 WHAT IS KNOWLEDGE ?

First, a broad description of what is, and what is not knowledge.

Knowledge is different from perception : the most basic element of knowledge is the belief (a state of mind) of an individual with regard to a subject. It can be initiated, or not, by a sensitive perception or by the measure of a physical phenomenon.

Knowledge is not necessarily justified : it can be a certain perception, or a plausible perception ("I think that I have seen..."), or a pure stated belief ("God exists"), or a hypothesis.

Knowledge is shared beliefs : if individual states of minds can be an interesting topic, knowledge is concerned with beliefs which can be shared with other human beings. So knowledge is expressed in conventional formats, which are generally accepted by a community of people interested by a topic. This is not a matter of the tongue which is used, it supposes the existence of common conventions, which enables the transmission of knowledge without loss of meaning.

Knowledge is a construct : this is more than an accumulation of beliefs, knowledge can be learnt and taught and for this purpose it uses basic concepts and rules, organized more or less tightly in theories addressing similar topics.

1.1.1 Circumstantial assertions

The most basic element of knowledge can be defined as a **circumstantial individual assertion**, which can be formatted as comprised of :

- the author of the assertion

- the specific case (the circumstances) about which the assertion is made. Even if it is often implicit, it is assumed that the circumstances, people, background,.. are known, this is a crucial part of the assertion.

- the content of the assertion itself : it can be simply a logical assertion (it has the value true or false) or be expressed in a value using a code or a number.

The assertion can be **justified** or not. The author may himself think that his assertion is only plausible, it is a hypothesis. An assertion can be justified by being shared by several persons. A stronger form of justification is a **factual justification**, when everybody who wants to check can share by himself the assertion : the assertion is justified by evidence. In Sciences factual justifications are grounded in measures, done according to precise and agreed upon procedures : the experiment can be repeated.

Examples of circumstantial individual assertions :

"Alice says that yesterday Bob had a blue hat", "I think that this morning the temperature was in the low 15 $°C$","I believe that the cure of Alice is the result of a miracle",...

Knowledge, and specially scientific knowledge, is more than individual circumstantial assertions : it is a method to build narratives from assertions. It proceeds by enlargement, by going from individuals to a community, from circumstantial to universal, and by linking together assertions.

1.1.2 Rational narrative and logic

By combining together several assertions one can build a narrative, and any kind of theory is based upon such construct. To be convincing, or only useful, a narrative must meet several criteria, which makes it rational. *Rationality is different from justification : it addresses the syntax of the narrative*, the rules that the combination of different assertions must follow in the construct, and does not consider a priori the validity of the assertions. The generally accepted rules come from **logic**. Aristotle has exposed the basis of logic but, since then, it has become a field of research on its own (for more see Maths.Part 1).

Formal logic deals with logical assertions, that is assertions which can take the value true (T) of false (F) exclusively. Any assertion can be put in this format.

Propositional logic builds propositions by linking assertions with four operators $\wedge$ (and), $\vee$ (or), $\urcorner$ (not), $\Rightarrow$ (implies). For each value T or F of the assertions the propositions resulting from the application of the operators take a precise value, T or F. For instance the proposition : $P = (A \Rightarrow B)$ is F if $A = T$ and $B = F$, and $P = T$ otherwise. Then one can combine propositions in the same way, and explore all their possible values by "table-truth", which are just tables listing the propositions in columns, and all their possible values in rows.

Demonstration in formal logic uses propositions, built as above, and deduces true propositions from a collection of propositions deemed true (the axioms). To do this it lists axioms, then row after row, new true propositions using a rule of inference : if A is T, and $(A \Rightarrow B)$ is T, then B is T. The last, true, proposition is then proven.

These two kinds of propositional logic can be formalized in the Boolean calculus, and automated.

Propositions deal with circumstantial assertions. To enlarge the scope of formal logic, **predicates** are propositions which enables the use of variables, belonging to some fixed collection. Assertions and propositions are then linked with the use of two additional operators : $\forall$ (whatever the value of the variable in the collection), $\exists$ (there is a value of the variable in the collection). In first order predicates, these operators act only on variables, which are previously listed, and not on predicates themselves. One can build table-truth in the same way as above, for all combinations of the variable. Demonstrations can be done in a similar way, with rules of inference which are a bit more complicated.

The Gödel's completeness theorem says that any true predicate can be proven, and conversely that only true predicates can be proven. The Gödel's compactness theorem says in addition that if a formula can be proven from a set of predicates, it can also be proven by a finite set of predicates : there is always a demonstration using a finite number of steps and predicates. These two theorems show that, so formalized, *formal logic is fully consistent, and can be accepted as a sound and solid basis to build rational narratives.*

This is only a sketch of logic,which has been developed in a sophisticated system, important in computer theory. Several alternate formal logics have been proposed, but they lead to more complicated, and less efficient, systems, and so are not commonly used. Other systems called also "logic", have been proposed in special fields, such as Quantum Mechanics (see Josef Jauch and Charles Francis for more) and information theory. Actually they are Formal Systems, similar to the Theories of Sets or Arithmetic in Mathematics : they do not introduce any new Calculus of Predicates, but use Mathematical Logic acting on a set of axioms and propositions.

Using the basic rules of formal logic, one can build a **rational narrative**, in any field. Notice that in the predicates the collections to which variables must belong are not sets, such as defined in Mathematics, and no special property is assumed about them. A variable can be a citizen, belonging to a country and indeed many laws could be formulated using formal logic.

Formal logic is not concerned about the justification or the veracity of the assertions. It tells only what can be logically deduced from a set of assertions, and of course can be used to refute propositions which cannot be right, given their premises. For instance the narrative :

$\forall X$ human being, $((X \text{ is ill}) \wedge (X \text{ prays}) \wedge (\text{God wills})) \Rightarrow (X \text{ is cured})$

is rational. It is F only if there is a X such that the first part is T and X is not cured. And one can deduce that God's will is F in this case. Without the proposition (God wills) it would be irrational.

Rational narrative are the ingredient of mystery books : at the end the detective comes with a set of assertions to unveil the criminal. A rational narrative can provide a plausible explanation, and a rational, justified, narrative, is the basis for a judgement in a court of law.

Scientific knowledge of course requires rational narratives, but it is more than that. A plausible explanation is rooted in the specific circumstances in which it has occurred : there is no reason why, under the same circumstances, the same facts would happen. To go further one needs a feature which is called necessitation by philosophers, and this requires to go from the circumstantial to the universal. And scientific knowledge is justified, which means that the evidences which support the explanation can be provided in a controlled way.

1.2 SCIENTIFIC KNOWLEDGE

1.2.1 Scientific laws

Let us take some examples of scientific laws :

A material body which is not submitted to any force keeps its motion.

For any ideal gas contained in a vessel there is a relation $PV = nRT$ between its pressure, volume, and temperature.

For any conductive material submitted to an electric field there is a relation $U = RI$ between the potential U and the intensity I of the current.

Any dominant allele is transmitted to the descenders.

Scientific laws are assertions, which have two key characteristics :

i) They are **universal** : they are valid whenever the circumstances are met. A plausible explanation if true in specific circumstances, a scientific law is true whenever some circumstances are met. Thus in formal logic they should be preceded by the operator $\forall$. This is a strong feature, because if it is false in only one circumstance then it is false : it is **falsifiable**. This falsifiability, which has been introduced by Popper, is a key criterion of scientificity.

ii) They are **justifiable** : what they express is linked to physical phenomena which can be reproduced, and the truth of the law can then be checked by anybody. In a justified plausible explanation, the evidences are specific and exist only in one realization. For a scientific law the evidences can be supplied at will, by following procedures. A scientific law is justified by the existence of reproducible experimental proofs. This feature, introduced by Kant, distinguishes scientific narratives from metaphysical narratives.

One subtle point of falsifiability, by checking a prediction, is that it requires the possibility, at least theoretically, to test and check any value of each initial assertion before the prediction. Take the explanation that we have seen above :

$\forall X$ human being, ((X is ill) $\wedge$ (X prays) $\wedge$ (God wills)) $\Rightarrow$ (X is cured)

For any occurrence, three of the assertions can be checked, and so one could assume that the value of the fourth (God's will) is defined by the final outcome in each occurrence, and we would have a scientific law. However falsifiability requires that one could test for different values of the God's will before measuring the outcome, so we do not have a scientific law. The requirement is obvious in this example but we have less obvious cases in Physics. Take the two slits experiment and the narrative :

(particles are targeted to a screen with two slits) $\wedge$ (particles behave as waves) $\Rightarrow$(we see a pattern of interferences)

Without the capability to predict which of the two, contradictory, behaviors, is chosen, we cannot have a scientific law.

These criteria are valid for any science. The capability to describe the circumstances, to reproduce or at least to observe similar occurrences, to check and whenever possible to measure the facts, are essential in any science. However falsifiability is usually a difficult criterion to meet in Social Sciences, even if one strives to control the environment, but this is close to impossible in Archeology or History, where the circumstances in which events happened are difficult or impossible to reproduce, and are usually not well known. The extinction of the dinosaurs by the consequences of the fall of an asteroid is a plausible explanation, it seems difficult to make a law of it.

1.2.2 Probability in scientific laws

The universality of scientific laws opens the way to probabilistic formalization : because one can reproduce, in similar or identical manner, the circumstances, one can compute the probability of a given outcome. But this is worth some clarification because it is closely linked to a big issue : are all physical processes determinist ?

In Social Sciences, which involve the behavior of individuals, the assumption of free will negates the possibility of determinist laws : the behavior of a man or woman cannot be determined by his or her biological, social or economic characteristics. This has been a lasting issue for philosophers such as Spinoza, with a following in Marxist ideas. Of course one could challenge the existence of free will, but it would not be without risk : the existence of free will is the basis for the existence of Human Rights and the Rule of Law. Anyway, from our point of view here, no scientific law has been proven which would negate this free will, just more or less strong correlations between variables, which can be used in empirical studies (such as market studies).

In the other fields, the discrepancy between the outcomes can be imputed to the fact that the circumstances are similar, but not identical :

- the measures are imperfect

- the properties of the objects (such as their shape) are not exactly what is assumed

- some phenomena are neglected, because it is assumed that their effect is small, but it is non null and unknown

This is a common case in Engineering, where phenomenological laws are usually sufficient for their practical use (for instance for assessing the strength of materials). In Biology the Mendel's heredity laws provide another example. As an extreme example, consider the distribution of the height of people in a given population. It seems difficult to accept that, for a given individual, this is a totally random variable. One could assume that biological processes determine (or quite so) the height from parameters such as the genetic structure, diet, way of life,... The distribution that one observes is the result of the distribution of the factors which are neglected, and it can be made more precise, for instance just by the distinction between male and female.

And similarly, at a macroscopic scale, probabilist laws are commonly used to represent physical processes which involve a great number of interacting microsystems (such as in Thermodynamics) whose behavior cannot be individually measured, or discontinuous processes such as the breakdown of a material, an earth-quake,...which are assumed to be the result of slow continuous processes.

In all these cases a probabilist law does not imply that the process which is represented is not determinist, just that all the factors involved have not been accounted for. I don't think that any geologist believes that earth-quakes are pure random phenomena.

However one knows of physical elementary processes which, in our state of knowledge, seem to be not determinist : the tunnel effect in semi-conductors, the disintegration of a nucleus or a particle, or conversely the spontaneous creation of a particle,..

Quantum Mechanics (QM) makes an extensive use of probability laws, and some of its interpretations postulate that at some scale physical laws are fundamentally not determinist. Up to now QM is still the only theory which can represent efficiently elementary non determinist phenomena. However, as we will see in the next chapter, the probabilist feature of the main axioms of QM does not come from some random behavior of natural objects, but from the discrepancy between the measures which can be done and their representation in our theories.

1.2.3 Models

To implement a scientific law, either to check it or to use it for practical purpose (to predict an outcome), scientists and engineers use models. A **model** can be seen as a general representation of the law. It comprises :

- a system : the area in which the system is located and the time frame during which it is observed, the list of the objects and of their properties which are considered

- the circumstances if they are specific (temperature, interference with the exterior of the system,...)

- the variables representing the properties, associated each to a mathematical object with more specific mathematical properties if necessary (a scalar can be positive, a function can be continuous,...)

- the procedures used to collect and analyze the data, notably if statistical methods are used.

Building and using models are a crucial part of the scientific work. Economists are familiar with the denomination models, either theoretical or as a forecasting tool. If they are not known by the name, any engineer or theoretical physicist use them, either to compute solutions of a problem from well established laws, or to explore the consequences of more general hypotheses. A model is a representation, usually simplified, of part of the reality, built from concepts, assumptions and accepted laws. The simplification helps to focus on the purpose, trading accuracy for efficiency. Models provide both a framework in which to make the computations, using some formalism in an ideal representation, and a practical procedure to organize the collection and analysis of the data. They are the embodiment of scientific laws, implemented in more specific circumstances, but still with a large degree of generality which enables to transpose the results from one realization to another. Actually most, if not all, scientific laws can be expressed in the framework of a model.

Models use a formalism, that is a way to represent the properties in terms of variables, which can take different values according to the specific realizations of the model, and which are used to make computations to predict a result. The main purpose of the formalism is efficiency, because it enables to use rules and theorems well established in a more specific field. If the variables are logic, then formal logic provides an adequate formalism. Usually in Physics the formalism is mathematical, but other formalisms exist. The most illuminating example is the atomic representation used in Chemistry. A set of symbols such as :

H2 + 1/2 O2 $\rightarrow$ H2O + 286 kJ/Mol

tells us almost everything which is useful to understand and work with most of chemical experiments. Similarly Economics uses the formalism of Accounting.

However the role of Mathematics in the formalism used in Physics leads us to have a look about the status of Mathematics itself in Science.

1.2.4 The status of Mathematics

It is usually acknowledged that Euclide founded Mathematics, with his Geometry, based on the definition of simple objects (points, lines,...) which are idealization of physical objects, a small number of axioms, and logic as the computational motor. For millennia it has been seen as the embodiment of rationality, and Mathematics has been developed in a patchwork of different fields : Algebra, Analysis, Differential Geometry... extending the scope of objects, endowed with more sophisticated properties. In the XIX° century mathematicians felt the need to unify this patchwork and to found a clean Mathematics, grounded in as few axioms as possible. This was also the consequence of discoveries, such as non euclidean geometries by Lobatchevski, and of paradoxes in the newly borne Cantor's set theory. And this was also the beginning of many controversies, which are not totally closed at this day.

However this endeavour (promoted by Hilbert) lead to the creation of Mathematical Logic. This is actually a vibrant field of Mathematics of its own, which aims at scrutinizing Mathematics with respect to its consistency. It became clear that, in order to progress, it was necessary to distinguish in the patchwork some mathematical theories, and the focus has been put on Arithmetic and Set Theory, as they are the starting point for all the other fields of Mathematics. Without attempting to give even an overview of Mathematical Logic, three main features emerge from its results :

- the need to define objects specific to each field (natural numbers, sets) through their properties which are then enshrined in the axioms of the theories;

- the fact that these objects are of an abstract nature, in the meaning that they cannot be seen simply as the idealized realization of some physical objects, as points, lines,... were in Euclidean Geometry;

- and this fact is compounded by the need to assume properties which cannot be the realization of physical objects : the key example is the axiom of infinity in the Set Theory which postulates the existence of a set with an infinite number of elements.

So Mathematics is essentially different from formal logic (even if it uses it to work on these objects) : it relies on the prior definitions of objects and precise axioms, and deal only with these objects and those which can be constructed from them. Formal logic is only syntax, Mathematics assumes a semantic part.

On these bases several sets of axioms have been proposed both for Arithmetic (Peano) and the Set Theory (Zermello, Frankael). They provide efficient systems, which have been generally accepted at the time and still nowadays, with some variants. However two results came as a big surprise:

- Gödel proved in 1931, with complements given by Gentzen in 1936 and Ackerman in 1940, that in any formal system powerful enough to represent Arithmetic, there are propositions which are true but cannot be proven.

- Church proved in 1936 that there cannot exist a fixed procedure to prove any problem in Arithmetic in a finite time (this is not a decidable theory).

The incompleteness Gödel's theorem is commonly misunderstood. Its meaning is that, to represent Arithmetic with all its usual properties that we know, we need a minimum set of axioms, but one could then add an infinite number of other axioms, which would not be inconsistent with the theory : they are true, because they are axioms, and they cannot be proven, because they are independent from the other axioms.

The Church's theorem is directly linked with computers (formalized as Turing's machines) : it cannot exist a program which would solve automatically all problems in Arithmetic.

Many similar or more sophisticated results have been proven in different fields of Mathematical Logic. For our purpose here, several conclusions can be drawn :

i) Mathematics can be seen as a science : it deals with objects and properties, using formal logic, to deduce laws which are scientific by the fact that they are always true for any realization of the objects. It has the great privilege to invent its own objects, however this comes with a price : the definition is not unique, other properties could be added or specified without harming Mathematics.

ii) The choice of the right axioms is not dictated by necessity, but by efficiency. Mathematics, as we know it, has not been created from scratch by an axiomatic construct, it is the product of centuries of work, sometimes not rigorous, and the axioms which emerge today are the ones which have been proven efficient for our needs. But perhaps, one day, we will find necessary to enlarge the set of axioms, as it has been done with the axiom of infinity.

iii) Because the objects are not simple idealization of physical realizations, and because there is no automated procedure to prove theorems, and so to extend Mathematics, it appears that it is a true product of the human mind. All mathematicians (as Poincaré noticed) have known these short periods of illumination, when intuition prevails over deduction, to find the right path to the truth. It seems that an artificial intelligence could not have arrived to the creativity that Mathematics requires.

1.3 THEORIES

Scientific laws are an improvement over circumstantial explanations, because they have the character of necessity and they are related to physical observable phenomena. Often philosophers view laws of nature as something which has to be discovered, as a new planet, hidden from our knowledge or perception. But science is more than a collection of laws, it has higher goals, it aims at providing a plausible explanation for as many cases as possible. Early on appeared the want to unify these laws, either to induce a cross fertilization process, or by the more holistic concern to understand what is the real world that they describe : *Science should provide more than efficient tools, it should explain what it is.*

Scientific laws rely on the definition of objects (material body, force, ideal gas,...) which have properties (motion, volume, pressure,...) related to observable physical phenomena and also represented by mathematical objects (scalar, functions,...). These concepts have emerged in each field, and have been organized in Theories : Mechanics, Fluid Mechanics, Thermodynamics, Electromagnetism, Theory of Fields,... and a similar process has been at work in Chemistry or Biology. And of course the want to unify further these fields has appeared. However the endeavour has not gone as well as in Mathematics. Many scientists are quite pleased with their tools and do not feel the need to go beyond what they use and know. A pervasive mood exists in Physics that the focus shall be put on experiments : if it works then it is true, whatever the way the computations are done. In an empiricist vision the concepts are nothing more than what is measured : a scientific law is essentially the repeated occurrences of observed facts, and one can accept a patchwork of laws. QM has greatly strengthened this approach, at first by casting a deep doubt about concepts which were thought to be strong (such as location, speed, matter,...) and the generalization of probabilist laws, and then by promoting the use of new concepts (fields, wave function, superposition of states...) which, from the beginning, were deemed to have no physical meaning, at least that we could understand. However the need for a more unified and consistent vision exist, even if it is met by unsatisfying construct, and one goes from a patchwork of scientific laws to theories.

1.3.1 What is a scientific theory ?

A scientific theory aims at giving a unified vision of a field, a framework in which scientific laws can be expressed, and a formalism which enables to deduce new laws that can be checked. So it comprises :

- a set of concepts, objects related to physical realizations, to which are attached properties which can be measured. These properties can be seen as defining the objects.

- a set of fundamental laws, or first principles : expressed in general terms, they are based on the observation of the physical world, and grounded in experiments, but they can or cannot be checked directly.

- a formalism, which provides the framework of models, and the computational tools to deduce new laws, forecast the results of experiments and check the laws.

Examples :

The atomist theory in Chemistry. Compounds are made of a combination of 118 elements with distinct chemical properties, chemical reactions occur without loss of elements and an exchange of energy, ruled by thermodynamics.

The Newton's Mechanics. Material bodies are composed of material points, in a solid they stay at a constant distance from each other. The motion of material bodies is represented in the Galilean Geometry, it depends on their inertia and on forces which are exerted by contact or at a distance, according to fundamental laws.

Special Relativity. The universe is a four dimensional affine space endowed with a fixed Lorentz metric. Material bodies move along world lines at a constant velocity and their kinematics is characterized by their energy-momentum vector. The speed of light is constant for any inertial observer.

The properties are crucial because, for each situation, they can identify generic objects with similar properties, and associate to these objects a set of well defined values, which can be measured in each occurrence : "all insects have three pairs of legs", "material bodies travel along a world line in the 4 dimensional universe", "for any gas there is a temperature T". But by themselves they do not have a predictive power. In some cases the value of the variable comes from the definition itself (the number of legs of an insect), but usually it does not provide the value of the variable (the temperature of a gas).

As said before, the formalism used is not necessarily mathematical, but it acquires a special importance. This is a matter of much controversies but it is clear that major steps in the theories would have been impossible without prior progresses in the formalism which is used : Chemistry with the atomist representation, Mechanics with differential and integral calculus, General Relativity with differential geometry, and even Economics with Statistics. The use of more powerful mathematical tools, and similarly of computational techniques, increases our capacity to check predictions, but also to build the theories. Inspired by Thermodynamics and QM, it has been proposed to give to Information Theory an unifying role in Physics. A step further, considering that many structures used in different fields have similar features, the Category Theory, a branch of Mathematics developed around 1945 (Eilenberg, Mac Lane) has been used as a formalism in Physics, notably in Quantum Computing (Heyting algebras).

Fundamental laws can be not justified experimentally, their validity stems essentially from the consequences which can be deduced from them. From this point of view this is the theory as a whole which is falsifiable : if any law that can be deduced in the framework of the theory is proven false, then this is the entire theory which is at risk. And actually this has been a recurring event : Maxwell's laws and Galilean Geometry had to be revised after the Michelson and Morley experiments, the Atomist theory has had to integrate radio-activity,...The process has not gone smoothly, and usually patches are proposed to sustain the existing theory. And indeed a good part of the job of scientists is to improve the theories, meaning to propose new theories which are then checked. What are the criteria in this endeavour ?

1.3.2 The criteria to improve scientific theories

Simplicity

The first criterion is simplicity. This is an extension of the Occam's razor rule : whenever we face several possible explanations, the fewer assumptions are made, the better. With our description of scientific theories it is easy to see what are the parameters to look for improvements. There must be as few kinds of objects as possible, themselves differentiated by a small number of properties or variables. There are 118 elements with distinct chemical properties, their nuclei are comprised of 12 fermions, there are millions of eukaryotes, but their main distinctive characteristics come from their DNA, organized in a small number of chromosomes, which are a combination of 4 bases. The electric and magnetic fields have been unified by the Maxwell's laws, and the unifications of all force fields including gravitation is the Graal of physicists. Similarly there should be as few fundamental hypotheses as possible. The Galilean system was not more accurate or legitimate (the assertions that Earth circles the Sun or that Sun moves around the Earth are both valid) than the Ptolemaic system, but it was simpler and provided a general theory to compute the trajectories of bodies around a star and paved the way to the Newton's gravitation law.

Enlarge the scope of phenomena addressed by the theory

The second criterion is the scope of the field which is addressed by the theory. Science is imperialist : it strives to find a rational explanation to everything. Lead by the Occam's razor rule it looks for more fundamental objects and theories, from which all the others could be deduced. This is a fact, and a legitimate endeavour. It has been developed in the different forms of positivism. In its earlier version (A.Comte) science had to deal only with and proceed from empirical evidence, scientific knowledge could be built by a logic formalization, which leads to a hierarchy of sciences giving preeminence to mathematics. In its more modern version positivism embraces the idea of the unity of science, that there is, underlying the various scientific disciplines, basically one science about one real world. Actually this is more complicated.

Starting with mathematics, as we have seen it could be seen as a science. True, mathematicians can invent their own objects. Quite often a narrative in Mathematics starts as "Let be a set such that...", but the first step required is to prove that such a set exists (as an example the definition of the tensorial product of vector space from an universal property). And if this is not possible one has to add another axiom (such as infinite sets), and support the consequences.

In natural sciences it is a sound requirement that there is a strong, unified background, explaining and reflecting the unity of the physical world. But in the different fields theories usually do not proceed from the most elementary laws. The atomic representation used in Chemistry precedes quantum field theories of particles. Biology acknowledges the role of chemical reactions, but its basic concepts are not embedded in chemistry. We do not have in Physics a theory which would be general and powerful enough to account for everything. And anyway in most practical cases specific theories suffice. They use a larger set of assumptions, which are simplified cases of general laws (Galilean Geometry replacing Relativist Geometry, Newton's laws substituted to General Relativity) or phenomenological laws based on experimental data. In doing this the main motivation of scientists is efficiency : they do not claim the independence of their fields, but acknowledge the necessity of simpler theories for their work. However one cannot ignore that this move from one level to the other may cover a part of mystery. We still do not understand what is life. We do not have any determinist model of irreversible elementary process (such as the disintegration of a particle).

Economics is by far the social science which has achieved the higher level of formalization, in theoretical studies, empirical predictive tools, and in the definition of a set of concepts which give a rigorous basis for the collection and organization of data. Through the accounting apparatus, at the company level, the state level as well as many specialized fields (welfare, health care, R&D,...) one can have a reliable and quantified explanation of facts, and be able to assess the potential consequences of decisions. Because of the stakes involved these concepts are controversial, but this is not an exclusivity of Economics [1]. Actually what hampers Economics, and more generally the Social Sciences, is the difficulty of experimentation. Most of the work of scientists in these fields relies on data about specific occurrences, past or related to a few number of cases. The huge number of factors involved, most of which cannot be controlled, weakens any prediction[2], and the frailty of phenomenological laws in return limits the power of the falsifiability check. But this does not prevent us to try.

So we are still far away from a theory of everything. But the imperialism of science is legitimate, and we should go with the Hilbert's famous saying : "Wir müssen wissen, wir werden

[1] Actually some philosophers (who qualify themselves as feminists, such as Antony) deny that science is objective, and is very much an instrument of oppression (in Turri about Quine).

[2] And anyway it would be difficult to justify the realization of an economic crisis in order to check a law. Quite often Economics predictions are no realized because the implementation of the Economics Theory have prevented them to happen.

wissen". It is backed by the pressing want of people to have explanations, even when they are not always willing to accept them. As a consequence it increases the pressure on scientists and more generally on those who claim to have knowledge. As G.B.Shaw said "All professions are a conspiracy against the laity". So it is a sound democratic principle that scientists should be kept accountable to the people who fund their work.

Conservative pragmatism

The third criterion in the choice of theories is that any new theory should account for the ones that it claims to replace. What one can call a conservative pragmatism. Sciences can progress by jumps, but most often they are revisions of present theories, which become embedded in new ones and are seen as special case occurring in more common circumstances. This process, well studied by G.Bachelard, is most obvious in Relativity : Special Relativity encompasses Galilean Geometry, valid when the speeds are weak, and General Relativity encompasses Special Relativity, valid when gravitation does not vary too much. Old theories have been established on an extended basis of experimental data, and backed by strong evidences which cannot be dismissed easily. New evidences appear in singular and exceptional occurrences and this leads to a quest for more difficult, and expansive, experimentations, which require more complex explanations. This is unavoidable but has drawbacks and the path is not without risks. The complexity of the proofs is often contrary to the first criterion - simplicity - all the more so when the new theory involves new objects with assumed, non checked, properties. The obvious examples are dark matter, or the Higgs boson. Of course it has happened in the past, with the nucleus, the neutrino, ... but it is difficult to feel comfortable in piling up enigma : the purpose of science is to provide answers, not to explain a mystery by a riddle. And when the new enigma requires more powerful tools the race may turn into a justification in itself.

1.4 FOUR QUESTIONS ABOUT SCIENCE

1.4.1 Is there a scientific method ?

It is commonly believed that one distinctive feature of the scientific work is that it proceeds according to a specific method. There is no doubt that the prerequisite of any scientific result is that it is justified for the scientific community. So the specificity of a scientific method would be guaranteed by higher ethical and professional standards. This claim is commonly associated to the "peer review" process : any result is deemed scientific if it has been approved for publication by at least two boffins of the field. Knowing the economics of this process, this criterion seems less reliable than what is usually required for an evidence in a court of justice, as recent troubles with published results show. The comparison is not fortuitous. For people who have dedicated years of their life to develop or to teach ideas, it is neither easy nor natural to challenge their beliefs, and all the more so when these beliefs are supported by the highest authorities in the field. Science has become a very competitive area, with great fame and financial stakes. Assume that fierce competition has increased the pressure to innovate is a bit optimistic. The real pressure comes from outside the scientific community, when quick economic return can be expected from a new discovery. This is no surprise that Computer Sciences or Biology have made gigantic progresses, meanwhile Particle Physics is still praising a Standard Model 40 years old. In any business, if the introduction of a new product was submitted to the anonymous judgment of your competitors, there would be no innovation. Only the interest of the customers should matter, but in Science this is a very distant concern, as well as the more direct interest of students who strive to understand theories that are reputed impossible to understand.

More generally this leads to question the existence of a science in fields such as History, Archeology,... Clearly there are criteria for the justification of assertions in these fields, which are more or less agreed upon by their communities, but it seems difficult that these assertions would ever be granted the status of scientific laws, at best they are plausible explanations.

So, and in agreement with most philosophers, I consider that scientific knowledge cannot be characterized by its method.

1.4.2 Is there a Scientific Truth ?

A justified assertion can be accepted as truth in a Court of justice. But not many people would endorse a scientific truth, and probably few scientists as well. Scientific theories are backed by a huge amount of checked evidences, and justified by their power to provide plausible explanations for a large scope of occurrences. So in many ways they are closer to the truth than most conceivable human assertions, but the purpose of science is not the quest for the truth, because science is a work in progress and doubt is a necessary condition for this progress. A striking example of this complex relation between science and truth is Marxism : Karl Marx made very valuable observations about the relations between technology, economic and political organizations, and claimed to have founded a new science, which enables people to make history. The fact that his followers accepted his claims to be the truth had dramatic consequences [3].

1.4.3 Science and Reality

Science requires the existence of a real world, which does not depend on our minds, without which it would be impossible to conceive universal assertions. Moreover it assumes that this reality is

[3] This aspect of marxism as the pretense of a science has been explained in my article published in 1982 in "les temps modernes".

unified, in a way that enables us to know its different faces, if any. Perhaps this is most obvious in social sciences : communities have very different organizations, beliefs and customs, but we strive to study them through common concepts because we see them as special occurrences of Human civilizations, with common needs and constraints. However this does not mean that we know what is reality : what we can achieve is the most accurate and plausible representation of reality, but it will stay temporary, subject to revision, and adjusted to the capability of our minds.

Because this representation is made through a formalization, the language which is used acquires a special importance. Some scientists resent this fact, perceived as an undue race towards abstraction, meanwhile they believe that empirical research should stay at the core of scientific progress. Actually the issue stems less from the use of more sophisticated mathematics than from the reluctance to adjust the concepts upon which the theories are based to take full advantage of the new tools. It is disconcerting to see physical concepts such as fields, particles, mass, energy, momentum,.. mixed freely with highly technical topological or algebraic tools. The discrepancy between the precision of the mathematical concepts and the crudeness of the physical concepts is source of confusion, and defiance. But the revision of the concepts will not come from the accumulation of empirical data, whatever the sophistication of the computational methods, it will come from fresh ideas.

From where do come these fresh ideas ? They are not the result of inference : a theory, with its collection of concepts and related formalism, has for purpose to provide models to explain specific occurrences. A continuous enlargement of the scope of experimental research provides more reliable laws, or conversely the proof of the failure of the theory, but it does not creates a new theory. New theories require a revision of the concepts, which may imply, but not necessarily, new hypotheses which are then checked. Innovation is not a linear, predictable process, it keeps some mystery, which, probably, is related to the genuine difference between computers and human intelligence. But it is obvious that a deep understanding of the concepts is a key to scientific progress.

1.4.4 Dogmatism and Hubris in Science

As the criteria for the validation of Scientific knowledge began to emerge, the implementation of the same criteria leads to two opposite dogmatisms, and their unavoidable hubris. And what is strange is that, in some areas of the present days Physics, these opposite succeeded to be packaged together, for the worst.

The first dogmatism is the identification of the real world with the concepts. This is what Euclide and generations of mathematicians did for millennia : a point, a line, exist really, as well as parallels lines : after all they are nothing more than the idealization of tangible objects whose properties can be studied as suited. The overwhelming place taken by the mathematical formalism and the power it gives to compute complicated predictions lead to believe in the adequation between models and the real world. If it can be computed, then it exists. And if something cannot be computed, it is not worth to be considered. The first challenges to this dogmatism appear with Relativity, then the Physics in the atomic world. Scientists had been used to consider natural a 3 dimensional euclidean universe, with an external time. The jump to a 4 dimensional representation, and worst a curved Universe, seemed intractable. If the Universe integrates time, do the past and future events exist all together ? Still today, even for some professionals physicists, it seems difficult to address these questions. They do not realize that, after all, the idea of an infinite, flat Universe, existing for ever, is also a controversial representation. Similarly Mechanics and its admirable mathematical apparatus, seemed to breakdown when confronted to experiments in the atomic world : particles cannot pass the test of the two slits experiments, electrons could not keep a stable orbit around the nucleus, even Chemistry was challenged with the non conservation of matter and elements. Of course Engineers had for centuries a more pragmatic approach to the problem, the clean idea of continuous, non dissipative, motion had been replaced by phenomenological laws which could deal with deformable solids, fluid, and gas. But this was only Engineering...

The second dogmatism appeared, and triumphed, in reaction to the disarray caused by this discrepancy between a comprehensive and consistent vision and the experiments. Since the facts are the ultimate jury in checking a Scientific Theory, let us put the measures at the starting point in the elaboration of the theories. And because experimentation is overall a matter of statistical evaluation, it is natural to give to probability the place that it should have had from the beginning. There is nothing wrong in acknowledging the actual practices of scientific experiments. After all a Scientific Law is no more than the repetition of occurrences. The formalism of Statistical Mechanics was available, and soon, with the support of some mathematical justification, Quantum Physics had been born, and stated in axioms, rules and computational methods.

The central issue, pushed by the supporters of the first dogmatism, was then to find a physical justification to the new formalism. As of today there has not been a unique answer. For some physicists Quantum Mechanics belong to a realm inaccessible to human understanding, a modern Metaphysics that it is vain to discuss, even if it can be marginally justified by mathematical considerations in simple cases. For others the want to find an interpretation is stronger, and the past century has been heralded with hundred of interpretations. They succeed actually in merging the two dogmatisms : if QM is stated in bizarre, non intuitive rules, it is because Reality itself is bizarre : it is discreet, non determinist. We retrieve the identification of the formalism, as convoluted as it is, with the real world, but at the price of an obvious lack of agreement in the Scientific community, and at best a muddled picture. One of the strangest example of this new dogmatism is given in Cosmology : because we can model the Universe, it is possible to compute the whole Universe, and adding some QM, even consider the wave function of the Universe, which could then assess the probability of occurrences of the parallel universes...

Dogmatism and hubris go together. The criterion of factual justification is replaced by the forced identification of the real world with the formalism : if the computation works, it is because this is how the physical reality is. Humility is not the most significant feature of the Human mind, happily so. We need concepts, broad, easy to understand, illuminating and consistent representations which can be implemented and developed, which can be understood, learnt and taught. They can only be the product of intuition, of the imagination of the Human brain, they will never come from a batch of data. These ideas must be kept in check by the facts, not suppressed by the facts. But in the same time we must keep in mind that these are our concepts, our ideas, and that reality is still there, waiting to be probed, not enlisted to our cause. This leads to the reintroduction of the Observer in Physics, an object to which the rest of the book will give a significant place, and to which it could be dedicated.

1.5 FUNDAMENTAL PRINCIPLES IN PHYSICS

Whatever the theory in Physics there are some fundamental principles which are generally accepted.

1.5.1 Principle of Relativity

Scientific laws in Physics require measures of physical phenomena. Each object identified in a model has properties which are associated to mathematical objects, and the measure of these properties implies that it is possible to associate figures, real scalars, to the properties. There are many ways to do this, and because Scientific laws are universal, it shall be possible to do the measures in a consistent way, in precise protocols, and because it shall be possible to check the law in different occurrences, the protocol must tell how to adapt the measures to different circumstances.

The Principle of Relativity is used with different meanings in the literature. Here I will state it as "Scientific laws do not depend on the observer". Which is the logical consequence of the definition of Scientific laws : they should be checked for any occurrence, as long as the proper protocols are followed, whoever do the experiment (the observers), whenever and wherever they are located. It has strong and important consequences in the mathematical formalization of the theories.

In any model the quantities which are measured are represented as mathematical objects, which have their own properties, and these properties are a defining part of the model, notably because they impose the format to collect the data. For instance in the Newton's law $\overrightarrow{F} = m\overrightarrow{\gamma}$ the quantities $\overrightarrow{F}, \overrightarrow{\gamma}$ are vectors, and we must know how their components change when one uses one frame or another. Similarly the laws should not depend on the units in which the quantities are expressed. As a general rule, if a law is expressed as a relation $Y = L(X)$ between variables X, Y and there are relations $X' = R(X)$, $Y' = S(Y)$ where R, S are fixed maps, given by the protocols under which two observers proceed and thus known, then the law L' shall be such that : $Y' = L'(X') \Leftrightarrow L' = S \circ L \circ R^{-1}$. This is of special interest when R, S vary according to some parameters, because the last relation must be met whatever the value of the parameter. This is the starting point for the gauge theories in Physics.

The Principle of Relativity assumes that there are observers. In its common meaning an observer is the scientist which makes the measures. But in a Theory it requires that one defines the properties of an observer : this is a concept as the others, and it is not always obvious to define precisely and in a consistent way what are these properties. One key property of observers is that they have free will, and this implies notably that they can change freely the conditions of an experiment (as the universality of scientific laws requires) : they can choose different units, spatial location of their devices, repeat the same experiment over and over,... Free will implies also that the observers are not subjected to the laws which rule the system they observe, however they are also subjected to physical laws but it is assumed that these laws do not interfere with the experiment they review. This raises some issues in Relativity, and a big issue in Cosmology, which is a theory of the whole Universe.

1.5.2 Principle of Conservation of Energy and Momentum

The principle is usually stated as "In any physical process the total quantity of energy and momentum of a system is conserved". But its interpretation raises many questions.

The first is about the definition of energy and momentum. They come from the intuitive notion that every physical object carries with it a capacity either to resist to a change, or to cause a change in other objects. So energy and momentum are attached to each object of the system : it is one of their properties. For localized objects such as material bodies, these quantities are localized as well. For objects which are spread over a vast area (fluids, force fields), energy and momentum are defined as density, related to some measure of volume of the area. Then the principle reads as the sum of energy and momentum for all the objects of the system is conserved.

The second is the distinction between energy and momentum. The former is expressed as a scalar, the latter as a vectorial like quantity. And both are intuitively linked.

For a material body the momentum is related to the motion. Motion is a purely geometric concept, corresponding to the location and disposition of a material body, and their change with time. The motion of a material body cannot be changed without an external action, and this resistance, characterizing its inertia, is related to the mass of the body and its motion, and represented by kinematic quantities : the translational momentum and the rotational momentum, which have the general form [mass] $\times$ [motion]. So momenta for material bodies are expressed by kinematic quantities, which express their inertia, and they require the introduction, besides the representation of the motion, of other variables similar to mass. Motion is a derivative of location, and it can be expressed at any given time : momenta are continuous variables, as long as the motion is continuous.

The instantaneous effort which is required to change the momenta is a force : $F = \frac{d}{dt}$ [Momentum] , which equals the force of inertia of the body. The total effort done in a physical process during which the momenta have been changed is the energy transferred to the body, expressed as Δ([Momentum]$\times$[Motion]) essentially because its most visible form is the kinetic energy of material bodies, which is also the source of heat at a microscopic level.

If the picture is fairly clear for localized body, what about objects which are spread, such as fluids or fields ? For a fluid it is assumed that the momentum, and then the energy, are transferred by contact. But it is acknowledged that there are forces which act without any contact, and then we have to assume that, if the transfers occurs, force fields also carry momentum, evidenced by their action on material bodies, and exchange energy.

A momentum seems to be a relative concept : the motion is measured with respect to some frame. However we will see that, in General Relativity, the concept of material body leads to acknowledge a quantity similar to an absolute momentum : the spinor, which incorporates both the translational and rotational motion. The momentum is then a geometric quantity, which exists independently of any frame, but its measure, by components in a frame, is relative, and changes with the observer.

In Thermodynamics the Energy is also a relative concept : this is a state variable, defined up to a constant, and which can be positive or negative (depending on which object is considered) : only the exchange of energy has a clear meaning. Relativity introduced the idea of equivalence of mass and energy with the celebrated formula $E = mc^2$. However it does not mean that there is an absolute level of energy : in a nuclear reaction the mass deficit is converted in energy which is transferred to other physical objects, particles or fields, and the mc^2 is the measure of a quantity which is exchanged, similarly to the kinetic energy $\frac{1}{2}mv^2$. The Quantum theory of Fields introduces a “ground state” which, by definition, would have a level 0 of energy, similar to the 3d law of Thermodynamics (the existence of a $0°K$ state). This issue is closely linked to the concept of “vacuum”, a region of the universe which would be devoid of any physical object, which is not without problems of its own. A force field propagates, and as far as one knows, the electromagnetic and the gravitational field have an infinite range, so a vacuum could exist only in an area out of reach of any field which have ever been emitted, which is, by definition, not accessible to any observation. And indeed in QTF the vacuum has quite an animated life with virtual particles. So it seems wiser to consider energy as a relative quantity, which is exchanged between a system, whatever it is, and its surroundings. This leads to the definition of potential energy : the energy which can be given or received by a physical object in the transformation of a system. So it has no meaning out of the context of a given system.

The concepts of energy, momentum, and of evolution of a system, require a clear definition of the time, which can depend on the observer. These issues were not solved until the advent of Relativity, which also gives a relation between energy and momentum. One feature to notice about this Principle is that it does not assume that the evolution is continuous : there are clearly two states of the system, differentiated by a time elapsed between the measures, but the process can be continuous or discontinuous. Then this is the difference between the values of energy and momentum at the beginning and at the end of the process which matters.

1.5.3 Principle of Least Action

As there are quantities which are globally conserved in a physical process, based upon experience, it is assumed that any system has privileged states, called states of equilibrium, from which it does not move without a change in its environment, for instance an external action. This concept still holds when one considers the evolution in time : equilibrium does not imply that the state of the system is frozen, it can move along a path from which it does not differ easily. This is the generalization of the idea that an isolated system is in the state of least energy. States of equilibrium can be achieved by a continuous or a discontinuous process : the Principle of Least Action does not tell how a state of equilibrium is reached, only what are its characteristics.

From Mechanics, this principle is usually represented in Physics by the fact that a scalar functional, the action, is stationary for the values corresponding to the state of equilibrium : $\ell\left(L\left(z^i, z^i_\alpha, z^i_{\alpha\beta}, ...\right)\right)$ where $Z = z^i, z^i_\alpha, z^i_{\alpha\beta}, ...$are the variables and their partial derivatives and L a scalar function (the scalar lagrangian).

It comes from Analytic Mechanics where $L = \sum_i \frac{1}{2} m_i v_i^2 - U$ is the total energy of the system (Kinetic and potential) and the lagrangian has the general meaning of a density of energy / momentum, as described above.

The stationary means that for any (small) changes δZ of the value of the variables around the equilibrium Z_0 the value of the functional ℓ is unchanged. So this is not necessarily a maximum or a minimum, even local. And a state of equilibrium is not necessarily unique.

Whenever the variables are maps or functions defined over the area Ω of a manifold endowed with a volume measure ϖ the functional is assumed to be an integral :

$\int_\Omega L\left(z^i(m), z^i_\alpha(m), z^i_{\alpha\beta}(m), ...\right)\varpi(m)$

More simply when the variables are functions of the time only, varying from t_1 to t_2, the action reads:

$\int_{t_1}^{t_2} L\left(z^i(t), z^i_\alpha(t), z^i_{\alpha\beta}(t), ...\right)dt$

This formulation is extensively used, and most of the laws in Physics can be expressed this way. The Principle does not tell anything about the lagrangian, in which lies the physical content. There are constraints on its expression, due to the Principle of Relativity, but the choice of the right lagrangian is mostly an art, which of course must be checked by the consequences that can be deduced.

The Principle seems to introduce a paradox in that the values taken by the variables at any moment depend on the values on the whole evolution of the system, that is on the values which will be taken in the future. But this paradox stems from the model itself : at the very beginning the physicist assumes that the variables which are measured or computed belong to some class of objects which are defined all over Ω. So *the variables are the maps and not the values that they take for each value of their arguments.*

The physical quantity represented by L in the action is usually seen as the total energy of the system but, it is actually the sum of the energy exchanged between the components of the system, and if actions exterior to the system are involved, they should be accounted for (the are then known) in L. So the concept of equilibrium is that of a global balance between the physical objects considered.

1.5.4 Second Principle of Thermodynamics and Entropy

The universality of scientific laws implies that experiments are reproducible, time after time, which requires either that the circumstances stay the same, or can be reproduced identically. This can be achieved only to some degree, controlled by checking all the parameters which could influence the results. It is assumed that the parameters which are not directly involved in the law which is tested are not significant, or keep a steady value, in time as well as in the domains which are exterior to the area which is studied. So universality implies some continuity of the phenomena.

Many discontinuous phenomena at a macroscopic scale can be explained as the result of continuous processes at a smaller scale : an earthquake is the result of the slow motion of tectonic plates. Others involve the transition between phases, which are themselves states of equilibrium, and can be explained, as we will see, by the interaction of microsystems. However discontinuous processes exist, but they are never totally discontinuous : what happens is the transition between different continuous processes. Brownian motion is modelled by patching continuous paths. And indeed totally discontinuous functions are a mathematical curiosity, not easy to build. So the maps involved in physical models can be safely assumed to be continuous, except at isolated points.

If a transition occurs between two states of equilibrium the Principle of Least action can be implemented for each of them. But this leaves several issues.

The first is about the concept of equilibrium itself. Stationary does not mean frozen, the Principle of Least Action encompasses systems whose state varies with the time, so it addresses also processes. But in Physics there are common restrictions imposed on these processes. In Thermodynamics equilibrium is identified with reversible processes, seen as slow processes : at any moment the system is close to equilibrium. In Theoretical Physics these are processes whose evolution is ruled by equations which are invariant by time reversal : if $X(t)$ is solution, then the replacement of t by $-t$ is still a solution. A reversible process is determinist (there is only one path to go from a state to another) but the converse is not true. The Second Principle of Thermodynamics is a way to study processes which do not meet these restrictions.

In Thermodynamics the Second Principle is based upon the equation :

$dU = TdS - pdV + \sum_c \mu_c dN_c$

where the internal energy U, the entropy S, the volume V and the number of moles of chemical species N_c are variables which characterize the state of the system. The key point is that they do not depend on the path which has been followed to arrive at a given state. In the evolution between states:

$dU = \delta Q + \delta W$

where $\delta Q, \delta W$ are the quantity of heat and work exchanged by the system with its surroundings during any evolution. The variable temperature T is a true thermodynamic variable : it has a meaning only at a macroscopic scale. The symbol d represents a differential, meaning that the corresponding state variables are differentiable, and thus continuous, and δ a variation, which can be discontinuous.

For a system in any process :

$dS \geq \frac{\delta Q}{T}$

so that for isolated systems $dS \geq 0$: their entropy can only increase and this defines an arrow of time. We have an equality only in reversible processes.

The Thermodynamics formulation can be generalized to the evolution of systems comprised of many interacting microsystems. The model, proposed first by Boltzmann and Gibbs, has been used with many variants, notably by E.T. Jayes in his Principle of Maximum Entropy in relation with Information Theory. Its most common formalization is the following. A system is comprised of N (a large number) identical microsystems. Their states are represented by a random variable $X = (X_a)_{a=1}^m$ valued in a domain Ω with an unknown probability law $\Pr(X_1 = x_1, ...X_N = x_N) = \rho(x_1, ..x_n)$. There are m macroscopic variables $(Y_k)_{k=1}^m$ which can be measured for the whole system, whose value depend on the states of the microsystems :$Y_k = f_k(x_1, ...x_N)$. Knowing the values $\left(\widehat{Y}_k\right)_{k=1}^m$ observed, the problem is to estimate ρ.

The Principle of Maximum Entropy states that the law ρ is such that the integral :

$S = \int_\Omega -\rho(x_1, ..x_N) \ln \rho(x_1, ..x_N) dx_1...dx_N$

over the domain Ω of the x_a is maximum, under the constraints :

$\widehat{Y}_k = f_k(x_1, ...x_N)$

$\int_\Omega \rho dx = 1$

The solution of this problem leads to the introduction of m new variables $(\theta_k)_{i=1}^m$ (the Lagrange parameters) dual of the observables Y_k which are truly thermodynamic : they have no meaning for the microsystems. Temperature is the dual of energy.

So formulated we have a classic problem of Statistics, and we can give a more precise definition of a reversible process. If the process is such that :

- the state of a microsystem does not depend on the state of other microsystems, only on the state of the global system

- the collection $(Y_k)_{k=1}^m$ is a complete statistic (one cannot expect to have more information on the system by adding another macroscopic variable)

then it is not difficult, using the Pitman-Koopman-Darmois theorem, to show that the solution given by the Principle of Maximum Entropy is indeed a good maximum of likehood estimator, and validates it, at least for computational purposes.

This scheme has been extended in the framework of QM, the quantity $-Tr[\rho \ln \rho]$ called the information entropy, becoming a functional and ρ an operator on the space of states.

So, for reversible processes, there is a model which is well grounded, and entropy has a clear meaning. It leads to the introduction of additional variables (notably temperature) which can be measured but have a meaning only for the whole system. But it is usually acknowledged that there is no satisfying general model for non reversible processes, or processes which involve disequilibrium (see G.Röpke for more).

However the concept of entropy has a real significance to answer the issue of the choice of a state of equilibrium, if several are possible in the implementation of the Principle of Least Action : it tells why one of them is privileged, and the driving force is the entropy.

If disequilibrium processes are of importance at a macroscopic scale, due to their prevalence in practical problems, this is less so from a theoretical point of view. There is no obvious reason to focus on processes which are modelled by equations invariant by time reversal, and actually they are not in Quantum Theory of Fields. The issue of determinism, closely related to discontinuous processes, seem more important.

There is another Principle acknowledged in Physics, which is related to the definition of reversible process : the laws of Physics are assumed to be invariant by the CPT operations. As their definition involves a precise framework, it will be given in the Chapter 5.

1.5.5 Principle of Locality

It can be stated as : "the outcome of any physical process occurring at a location depends only of the values of the involved physical quantities at this location". So it prohibits actions at a distance. This is obvious in the lagrangian formulation of the Principle of Least Action : the integral is computed from data whose values are taken at each point m (but one can conceive of other functionals ℓ).

Any physical theory assumes the existence of material objects, whose main characteristic is that they are localized : they are at a definite place at a given time. To account for phenomena such as electromagnetism or gravity, the principle requires the existence of physical objects, the force fields, which have a value at any point. Thus this principle is consubstantial to the distinction matter / fields. It does not prohibit by itself the existence of objects which are issued from fields and behave like matter (the bosons). And similarly it does not forbid the representation of material bodies in a formalism which is defined at any point : in Mechanics the trajectory of a material point is a map $x(t)$ defined over a period of time. But these features appear in the representation of the objects, and do not imply physical action at a distance. The validity of this principle has been challenged by the entanglement of states of bosons, but it seems difficult to accept that it is false, as most of the Physics use it.

Because any measurement involves a physical process, the principle of locality implies that the measures shall be done locally, that is by observers at each location. This does not preclude the observers to exchange their information, but requires a procedure to collect and compare these measures. This procedure is part of the system, and the laws that they represent. As it has been said before, the observer, if he is not by himself submitted to the phenomena that he measures (he has free will), has distinctive characteristics which must be accounted for in the formulation of a law. So the Principle of Locality requires the definition of rules which tell how measures done by an observer at a location can be compared to measures done by an observer at another location.

1.5.6 Principle of causality

The Principle of Causality exists since the beginning of Philosophy, and it would seem to belong more to the rules for rationale discourse than to Physics. However it introduces, one way or another, a critical component which is a relation in time. A phenomenon A is the cause of another B if it manifests before B. And this is more than a simple timing : it is accepted than two phenomena can be not related. In Classic Physics the use of a time coordinates is usually sufficient to account for the potential causality. In Quantum Mechanics this is more sensitive : almost all reasoning is based on the comparison between an initial state and a final state, which requires the possibility to identify non ambiguously these states, that is a set of measures, related to a set of phenomena, which can be considered as the potential causes or the results of an experiment. Relativity introduced a disturbing element : simultaneity is no longer universal and depends on the observer. The Principle of Causality adds a specific structure in the representation of the geometry of the universe, which is clearly explained by the existence of a metric. However this leads to much complication in Quantum Mechanics.

The rest of this book will be in some way a practical illustration of this first chapter. We will successfully expose the Geometry of General Relativity, the Kinematics of material bodies, the

Force fields, the Interactions Fields / Particles, the Bosons. Starting from facts, common or scientific known facts, we will make assumptions, then, using the right mathematical formalism and Fundamental principles, we will predict scientific laws, as theorems. And this is the experimental verification of these laws that will provide the validity of the theory. So this is very different, almost the opposite, of what is usually done in Physics Books, such as Feynman's, where the starting point is almost always an experiment. The next chapter, dedicated to Quantum Theory, is purely mathematical but, as we will see, it starts also by the construction of its own objects : physical models.

Chapter 2

QUANTUM MECHANICS

Quantum Physics encompasses several theories, with three distinct areas:

i) Quantum Mechanics (QM) proper, which, since the seminal von Neumann's book, is expressed as a collection of axioms, such as summarized by Weinberg :

- Physical states of a system are represented by vectors ψ in a Hilbert space H, defined up to a complex number (a ray in a projective Hilbert space)

- Observables are represented by Hermitian operators

- The only values that can be observed for an operator are one of its eigen values λ_k corresponding to the eigen vector ψ_k

- The probability to observe λ_k if the system is in the state ψ is proportional to $|\langle\psi, \psi_k\rangle|^2$

- If two systems with Hilbert space H_1, H_2 interact, the states of the total system are represented in $H_1 \otimes H_2$

and, depending on the authors, the Schrödinger's equation.

ii) Wave Mechanics, which states that particles can behave like fields which propagate, and conversely force fields can behave like pointwise particles. Moreover particles are endowed with a spin. In itself it constitutes a new theory, with the introduction of new concepts related to physical objects (spin, photon), for which QM is the natural formalism. Actually this is essentially a theory of electromagnetism, and is formalized in Quantum Electrodynamics (QED).

iii) The Quantum Theory of Fields (QTF) is a theory which encompasses theoretically all the phenomena at the atomic or subatomic scale, but has been set up mainly to deal with the other forces (weak and strong interactions) and the organization of elementary particles. It uses additional concepts (such as gauge fields) and formalism and computation rules (Feynman diagrams, path integrals).

I will address in this chapter *QM only*. It would seem appropriate to begin the Physics part of this book by QM, as it has been dominant and pervasive since 70 years. But actually it is the converse : the place of this chapter comes from the fact that QM is not a physical theory. This is obvious with a look at the axioms : they do not define any physical object, or physical property (if we except the Schrödinger's equation which is or not part of the corpus). They are deemed valid for any system and, actually, they would not be out of place in a book on Economics. These axioms, which are used commonly, are not Physical Laws, and indeed they are not falsifiable (how could we check that an observable is a Hermitian operator ?). Some, whose wording is general, could be seen as Fundamental Laws, similar to the Principle of Least Action, but others have an almost supernatural precision (the eigen vectors). Nevertheless they are granted with a total infallibility, supported by an unshakable faith, lauded by the media as well as the Highest Academic Authorities, reputed to make incredibly precise predictions. Their power is limited only by a scale which is not even mentioned and which is impossible to compute.

This strange status, quite unique in Science, is at the origin of the search for interpretations, and for the same reason, makes so difficult any sensible discussion on the topic. Actually these axioms have emerged slowly from the practices of great physicists, kept without any change in the last decenniums, and endorsed by the majority, mostly because, from their first Physics 101 to the software that they use, it is part of their environment. I will not enter into a debate about the interpretations of these axioms, but it is necessary to evoke the attempts which have been made to address directly their foundations.

In seminal books and articles, von Neumann and Birkhoff have proposed a new direction to understand and justify these axioms. Their purpose was, from general considerations, to set up a Formal System, actually similar to what is done in Mathematics for Arithmetic or Sets Theory, in which the assertions done in Physics can be expressed and used in the predictions of experiments, and so granting to Physics a status which would be less speculative and more respecting of the facts as they can actually be established. This work has been pursued, notably by Jauch, Haag, Varadarajan and Francis in the recent years. An extension which accounts for Relativity has been proposed by Wightman and has been developed as an Axiomatic Quantum Field Theory (Haag, Araki, Halvorson, Borchers, Doplicher, Roberts, Schroer, Fredenhagen, Buchholz, Summers, Longo,...). It assumes the existence of the formalism of Hilbert space itself, so the validity of most of the axioms, and emphasizes the role of the algebra of operators. Since all the information which can be extracted from a system goes through operators, it can be conceived to define the system itself as the set of these operators. This is a more comfortable venue, as it is essentially mathematical, which has been studied by several authors (Bratelli and others). Recently this approach has been completed by attempts to link QM with Information Theory, either in the framework of Quantum Computing, or through the use of the Categories Theory.

These works share some philosophical convictions, supported with a strength depending on the authors, but which are nonetheless present :

i) A deep mistrust with regard to realism, the idea that there is a real world, which can be understood and described through physical concepts such as particles, location,...At best they are useless, at worst they are misleading.

ii) A great faith in the mathematical formalism, which should ultimately replace the concepts.

iii) The preeminence of experimentation over theories : experimental facts are seen as the unique source of innovation, physical laws are essentially the repeated occurrences of events whose correlation must be studied by statistical methods, the imperative necessity to consider the conditions in which the experiments can or cannot be made.

As any formal system, the axiomatic QM defines its own objects, which are basically the assertions that a physicist can make from the results of experiments ("the yes-no experiments" of Jauch), and set up a system of rules of inference according to which other assertions can be made, with a special attention given to the possibility to make simultaneous measures, and the fact that any measure is the product of a statistical estimation. With the addition of some axioms, which obviously cannot reflect any experimental work (it is necessary to introduce infinity), the formal system is then identified, by a kind of structural isomorphism, with the usual Hilbert space and its operators of Mathematics. And from there the axioms of QM are deemed to be safely grounded.

One can be satisfied or not by this approach. But some remarks can be done.

In many ways this attempt is similar to the one by which mathematicians tried to give an ultimate, consistent and logical basis to Mathematics. Their attempt has not failed, but have shown the limits of what can be achieved : the necessity to detach the objects of the formal system from any idealization of physical objects, the non unicity of the axioms, and the fact that they are justified by experience and efficiency and not by a logical necessity. The same limits are obvious in axiomatic QM. If to acknowledge the role of experience and efficiency in the foundations of the system should not be disturbing, the pretense to enshrine them in axioms, not refutable and not subject to verification, places a great risk to the possibility of any evolution. And indeed the axioms have not changed for more than 50 years, without stopping the controversies about their meaning. The unavoidable replacement of physical concepts, identification of physical objects and their properties, by formal and abstract objects, which is consistent with the philosophical premises, is specially damaging in Physics. Because there is always a doubt about the meaning of the objects (for instance it is quite impossible to find the definition of a "state") the implementation of the system sums up practically to a set of "generally accepted computations", it makes its learning and teaching perilous (the Feynmann's affirmation that it cannot be understood), and eventually to the recurring apparitions of "unidentified physical objects" whose existence is supposed to fill the gap. In many ways the formal system has replaced the Physical Theories, that is a set of objects, properties and behaviors, which can be intuitively identified and understood. The Newton's laws of motion are successful, not only because they can be checked, but also because it is easy to understand them. This is not the case for the decoherence of the wave function...

Nevertheless, this attempt is right in looking for the origin of these axioms in the critique (in the Kantian meaning) of the method specific to Physics. But it is aimed at the wrong target : the concepts are not the source of the problems, they are and will stay necessary because they make the link between formalism and real world, and are the field in which new ideas can germinate. And the solution is not in a sanctification of the experiments, which are too diverse to be submitted to any analytical method. Actually these attempts have missed a step, which always exists between the concepts and the collection of data : the mathematical formalization itself, in models. Models, because they use a precise formalism, can be easily analyzed and it is possible to show that, indeed, they have specific properties of their own, which do not come from the reality they represent, but from their mathematical properties and the way they are used. The objects of an axiomatic QM, if one wishes to establish one, are then clearly identified, without disturbing the elaboration or the implementation of theories. The axioms can then be proven, they can also be safely used.

QM is about the representation of physical phenomena, and not a representation of these phenomena (as can be Wave Mechanics, QED or QTF). It expresses properties of the data which can be extracted from measures of physical phenomena but not properties of physical objects. To sum up : QM is not about how the physical world works, it is about how it looks.

2.1 HILBERT SPACE

2.1.1 Representation of a system

Models play a central role in the practical implementation of a theory to specific situations. It will be our starting point.

Let us start with common Analytic Mechanics. A system, meaning a delimited area of space comprising material bodies, is represented by scalar generalized coordinates $q = (q_1, ..., q_N)$ its evolution by the derivatives $q' = (q_1', ..., q_N')$. By extension q can be the coordinates of a point Q of some manifold M to account for additional constraints, and then the state of the system at a given time is fully represented by a point of the vector bundle $TM : W = (Q, V_Q)$. By mathematical transformations the derivatives q' can be exchanged with conjugate momenta, and the state of the system is then represented in the phase space, with a symplectic structure. But we will not use this addition and stay at the very first step, that is the representation of the system by (q, q') .

Trouble arises when one considers the other fundamental objects of Physics : force fields. By definition their value is defined all over the space x time. So in the previous representation one should account, at a given time, for the value of the fields at each point, and introduce unaccountably infinitely many coordinates. This issue has been at the core of many attempts to improve Analytic Mechanics.

But let us consider two facts :

- Analytic Mechanics, as it is usually used, is aimed at representing the evolution of the system over a period of time $[0, T]$, as it is clear in the Lagrangian formalism : the variable are accounted, together, for the duration of the experiment;

- the state of the system is represented by a map $W : [0, T] \to (Q, V_Q)$: the knowledge of this map sums up all that can be said on the system, the map itself represents the state of the system.

Almost all the problems in Physics involve a model which comprises the following :

i) a set of physical objects (material bodies or particles, force fields) in a delimited area Ω of space x time (it can be in the classical or the relativist framework) called the system;

ii) the state of the system is represented by a fixed finite number N of variables $X = (X_k)_{k=1}^N$ which can be maps defined on Ω , with their derivatives;

so that the state of the system is defined by a finite number of maps, which usually belong themselves to infinite dimensional vector spaces.

And it is legitimate to substitute the maps to the coordinates in Ω. We still have infinite dimensional vector spaces, but by proceeding first to an aggregation by maps, the vector space is more manageable, and we have some mathematical tools to deal with it. But we need to remind the definition of a manifold, a structure that we will use abundantly in the following (more in Maths.15.1.1).

2.1.2 Manifold

Let M be a set, E a topological vector space, an atlas, denoted $A = (O_i, \varphi_i, E)_{i \in I}$ is a collection of :

subsets $(O_i)_{i \in I}$ of M such that $\cup_{i \in I} O_i = M$ (this is a cover of M)

maps $(\varphi_i)_{i \in I}$ called **charts**, such that :

i) $\varphi_i : O_i \to U_i :: \xi = \varphi_i(m)$ is bijective and ξ are the coordinates of M in the chart

ii) U_i is an open subset of E

iii) $\forall i, j \in I : O_i \cap O_j \neq \varnothing$:

$\varphi_i(O_i \cap O_j), \varphi_j(O_i \cap O_j)$ are open subsets of E, and there is a bijective, continuous map, called a transition map :

$\varphi_{ij} : \varphi_i(O_i \cap O_j) \to \varphi_j(O_i \cap O_j)$

Notice that no mathematical structure of any kind is required on M. A topological structure can be imported on M, by telling that all the charts are continuous, and conversely if there is a topological structure on M the charts must be compatible with it. But the set M has no algebraic structure : a combination such as $am + bm'$ has no meaning.

Two atlas $A = (O_i, \varphi_i, E)_{i \in I}, A' = \left(O'_j, \varphi'_j, E\right)_{j \in J}$ of M are said to be compatible if their union is still an atlas. Which implies that :

$\forall i \in I, j \in J : O_i \cap O'_j \neq \varnothing : \exists \varphi_{ij} : \varphi_i\left(O_i \cap O'_j\right) \to \varphi'_j\left(O_i \cap O'_j\right)$ is a homeomorphism

The relation $A, A'$are compatible atlas of M, is a relation of equivalence. A class of equivalence is a **structure of manifold** on the set M.

The key points are :

- there can be different structures of manifold on the same set. On $\mathbb{R}^4$ there are unaccountably many non equivalent structures of smooth manifolds (this is special to $\mathbb{R}^4$: on $\mathbb{R}^n, n \neq 4$ all the smooth structures are equivalent !).

- all the interesting properties on M come from E : the dimension of M is the dimension of E (possibly infinite); if E is a Fréchet space we have a Fréchet manifold, if E is a Banach space we have a Banach manifold and then we can have differentials, if E is a Hilbert space we have a Hilbert manifold, but these additional properties require that the transition maps φ_{ij} meet additional properties.

- for many sets several charts are required (a sphere requires at least two charts) but an atlas can have only one chart, then the manifold structure is understood as the same point M will be defined by a set of compatible charts.

The usual, euclidean, 3 dimensional space of Physics is an affine space. It has a structure of manifold, which can use an atlas with orthonormal frames, or with curved coordinates (spherical or cylindrical). Passing from one system of coordinates to another is a change of charts, and represented by transition maps φ_{ij}.

2.1.3 Fundamental theorem

In this chapter we will consider models which meet the following conditions:

Condition 1 *i) The system is represented by a fixed finite number N of variables* $(X_k)_{k=1}^N$

ii) Each variable belongs to an open subset O_k *of a separable Fréchet real vector space* V_k

iii) At least one of the vector spaces $(V_k)_{k=1}^N$ *is infinite dimensional*

iv) For any other model of the system using N variables $(X'_k)_{k=1}^N$ *belonging to open subset* O'_k *of* V_k*, and for* $X_k, X'_k \in O_k \cap O'_k$ *there is a continuous map :* $X'_k = \digamma_k(X_k)$

Remarks :

i) The *variables must be vectorial.* This condition is similar to the superposition principle which is assumed in QM. This is one of the most important condition. By this we mean that the associated physical phenomena can be represented by vectors (or tensors, or scalars). The criterion, to check if this is the case, is : if the physical phenomenon can be represented by X and X', does the phenomenon corresponding to any linear combination $\alpha X + \beta X'$ has a physical meaning ?

Are usually vectorial variables : the speed of a material point, the electric or magnetic field, a force, a moment,...and the derivatives, which are, by definition, vectors.

Are not usually vectorial variables : qualitative variables (which take discrete values), a point in the euclidean space or on a circle, or any surface. The point can be represented by coordinates, but these coordinates are not the physical object, which is the material point. For instance in Analytic Mechanics the coordinates $q = (q_1, ..., q_N)$ are not a geometric quantity : usually a linear combination $\alpha q + \beta q'$ has no physical meaning (think to polar coordinates). The issue arises because physicists are used to think in terms of coordinates (in euclidean or relativist Lorentz frame) which leads to forget that the coordinates are just a representation of an object which, even in its mathematical form (a point in an affine space) is not vectorial.

So this condition, which has a simple mathematical expression, has a deep physical meaning : it requires to understand clearly why the properties of the physical phenomena can be represented by a vectorial variable, and reaches the most basic assumptions of the theory. The status, vectorial or not, of a quantity is not something which can be decided at will by the Physicist : it is part of the Theory which he uses to build his model. However the addition of a variable which is not a vector can be very useful (Theorem 24).

ii) The variables are assumed to be independent, in the meaning that there is no given relation such that $\sum_k X_k = 1$. Of course usually the model is used with the purpose to compute or check relations between the variables, but these relations do not matter here. Actually to check the validity of a model one considers all the variables, those which are given and those which can be computed, they are all subject to measures and this is the comparison, after the experiment, between computed values and measured values which provides the validation. So in this initial stage of specification of the model there is no distinction between the variables, which are on the same footing.

Similarly there is no distinction between variables internal and external to the system : if the evolution of a variable is determined by the observer or by phenomena out of the system (it is external) its value must be measured to be accounted for in the model, so it is on the same footing as any other variable. And it is assumed that the value of all variables can be measured.

The derivative $\frac{dX_k}{dt}$ (or partial derivative at any order) of a variable X_k is considered as an independent variable, as it is usually done in Analytic Mechanics and in the mathematical formalism of r-jets.

iii) The variables can be restricted to take only some range (for instance it must be positive). The vector spaces are infinite dimensional whenever the variables are functions. The usual case is when they represent the evolution of the system with the time t : then X_k is the function itself : $X_k : \mathbb{R} \to O_k :: X_k(t)$. What we consider here are variables which cover the whole evolution of the system over the time, and not only just a snapshot $X_k(t)$ at a given time. But the condition encompasses other cases, notably fields F which are defined over a domain Ω. *The variables are the maps $F_k : \Omega \to O_k$ and not their values $F_k(\xi)$ at a given point $\xi \in \Omega$.*

iv) A Fréchet space is a Hausdorff, complete, topological space endowed with a countable family of semi-norms (Maths.971). It is locally convex and metric.

Are Fréchet spaces :

- any Banach vector space : the spaces of bounded functions, the spaces $L^p(E, \mu, \mathbb{C})$ of integrable functions on a measured space (E, μ) (Maths.2270), the spaces $L^p(M, \mu, E)$ of integrable sections of a vector bundle (valued in a Banach E) (Maths.2276)

- the spaces of continuously differentiable sections on a vector bundle (Maths.2310), the spaces of differentiable functions on a manifold (Maths.2314).

A topological vector space is separable if it has a dense countable subset (Maths.590) which, for a Fréchet space, is equivalent to be second countable (Maths.698). A totally bounded ($\forall r > 0$ there is a finite number of balls which cover V), or a connected locally compact Fréchet space, is separable (Maths.702, 703). The spaces $L^p(\mathbb{R}^n, dx, \mathbb{C})$ of integrable functions for $1 \leq p < \infty$, the spaces of continuous functions on a compact domain, are separable (Lieb).

Thus this somewhat complicated specification encompasses most of the usual cases.

In the following of this book we will see examples of these spaces : they are mostly maps :

$X : \Omega \to E$ from a relatively compact subset Ω of a manifold M to a finite dimensional vector space, endowed with a norm. Then the space of maps such that $\int_{\Omega} \|X(m)\| \varpi(m) < \infty$ where ϖ is a measure on M (a volume measure) is an infinite dimensional, separable, Fréchet space.

v) The condition iv addresses the case when the variables are defined over connected domains. But it implicitly tells that any other set of variables which represent the same phenomena are deemed compatible with the model.

The set of all potential states of the system is then given by the set $S = \left\{ (X_k)_{k=1}^N, X_k \in O_k \right\}$. If there is some relation between the variables, stated by a physical law or theory, its consequence is to restrict the domain in which the state of the system will be found, but as said before we stay at the step before any experiment, so O_k represents the set of all possible values of X_k.

Theorem 2 *For any system represented by a model meeting the conditions 1, there is a separable, infinite dimensional, Hilbert space H, defined up to isomorphism, such that $\mathcal{S}$ can be embedded as an open subset $\Omega \subset H$ which contains 0 and a convex subset.*

Proof. i) Each value of the set $\mathcal{S}$ of variables defines a state of the system, denoted X, belonging to the product $O = \prod_{k=1}^{N} O_k \subset V = \prod_{k=1}^{N} V_k$. The couple (O, X), together with the property iv) defines the structure of a Fréchet manifold M on the set $\mathcal{S}$, modelled on the Fréchet space $V = \prod_{k1}^{N} V_k$. The coordinates are the values $(x_k)_{k=1}^N$ of the functions X_k. This manifold is infinite dimensional. Any Fréchet space is metric, so V is a metric space, and M is metrizable.

ii) As M is a metrizable manifold, modelled on an infinite dimensional separable Fréchet space, the Henderson's theorem (Henderson - corollary 5, Maths.1386) states that it can be embedded as a open subset Ω of an infinite dimensional separable Hilbert space H, defined up to isomorphism. Moreover this structure is smooth, the set $H - \Omega$ is homeomorphic to H, the border $\partial\Omega$ is homeomorphic to Ω and its closure $\overline{\Omega}$.

iii) Translations by a vector are isometries. Let us denote $\langle\rangle_H$ the scalar product on H (this is a bilinear symmetric positive definite form). The map : $\Omega \to \mathbb{R} :: \langle \psi, \psi \rangle_H$ is bounded from below and continuous, so it has a minimum (possibly not unique) ψ_0 in Ω. By translation of H with ψ_0 we can define an isomorphic structure, and then assume that 0 belongs to Ω. There is a largest convex subset of H contained in Ω, defined as the intersection of all the convex subset contained in Ω. Its interior is an open convex subset C. It is not empty : because 0 belongs to Ω which is open in H, there is an open ball $B_0 = (0, r)$ contained in Ω. ■

So the state of the system can be represented by a single vector ψ in a Hilbert space.

From a practical point of view, often V itself can be taken as the product of Hilbert spaces, notably of square summable functions such as $L^2(\mathbb{R}, dt)$ which are separable Hilbert spaces and then the proposition is obvious.

If the variables belong to an open O' such that $O \subset O'$ we would have the same Hilbert space, and an open Ω' such that $\Omega \subset \Omega'$. V is open so we have a largest open $\Omega_V \subset H$ which contains all the Ω.

Notice that this is a real vector space.

The interest of Hilbert spaces lies with Hilbertian basis, and we now see how to relate such basis of H with a basis of the vector space V. It will enable us to show a linear chart of the manifold M.

2.1.4 Basis

Theorem 3 *For any basis $(e_i)_{i\in I}$ of V contained in O, there are unique families $(\varepsilon_i)_{i\in I}, (\phi_i)_{i\in I}$ of independent vectors of H, a linear isometry $\Upsilon : V \to H$ such that :*
$\forall X \in O : \Upsilon(X) = \sum_{i\in I} \langle \phi_i, \Upsilon(X)\rangle_H \varepsilon_i \in \Omega$
$\forall i \in I : \varepsilon_i = \Upsilon(e_i)$
$\forall i, j \in I : \langle \phi_i, \varepsilon_j\rangle_H = \delta_{ij}$
and Υ is a compatible chart of M.

Proof. i) Let $(e_i)_{i\in I}$ be a basis of V such that $e_i \in O$ and $V_0 = Span(e_i)_{i\in I}$. Thus $O \subset V_0$.

Any vector of V_0 reads : $X = \sum_{i\in I} x_i e_i$ where only a finite number of x_i are non null. Or equivalently the following map is bijective :

$\pi_V : V_0 \to \mathbb{R}^I_0 :: \pi_V\left(\sum_{i\in I} x_i e_i\right) = x = (x_i)_{i\in I}$

where the set $\mathbb{R}^I_0 \subset \mathbb{R}^I$ is the subset of maps $I \to \mathbb{R}$ such that only a finite number of components x_i are non null.

(O, X) is an atlas of the manifold M and M is embedded in H, let us denote $\Xi : O \to \Omega$ a homeomorphism accounting for this embedding.

The inner product on H defines a positive kernel :

$K : H \times H \to \mathbb{R} :: K(\psi_1, \psi_2) = \langle \psi_1, \psi_2\rangle_H$

Then $K_V : O \times O \to \mathbb{R} :: K_V(X, Y) = K(\Xi(X), \Xi(Y))$ defines a positive kernel on O (Math.1196).

K_V defines a definite positive symmetric bilinear form on V_0, denoted $\langle\rangle_V$, by :

$\left\langle \sum_{i\in I} x_i e_i, \sum_{i\in I} y_i e_i\right\rangle_V = \sum_{i,j\in I} x_i y_j K_{ij}$ with $K_{ij} = K_V(e_i, e_j)$

which is well defined because only a finite number of monomials $x_i y_j$ are non null. It defines a norm on V_0.

ii) Let : $\varepsilon_i = \Xi(e_i) \in \Omega$ and $H_0 = Span(\varepsilon_i)_{i\in I}$ the set of finite linear combinations of vectors $(\varepsilon_i)_{i\in I}$. It is a vector subspace (Math.901) of H. The family $(\varepsilon_i)_{i\in I}$ is linearly independent, because, for any finite subset J of I, the determinant

$\det\left[\langle \varepsilon_i, \varepsilon_j\rangle_H\right]_{i,j\in J} = \det\left[K_V(e_i, e_j)\right]_{i,j\in J} \neq 0.$

Thus $(\varepsilon_i)_{i\in I}$ is a non Hilbertian basis of H_0.

H_0 can be defined similarly by the bijective map :

$\pi_H : H_0 \to \mathbb{R}^I_0 :: \pi_H\left(\sum_{i\in I} y_i \varepsilon_i\right) = y = (y_i)_{i\in I}$

iii) By the Gram-Schmidt procedure (which works for infinite sets of vectors) it is always possible to built an orthonormal basis $(\widetilde{\varepsilon}_i)_{i\in I}$ of H_0 starting with the vectors $(\varepsilon_i)_{i\in I}$ indexed on the same set I (as H is separable I can be assimilated to $\mathbb{N}$).

$\ell^2(I) \subset \mathbb{R}^I$ is the set of families $y = (y_i)_{i\in I} \subset \mathbb{R}^I$ such that :

$\sup\left(\sum_{i\in J} (y_i)^2\right) < \infty$ for any countable subset J of I.

$\mathbb{R}^I_0 \subset \ell^2(I)$

The map : $\chi : \ell^2(I) \to H_1 :: \chi(y) = \sum_{i\in I} y_i \widetilde{\varepsilon}_i$ is an isomorphism to the closure $H_1 = \overline{Span(\widetilde{\varepsilon}_i)_{i\in I}} = \overline{H_0}$ of H_0 in H (Math.1121). H_1 is a closed vector subspace of H, so it is a Hilbert space. The linear span of $(\widetilde{\varepsilon}_i)_{i\in I}$ is dense in H_1, so it is a Hilbertian basis of H_1 (Math.1122).

Let $\pi : H \to H_1$ be the orthogonal projection on H_1 : $\|\psi - \pi(\psi)\|_H = \min_{u\in H_1} \|\psi - u\|_H$ then :

$\psi = \pi(\psi) + o(\psi)$ with $o(\psi) \in H_1^\perp$ which implies : $\|\psi\|^2 = \|\pi(\psi)\|^2 + \|o(\psi)\|^2$

There is a open convex subset, containing 0, which is contained in Ω so there is $r > 0$ such that :

$\|\psi\| < r \Rightarrow \psi \in \Omega$ and as $\|\psi\|^2 = \|\pi(\psi)\|^2 + \|o(\psi)\|^2 < r^2$

then $\|\psi\| < r \Rightarrow \pi(\psi), o(\psi) \in \Omega$

$o(\psi) \in H_1^{\perp}, H_0 \subset H_1 \Rightarrow o(\psi) \in H_0^{\perp}$
$\Rightarrow \forall i \in I : \langle \varepsilon_i, o(\psi)\rangle_H = 0 = K_V\left(\Xi^{-1}(\varepsilon_i), \Xi^{-1}(o(\psi))\right) = K_V\left(e_i, \Xi^{-1}(o(\psi))\right)$
$\Rightarrow \Xi^{-1}(o(\psi)) = 0 \Rightarrow o(\psi) = 0$
$H_1^{\perp} = 0$ thus H_1 is dense in H (Math.1115), and as it is closed : $H_1 = H$
$(\widetilde{\varepsilon}_i)_{i\in I}$ is a Hilbertian basis of H and
$\forall \psi \in H : \psi = \sum_{i\in I} \langle \widetilde{\varepsilon}_i, \psi\rangle_H \widetilde{\varepsilon}_i$ with $\sum_{i\in I} |\langle \widetilde{\varepsilon}_i, \psi\rangle_H|^2 < \infty$
$\Leftrightarrow (\langle \widetilde{\varepsilon}_i, \psi\rangle_H)_{i\in I} \in \ell^2(I)$
H_0 is the interior of H, it is the union of all open subsets contained in H, so $\Omega \subset H_0$
$H_0 = Span\left((\widetilde{\varepsilon}_i)_{i\in I}\right)$ thus the map :
$\widetilde{\pi}_H : H_0 \to \mathbb{R}_0^I :: \widetilde{\pi}_H\left(\sum_{i\in I} \widetilde{y}_i \widetilde{\varepsilon}_i\right) = \widetilde{y} = (\widetilde{y}_i)_{i\in I}$
is bijective and : $\widetilde{\pi}_H(H_0) = \widetilde{R}_0 \subset \mathbb{R}_0^I \subset \ell^2(I)$
Moreover : $\forall \psi \in H_0 : \widetilde{\pi}_H(\psi) = (\langle \widetilde{\varepsilon}_i, \psi\rangle_H)_{i\in I} \in \mathbb{R}_0^I$
Thus :
$\forall X \in O : \Xi(X) = \sum_{i\in I} \langle \widetilde{\varepsilon}_i, \Xi(X)\rangle_H \widetilde{\varepsilon}_i \in \Omega$
and $\widetilde{\pi}_H(\Xi(X)) = (\langle \widetilde{\varepsilon}_i, \Xi(X)\rangle_H)_{i\in I} \in \widetilde{R}_0$
$\forall i \in I, e_i \in O \Rightarrow \Xi(e_i) = \varepsilon_i = \sum_{j\in I} \langle \widetilde{\varepsilon}_j, \varepsilon_i\rangle_H \widetilde{\varepsilon}_j$
and $\widetilde{\pi}_H(\varepsilon_i) = \left(\langle \widetilde{\varepsilon}_j, \varepsilon_i\rangle_H\right)_{j\in I} \in \widetilde{R}_0$
iv) Let be : $\widetilde{e}_i = \Xi^{-1}(\widetilde{\varepsilon}_i) \in V_0$ and $\mathcal{L}_V \in GL(V_0; V_0) :: \mathcal{L}_V(e_i) = \widetilde{e}_i$
We have the following diagram :

$$\begin{array}{ccccccc} & \Xi & & \mathcal{L}_H^{-1} & & & \\ e_i & \to & \varepsilon_i & \to & \widetilde{\varepsilon}_i & & \\ & \searrow & & & \downarrow & & \\ & \mathcal{L}_V & \searrow & & \downarrow & \Xi^{-1} & \\ & & & \searrow & \downarrow & & \\ & & & & \widetilde{e}_i & & \end{array}$$

$\langle \widetilde{e}_i, \widetilde{e}_j\rangle_V = \langle \Xi(\widetilde{e}_i), \Xi(\widetilde{e}_j)\rangle_H = \langle \widetilde{\varepsilon}_i, \widetilde{\varepsilon}_j\rangle_H = \delta_{ij}$
So $(\widetilde{e}_i)_{i\in I}$is an orthonormal basis of V_0 for the scalar product K_V
$\forall X \in V_0 : X = \sum_{i\in I} \widetilde{x}_i \widetilde{e}_i = \sum_{i\in I} \langle \widetilde{e}_i, X\rangle_V \widetilde{e}_i$ and $(\langle \widetilde{e}_i, X\rangle_V)_{i\in I} \in \mathbb{R}_0^I$
The coordinates of $X \in O$ in the basis $(\widetilde{e}_i)_{i\in I}$ are $(\langle \widetilde{e}_i, X\rangle_V)_{i\in I} \in \mathbb{R}_0^I$
The coordinates of $\Xi(X) \in H_0$ in the basis $(\widetilde{\varepsilon}_i)_{i\in I}$ are $(\langle \widetilde{\varepsilon}_i, \Xi(X)\rangle_H)_{i\in I} \in \mathbb{R}_0^I$
$\langle \widetilde{\varepsilon}_i, \Xi(X)\rangle_H = \langle \Xi(\widetilde{e}_i), \Xi(X)\rangle_H = \langle \widetilde{e}_i, X\rangle_V$
Define the maps :
$\widetilde{\pi}_V : V_0 \to \mathbb{R}_0^I :: \widetilde{\pi}_V\left(\sum_{i\in I} \widetilde{x}_i \widetilde{e}_i\right) = \widetilde{x} = (\widetilde{x}_i)_{i\in I}$
$\Upsilon : V_0 \to H_0 :: \Upsilon = \widetilde{\pi}_H^{-1} \circ \widetilde{\pi}_V^{-1}$
which associates to each vector of V the vector of H with the same components in the orthonormal bases, then :
$\forall X \in O : \Upsilon(X) = \Xi(X)$
and Υ is a bijective, linear map, which preserves the scalar product, so it is continuous and is an isometry.
v) There is a bijective linear map : $\mathcal{L}_H \in GL(H_0; H_0)$ such that : $\forall i \in I : \varepsilon_i = \mathcal{L}_H(\widetilde{\varepsilon}_i)$.
$(\widetilde{\varepsilon}_i)_{i\in I}$ is a basis of H_0 thus $\varepsilon_i = \sum_{j\in I} [\mathcal{L}_H]_i^j \widetilde{\varepsilon}_j$ where only a finite number of coefficients $[\mathcal{L}_H]_i^j$ is non null.
Let us define : $\varpi_i : H_0 \to \mathbb{R} :: \varpi_i\left(\sum_{j\in I} \psi_j \varepsilon_j\right) = \psi_i$
This map is continuous at $\psi = 0$ on H_0 :
take $\psi \in H_0, \|\psi\| \to 0$

then $\psi = \sum_{i\in I} \langle \widetilde{\varepsilon}_i, \psi\rangle_H \widetilde{\varepsilon}_i$ and $\widetilde{\psi}_j = \langle \widetilde{\varepsilon}_i, \psi\rangle_H \to 0$

so if $\|\psi\| < r$ then $\|\psi\|^2 = \sum_{j\in I} \left|\widetilde{\psi}_j\right|^2 < r^2$ and $\forall j \in I : \left|\widetilde{\psi}_j\right| < r$

$\psi_i = \sum_{j\in J} [\mathcal{L}_H]^j_i \widetilde{\psi}_j \Rightarrow |\psi_i| < \varepsilon \sum_{j\in I} \max \left|[\mathcal{L}_H]^j_i\right|$ and $\left(\left|[\mathcal{L}_H]^j_i\right|\right)_{j\in I}$ is bounded $\Rightarrow |\psi_i| \to 0$

Thus ϖ_i is continuous and belongs to the topological dual H'_0 of H_0. It can be extended as a continuous map $\overline{\varpi}_i \in H'$ according to the Hahn-Banach theorem (Maths.958). Because H is a Hilbert space, there is a vector $\phi_i \in H$ such that : $\forall \psi \in H : \overline{\varpi}_i(\psi) = \langle \phi_i, \psi\rangle_H$ so that :

$\forall X \in O : \Upsilon(X) = \Xi(X) = \sum_{i\in I} \psi_i \varepsilon_i$

$= \sum_{i\in I} \langle \phi_i, \psi\rangle_H \varepsilon_i = \sum_{i\in I} \langle \phi_i, \Xi(X)\rangle_H \varepsilon_i$

$\forall i \in I :$

$\Xi(e_i) = \varepsilon_i = \Upsilon(e_i) = \sum_{j\in I} \langle \phi_j, \varepsilon_i\rangle_H \varepsilon_j \Rightarrow \langle \phi_j, \varepsilon_i\rangle_H = \delta_{ij}$

$\Xi(\widetilde{e}_i) = \sum_{j\in I} \langle \phi_j, \Xi(\widetilde{e}_i)\rangle_H \varepsilon_j = \widetilde{\varepsilon}_i = \sum_{j\in I} \langle \phi_j, \widetilde{\varepsilon}_i\rangle_H \varepsilon_j$

vi) The map $\Upsilon : O \to \Omega$ is a linear chart of M, using two orthonormal bases : it is continuous, bijective so it is an homeomorphism, and is obviously compatible with the chart Ξ. ■

Remarks

i) Because $(\widetilde{\varepsilon}_i)_{i\in I}$ is a Hilbertian basis of the separable infinite dimensional Hilbert space H, I is a countable set which can be identified to $\mathbb{N}$. The assumption about $(e_i)_{i\in I}$ is that it is a Hamel basis, which is the most general because any vector space has one. From the proposition above we see that this basis must be of cardinality $\aleph_0$. Hamel bases of infinite dimensional normed vector spaces must be uncountable, however our assumption about V is that it is a Fréchet space, which is a metrizable but not a normed space, and this distinction matters. If V is a Banach vector space then, according to the Mazur theorem, it implies that there it has an infinite dimensional vector subspace W which has a Shauder basis : $\forall X \in W : X = \sum_{i\in I} x_i e_i$ where the sum is understood in the topological limit. Then the same reasoning as above shows that the closure of W is itself a Hilbert space. Moreover it has been proven that any separable Banach space is homeomorphic to a Hilbert space, and most of the applications will concern spaces of integrable functions (or sections of vector bundle endowed with a norm) which are separable Fréchet spaces.

One interesting fact is that we assume that the variables belong to an open subset O of V. The main concern is to allow for variables which can take values only in some bounded domain. But this assumption addresses also the case of a Banach vector space which is "hollowed out" : O can be itself a vector subspace (in an infinite dimensional vector space a vector subspace can be open), for instance generated by a countable subbasis of a Hamel basis, and we assume explicitly that the basis $(e_i)_{i\in I}$ belongs to O.

ii) For $O = V$ we have a largest open Ω_V and a linear map $\Upsilon : V \to \Omega_V$ with domain V.

iii) To each (Hamel) basis on V is associated a linear chart Υ of the manifold, such that a point of M has the same coordinates both in V and H. So Υ depends on the choice of the basis, and similarly the positive kernel K_V depends on the basis.

iv) In the proof we have introduced a map : $K_V : O \times O \to \mathbb{R} :: K_V(X,Y)$ which is not bilinear, but is definite positive in a precise way. It plays an important role in several following demonstrations. From a physical point of view it can be seen as related to the probability of transition between two states X, Y often used in QM.[1]

[1] We will see that this positive kernel plays an important role in the proofs of other theorems. The transitions maps are a key characteristics of the structure of a manifold, and it seems that the existence of a positive kernel is a characteristic of Fréchet manifold. This is a point to be checked by mathematicians.

2.1.5 Complex structure

The variables X and vector space V are real and H is a real Hilbert space. The condition that the vector space V is real is required only in Theorem 2 to prove the existence of a Hilbert space, because the Henderson's theorem holds only for real structures. However, as it is easily checked, if H exists, all the following theorems hold even if H is a complex Hilbert space. This is specially useful when the space V over which the maps X are defined is itself a complex Hilbert space, as this is often the case.

Moreover it can be useful to endow H with the structure of a complex Hilbert space : the set does not change but one distinguishes real and imaginary components, and the scalar product is given by a Hermitian form. Notice that this is a convenience, not a necessity.

Theorem 4 *Any real separable infinite dimensional Hilbert space can be endowed with the structure of a complex separable Hilbert space*

Proof. H has a infinite countable Hilbertian basis $(\varepsilon_\alpha)_{\alpha\in\mathbb{N}}$ because it is separable.

A complex structure is defined by a linear map : $J \in \mathcal{L}(H;H)$ such that $J^2 = -Id$. Then the operation : $i \times \psi$ is defined by : $i\psi = J(\psi)$.

Define :

$J(\varepsilon_{2\alpha}) = \varepsilon_{2\alpha+1}; J(\varepsilon_{2\alpha+1}) = -\varepsilon_{2\alpha}$

$\forall\psi \in H : i\psi = J(\psi)$

So : $i(\varepsilon_{2\alpha}) = \varepsilon_{2\alpha+1}; i(\varepsilon_{2\alpha+1}) = -\varepsilon_{2\alpha}$

The bases $\varepsilon_{2\alpha}$ or $\varepsilon_{2\alpha+1}$ are complex bases of H :

$\psi = \sum_\alpha \psi^{2\alpha}\varepsilon_{2\alpha} + \psi^{2\alpha+1}\varepsilon_{2\alpha+1} = \sum_\alpha \left(\psi^{2\alpha} - i\psi^{2\alpha+1}\right)\varepsilon_{2\alpha}$

$= \sum_\alpha \left(-i\psi^{2\alpha} + \psi^{2\alpha+1}\right)\varepsilon_{2\alpha+1}$

$\|\psi\|^2 = \sum_\alpha \left|\psi^{2\alpha} - i\psi^{2\alpha+1}\right|^2$

$= \sum_\alpha \left|\psi^{2\alpha}\right|^2 + \left|\psi^{2\alpha+1}\right|^2 + i\left(-\overline{\psi}^{2\alpha}\psi^{2\alpha+1} + \psi^{2\alpha}\overline{\psi}^{2\alpha+1}\right)$

$= \sum_\alpha \left|\psi^{2\alpha}\right|^2 + \left|\psi^{2\alpha+1}\right|^2 + i\left(-\psi^{2\alpha}\psi^{2\alpha+1} + \psi^{2\alpha}\psi^{2\alpha+1}\right)$

Thus $\varepsilon_{2\alpha}$ is a Hilbertian complex basis

H has a structure of complex vector space that we denote $H_\mathbb{C}$

The map : $T : H \to H_\mathbb{C} : T(\psi) = \sum_\alpha \left(\psi^{2\alpha} - i\psi^{2\alpha+1}\right)\varepsilon_{2\alpha}$ is linear and continuous

The map : $\overline{T} : H \to H_\mathbb{C} : \overline{T}(\psi) = \sum_\alpha \left(\psi^{2\alpha} + i\psi^{2\alpha+1}\right)\varepsilon_{2\alpha}$ is antilinear and continuous

Define : $\gamma(\psi,\psi') = \left\langle \overline{T}(\psi), T(\psi')\right\rangle_H$

γ is sesquilinear

$\gamma(\psi,\psi') = \left\langle \sum_\alpha \left(\psi^{2\alpha} + i\psi^{2\alpha+1}\right)\varepsilon_{2\alpha}, \sum_\alpha \left(\psi'^{2\alpha} - i\psi'^{2\alpha+1}\right)\varepsilon_{2\alpha}\right\rangle_H$

$= \sum_\alpha \left(\psi^{2\alpha} + i\psi^{2\alpha+1}\right)\left(\psi'^{2\alpha} - i\psi'^{2\alpha+1}\right)$

$= \sum_\alpha \psi^{2\alpha}\psi'^{2\alpha} + \psi^{2\alpha+1}\psi'^{2\alpha+1} + i\left(\psi^{2\alpha+1}\psi'^{2\alpha} - \psi^{2\alpha}\psi'^{2\alpha+1}\right)$

$\gamma(\psi,\psi) = 0 \Rightarrow \langle\psi,\psi\rangle_H = 0 \Rightarrow \psi = 0$

Thus γ is definite positive ■

2.1.6 Decomposition of the Hilbert space

V is the product $V = V_1 \times V_2 ... \times V_N$ of vector spaces, thus the proposition implies that the Hilbert space H is also the direct product of Hilbert spaces $H_1 \times H_2 ... \times H_N$ or equivalently $H = \oplus_{k=1}^N H_k$ where H_k are Hilbert vector subspaces of H. More precisely :

Theorem 5 *If the model is comprised of N continuous variables* $(X_k)_{k=1}^N$ *, each belonging to a separable Fréchet vector space* V_k*, then the real Hilbert space* H *of states of the system is the Hilbert sum of N Hilbert space* $H = \oplus_{k=1}^N H_k$ *and any vector* ψ *representing a state of the system is uniquely the sum of N vectors* ψ_k*, each image of the value of one variable* X_k *in the state* ψ

Proof. By definition $V = \prod_{k=1}^N V_k$ .The set $V_k^0 = \{0, .., V_k, ...0\} \subset V$ is a vector subspace of V. A basis of V_k^0 is a subfamily $(e_i)_{i \in I_k}$ of a basis $(e_i)_{i \in I}$ of V. V_k^0 has for image by the continuous linear map Υ a closed vector subspace H_k of H. Any vector X of V reads : $X \in \prod_{k=1}^N V_k : X = \sum_{k=1}^N \sum_{i \in I_k} x^i e_i$ and it has for image by $\Upsilon : \psi = \Upsilon(X) = \sum_{k=1}^N \sum_{i \in I_k} x^i \varepsilon_i = \sum_{k=1}^N \psi_k$ with $\psi_k \in H_k$ .This decomposition of $\Upsilon(X)$ is unique.

Conversely, the family $(e_i)_{i \in I_k}$ has for image by Υ the set $(\varepsilon_i)_{i \in I_k}$ which are linearly independent vectors of H_k.It is always possible to build an orthonormal basis $(\widetilde{\varepsilon}_i)_{i \in I_k}$ from these vectors as done previously. H_k is a closed subspace of H, so it is a Hilbert space. The map : $\widehat{\pi}_k : \ell^2(I_k) \to H_k :: \widehat{\pi}_k(x) = \sum_{i \in I_k} x^i \widetilde{\varepsilon}_i$ is an isomorphism of Hilbert spaces and :$\forall \psi \in H_k : \psi = \sum_{i \in I_k} \langle \widetilde{\varepsilon}_i, \psi \rangle_H \widetilde{\varepsilon}_i$.

$\forall \psi_k \in H_k, \psi_l \in H_l, k \neq l : \langle \psi_k, \psi_l \rangle_H = \left\langle \Upsilon^{-1}(\psi_k), \Upsilon^{-1}(\psi_l) \right\rangle_E = 0$

Any vector $\psi \in H$ reads : $\psi = \sum_{k=1}^N \pi_k(\psi)$ with the orthogonal projection $\pi_k : H \to H_k :: \pi_k(\psi) = \sum_{i \in I_k} \langle \widetilde{\varepsilon}_i, \psi \rangle_H \widetilde{\varepsilon}_i$ so H is the Hilbert sum of the H_k ■

As a consequence the definite positive kernel of (V, Υ) decomposes as :

$K((X_1, ...X_N), (X_1', ...X_N'))$

$= \sum_{k=1}^N K_k(X_k, X_k')$

$= \sum_{k=1}^N \langle \Upsilon(X_k), \Upsilon(X_k') \rangle_{H_k}$

This decomposition comes handy when we have to translate relations between variables into relations between vector states, notably it they are linear. But it requires that we keep the real Hilbert space structure.

2.1.7 Discrete variables

It is common in a model to have discrete variables $(D_k)_{k=1}^K$, taking values in a finite discrete set. They correspond to different cases:

i) the discrete variables identify different elementary systems (such as different populations of particles) which coexist simultaneously in the same global system, follow different rules of behavior, but interact together. We will see later how to deal with these cases (tensorial product).

ii) the discrete variables identify different populations, whose interactions are not relevant. Actually one could consider as many different systems but, by putting them together, one increases the size of the samples of data and improve the statistical estimations. They are not of great interest here, in a study of formal models.

iii) the discrete variables represent different kinds of behaviors, which cannot be strictly identified with specific populations. Usually a discrete variable is then used as a proxy for a quantitative parameter which tells how close the system is from a specific situation.

We will focus on this third case. The system is represented as before by quantitative variables X, whose possible values belong to some set M, which has the structure of an infinite dimensional manifold. The general idea in the third case is that the possible states of the system can be regrouped in two distinct subsets. That we formalize in the following assumption : the set O of possible states of the system has two connected components O_1, O_2

Theorem 6 *If the condition of the theorem 2 are met, and the set O of possible states of the system has two connected components O_1, O_2 then there is a continuous function $f : H \to [0,1]$ such that $f(\Upsilon(X)) = 1$ in O_1 and $f(\Upsilon(X)) = 0$ in O_2*

Proof. The connected components O_1, O_2 of a topological space are closed, so O_1, O_2 are disjoint and both open and closed in V (Maths.624). Using a linear continuous map Υ then Ω has itself two connected components, $\Omega_1 = \Upsilon^{-1}(O_1), \Omega_2 = \Upsilon^{-1}(O_2)$ both open and closed, and disjoint. H is metric, so it is normal (Maths.705). Ω_1, Ω_2 are disjoint and closed in H. Then, by the Urysohn's Theorem (Maths.596) there is a continuous function f on H valued in $[0,1]$ such that $f(\psi) = 1$ in H_1 and $f(\psi) = 0$ in H_2. ■

The set of continuous, bounded functions is a Banach vector space, so it is always possible, in these conditions, to replace a discrete variable by a quantitative variable with the same features.

2.2 OBSERVABLES

The key point in the conditions 1 above is that the variables are maps, which take an infinite number of values (usually non countable). So the variables would require the same number of data to be totally known, which is impossible. The physicist estimates the variable by statistical methods. But any practical method involves a first step : the scope of all maps is reduced from V to a smaller subset W, so that any map of W can be characterized by a finite number of parameters. The procedure sums up to replace X by another variable $\Phi(X)$ that we will call an observable, which is then estimated from a finite batch of data. The mechanism of estimating the variables $X \subset V$ is then the following :

- the observer collects data, as a set $Y = \{x_p\}_{p=1}^N$ of values assumed to be taken by the variable X, in the mathematical format fitted to X (scalars, vectors,..) for different values of the arguments

- he proceeds to the estimation $\widehat{X}$ of the map $\Phi(X)$ by statistical adjustment to the data $\{x_p\}_{p=1}^N$. Because there are a finite number of parameters (the coordinates of $\Phi(X)$ in W) this is possible

- the estimation is : $\widehat{X} = \varphi(Y) \in W$: this is a map which is a simplified version of X.

The procedure of the replacement of X by $\Phi(X)$, called the choice of a **specification**, is done by the physicist, and an observable is not unique. However we make three general assumptions about Φ :

Definition 7 *i) an* ***observable*** *is a linear map :* $\Phi \in L(V;V)$

ii) the range of an observable is a finite dimensional vector subspace W *of* V *:* $W \subset V, \dim \Phi(W) < \infty$

iii) $\forall X \in O, \Phi(X)$ *is an admissible value, that is* $\Phi(O) \subset O$.

Using the linear chart Υ given by any basis, to Φ one can associate a map :

$$\widehat{\Phi} : H \to H :: \widehat{\Phi} = \Upsilon \circ \Phi \circ \Upsilon^{-1} \tag{2.1}$$

and $\widehat{\Phi}$ is an operator on H. And conversely.

The image of W by Υ is a finite dimensional vector subspace $H_\Phi = \Upsilon(W)$ of H, so it is closed and a Hilbert space : $\widehat{\Phi} \in \mathcal{L}(H; H_\Phi)$

$$\begin{array}{ccccccc} & & & \Phi & & & \\ & V & \to & \to & \to & W & \\ & \downarrow & & & & \downarrow & \\ \Upsilon & \downarrow & & & & \downarrow & \Upsilon \\ & \downarrow & & \widehat{\Phi} & & \downarrow & \\ & H & \to & \to & \to & H_\Phi & \end{array}$$

2.2.1 Primary observables

The simplest specification for an observable is, given a basis $(e_i)_{i\in I}$, to define Φ as the projection on the subspace spanned by a finite number of vectors of the basis. For instance if X is a function $X(t)$ belonging to some space such as : $X(t) = \sum_{n\in\mathbb{N}} a_n e_n(t)$ where $e_n(t)$ are fixed functions, then a primary observable would be $Y_J(X(t)) = \sum_{n=0}^{N} a_n e_n(t)$ meaning that the components $(a_n)_{n>N}$ are discarded and the data are used to compute $(a_n)_{n=0}^{N}$. To stay at the most general level, we define :

Definition 8 *A **primary observable** $\Phi = Y_J$ is the projection of $X = \{X_k, k = 1...N\}$ on the vector subspace V_J spanned by the vectors $(e_i)_{i\in J} \equiv \left(e_i^k\right)_{i\in J_k}$ where $J = \prod_{k=1}^{N} J_k \subset I = \prod_{k=1}^{N} I_k$ is a finite subset of I and $(\varepsilon_i)_{i\in I} = \prod_{k=1}^{N} \left(e_i^k\right)_{i\in I_k}$ is a basis of V.*

So the procedure can involve simultaneously several variables. It requires *the choice of a finite set of independent vectors of V.*

Theorem 9 *To any primary observable Y_J is associated uniquely a self-adjoint, compact, trace-class operator $\widehat{Y}_J$ on H : $Y_J = \Upsilon^{-1} \circ \widehat{Y}_J \circ \Upsilon$ such that the measure $Y_J(X)$ of the primary observable Y_J, if the system is in the state $X \in O$, is*

$Y_J(X) = \sum_{i\in I} \left\langle \phi_i, \widehat{Y}_J(\Upsilon(X)) \right\rangle_H e_i$

Proof. i) We use the notations and definitions of the previous section. The family of variables $X = (X_k)_{k=1}^{N}$ define the charts : $\Xi : O \to \Omega$ and the basis $(e_i)_{i\in I}$ defines the bijection $\Upsilon : V \to H$

$\forall X = \sum_{i\in I} x_i e_i \in O$:

$\Upsilon(X) = \sum_{i\in I} x_i \Upsilon(e_i) = \sum_{i\in I} x_i \varepsilon_i = \sum_{i\in I} \langle \phi_i, \Upsilon(X)\rangle_H \varepsilon_i$

$\Leftrightarrow x_i = \langle \phi_i, \Upsilon(X)\rangle_H$

$\forall i, j \in I : \langle \phi_i, \varepsilon_j\rangle_H = \delta_{ij}$

ii) The primary observable Y_J is the map :

$Y_J : V \to V_J :: Y_J(X) = \sum_{j\in J} x_j e_j$

This is a projection : $Y_J^2 = Y_J$

$Y_J(X) \in O$ so it is associated to a vector of H :

$\Upsilon(Y_J(X)) = \Upsilon\left(\sum_{j\in J} x_j e_j\right) = \sum_{j\in J} \langle \phi_j, \Upsilon(Y_J(X))\rangle_H \varepsilon_j$

$= \sum_{j\in J} \langle \phi_j, \Upsilon(X)\rangle_H \varepsilon_j$

iii) $\forall X \in O : \Upsilon(Y_J(X)) \in H_J$ where H_J is the vector subspace of H spanned by $(\varepsilon_j)_{j\in J}$. It is finite dimensional, thus it is closed in H. There is a unique map (Math.1111) :

$\widehat{Y}_J \in \mathcal{L}(H; H) :: \widehat{Y}_J^2 = \widehat{Y}_J, \widehat{Y}_J = \widehat{Y}_J^*$

$\widehat{Y}_J$ is the orthogonal projection from H onto H_J. It is linear, self-adjoint, and compact because its range is a finite dimensional vector subspace. As a projection : $\left\|\widehat{Y}_J\right\| = 1$.

$\widehat{Y}_J$ is a Hilbert-Schmidt operator (Maths.1143) : take the Hilbertian basis $\widetilde{\varepsilon}_i$ in H:

$\sum_{i\in I} \left\|\widehat{Y}_J(\widetilde{\varepsilon}_i)\right\|^2 = \sum_{ij\in J} |\langle \phi_j, \widetilde{\varepsilon}_i\rangle|^2 \|\varepsilon_j\|^2 = \sum_{j\in J} \|\phi_j\|^2 \|\varepsilon_j\|^2 < \infty$

$\widehat{Y}_J$ is a trace class operator (Maths.1147) with trace $\dim H_J$

$\sum_{i\in I} \left\langle \widehat{Y}_J(\widetilde{\varepsilon}_i), \widetilde{\varepsilon}_i \right\rangle = \sum_{ij\in J} \langle \phi_j, \widetilde{\varepsilon}_i\rangle \langle \varepsilon_j, \widetilde{\varepsilon}_i\rangle$

$= \sum_{j\in J} \langle \phi_j, \varepsilon_j\rangle = \sum_{j\in J} \delta_{jj} = \dim H_J$

iv) $\forall \psi \in H_J : \widehat{Y}_J(\psi) = \psi$

$\forall X \in O : \Upsilon(Y_J(X)) \in H_J$

$\forall X \in O : \Upsilon(Y_J(X)) = \widehat{Y}_J(\Upsilon(X)) \Leftrightarrow Y_J(X) = \Upsilon^{-1} \circ \widehat{Y}_J(\Upsilon(X)) \Leftrightarrow Y_J = \Upsilon^{-1} \circ \widehat{Y}_J \circ \Upsilon$

v) The value of the observable reads : $Y_J(X) = \sum_{i\in I} \left\langle \phi_i, \widehat{Y}_J(\Upsilon(X)) \right\rangle_H e_i$ ■

2.2.2 von Neumann algebras

There is a bijective correspondence between the projections, meaning the maps $P \in \mathcal{L}(H;H)$: $P^2 = P, P = P^*$ and the closed vector subspaces of H (Maths.1111). Then P is the orthogonal projection on the vector subspace. So the operators $\widehat{Y}_J$ for any finite subset J of I are the orthogonal projections on the finite dimensional, and thus closed, vector subspace H_J spanned by $(\varepsilon_j)_{j\in J}$.

We will enlarge the family of primary observables in several steps, in keeping the same basis $(e_i)_{i\in I}$ of V.

1. For any given basis $(e_i)_{i\in I}$ of V, we extend the definition of these operators $\widehat{Y}_J$ to any finite or infinite, subset of I by taking $\widehat{Y}_J$ as the orthogonal projection on the closure $\overline{H_J}$ in H of the vector subspace H_J spanned by $(\varepsilon_j)_{j\in J}$: $\overline{H_J} = \overline{Span\,(\varepsilon_j)_{j\in J}}$.

Theorem 10 *The operators* $\left\{\widehat{Y}_J\right\}_{J\subset I}$ *are self-adjoint and commute*

Proof. Because they are projections the operators $\widehat{Y}_J$ are such that : $\widehat{Y}_J^2 = \widehat{Y}_J, \widehat{Y}_J^* = \widehat{Y}_J$

$\widehat{Y}_J$ has for eigen values :

1 for $\psi \in \overline{H_J}$

0 for $\psi \in \left(\overline{H_J}\right)^\perp$

For any subset J of I, by the Gram-Schmidt procedure one can built an orthonormal basis $(\widetilde{\varepsilon}_i)_{i\in J}$ of H_J starting with the vectors $(\varepsilon_i)_{i\in J}$ and an orthonormal basis $(\widetilde{\varepsilon}_i)_{i\in J^c}$ of H_{J^c} starting with the vectors $(\varepsilon_i)_{i\in J^c}$

Any vector $\psi \in H$ can be written :

$\psi = \sum_{j\in I} x_j\widetilde{\varepsilon}_j = \sum_{j\in J} x_j\widetilde{\varepsilon}_j + \sum_{j\in J^c} x_j\widetilde{\varepsilon}_j$ with $(x_j)_{j\in I} \in \ell^2(I)$

$\overline{H_J}$ is defined as $\sum_{j\in J} x_j\widetilde{\varepsilon}_j$ with $(x_j)_{j\in J} \in \ell^2(J)$ and similarly $\overline{H_{J^c}}$ is defined as $\sum_{j\in J^c} x_j\widetilde{\varepsilon}_j$ with $(x_j)_{j\in J^c} \in \ell^2(J^c)$

So $\widehat{Y}_J$ can be defined as : $\widehat{Y}_J\left(\sum_{j\in I} x_j\widetilde{\varepsilon}_j\right) = \sum_{j\in J} x_j\widetilde{\varepsilon}_j$

For any subsets $J_1, J_2 \subset I$:

$\widehat{Y}_{J_1} \circ \widehat{Y}_{J_2} = \widehat{Y}_{J_1\cap J_2} = \widehat{Y}_{J_2} \circ \widehat{Y}_{J_1}$

$\widehat{Y}_{J_1\cup J_2} = \widehat{Y}_{J_1} + \widehat{Y}_{J_2} - \widehat{Y}_{J_1\cap J_2} = \widehat{Y}_{J_1} + \widehat{Y}_{J_2} - \widehat{Y}_{J_1} \circ \widehat{Y}_{J_2}$

So the operators commute. ■

2. Let us define $W = Span\left\{\widehat{Y}_i\right\}_{i\in I}$ the vector subspace of $\mathcal{L}(H;H)$ comprised of finite linear combinations of $\widehat{Y}_i$ (as defined in 1 above). The elements $\left\{\widehat{Y}_i\right\}_{i\in I}$ are linearly independent and constitute a basis of W.

The operators $\widehat{Y}_j, \widehat{Y}_k$ are mutually orthogonal for $j \neq k$:

$\widehat{Y}_j \circ \widehat{Y}_k(\psi) = \langle\phi_k, \psi\rangle \langle\phi_j, \varepsilon_k\rangle \varepsilon_j = \langle\phi_k, \psi\rangle \delta_{jk} = \delta_{jk}\widehat{Y}_j(\psi)$

Let us define the scalar product on W :

$\left\langle\sum_{i\in I} a_i\widehat{Y}_i, \sum_{i\in I} b_i\widehat{Y}_i\right\rangle_W = \sum_{i\in I} a_ib_i$

$\left\|\sum_{i\in I} a_i\widehat{Y}_i\right\|_W^2 = \sum_{i\in I} a_i^2 \left\|\widehat{Y}_i\right\|_W^2 = \sum_{i\in I} a_i^2$

W is isomorphic to $\mathbb{R}_0^I$ and its closure in $\mathcal{L}(H;H)$: $\overline{W} = \overline{Span\left\{\widehat{Y}_i\right\}_{i\in I}}$ is isomorphic to $\ell^2(I)$, and has the structure of a Hilbert space with :

$\overline{W} = \left\{ \sum_{i \in I} a_i \widehat{Y}_i, (a_i)_{i \in I} \in \ell^2 (I) \right\}$

3. Let us define A as the algebra generated by any finite linear combination or products of elements $\widehat{Y}_J, J$ finite or infinite, and $\overline{A}$ as the closure of A in $\mathcal{L}(H;H)$: $\overline{A} = \overline{Span \left\{ \widehat{Y}_J \right\}_{J \subset I}}$ with respect to the strong topology, that is in norm.

Theorem 11 *$\overline{A}$ is a commutative von Neumann algebra of $\mathcal{L}(H,H)$*

Proof. It is obvious that A is a *subalgebra of $\mathcal{L}(H,H)$ with unit element $Id = \widehat{Y}_I$.

Because its generators are projections, $\overline{A}$ is a von Neumann algebra (Maths.1190).

The elements of $A = Span \left\{ \widehat{Y}_J \right\}_{J \subset I}$ that is of finite linear combination of $\widehat{Y}_J$ commute

$Y, Z \in \overline{A} \Rightarrow \exists (Y_n)_{n \in \mathbb{N}}, (Z_n)_{n \in \mathbb{N}} \in A^{\mathbb{N}} : Y_n \to_{n \to \infty} Y, Z_n \to_{n \to \infty} Z$

The composition is a continuous operation.

$Y_n \circ Z_n = Z_n \circ Y_n \Rightarrow \lim (Y_n \circ Z_n) = \lim (Z_n \circ Y_n) = \lim Y_n \circ \lim Z_n = \lim Z_n \circ \lim Y_n = Z \circ Y = Y \circ Z$

So $\overline{A}$ is commutative.

$\overline{A}$ is identical to the bicommutant of its projections, that is to $\overline{A}$" (Maths.1189) ∎

This result is of interest because commutative von Neumann algebras are classified : they are isomorphic to the space of functions $f \in L^\infty (E, \mu)$ acting by pointwise multiplication $\varphi \to f\varphi$ on functions $\varphi \in L^2 (E, \mu)$ for some set E and measure μ (not necessarily absolutely continuous). They are the topic of many studies, notably in ergodic theory. The algebra $\overline{A}$ depends on the choice of a basis $(e_i)_{i \in I}$ and, as can be seen in the formulation through $(\widetilde{\varepsilon}_i)_{i \in I}$, is defined up to a unitary transformation.

Taking the axioms of QM as a starting point, one can define a system itself by the set of its observables : this is the main idea of the Axiomatic QM Theories. This is convenient to explore further the behavior of systems or some sensitive issues such as the continuity of the operators. But this approach has a fundamental drawback : it leads further from an understanding of the physical foundations of the theory itself. To tell that a system should be represented by a von Neumann algebra does not explain more why a state should be represented in a Hilbert space at the beginning.

We see here how such an algebra appears naturally. However the algebra $\overline{A}$ is commutative, and this property is the consequence of the choice of a unique basis $(e_i)_{i \in I}$. It would not hold for primary observables defined through different bases : they do not even constitute an algebra. Any von Neumann algebra is the closure of the linear span of its projections (Maths.1190), and any projection can be defined through a basis, thus one can say that the "observables" (with their usual definition) of a system are the collection of all primary observables (as defined here) for all bases of V.

2.2.3 Secondary observables

Beyond primary observables, general observables Φ can be studied using spectral theory.

1. A spectral measure defined on a measurable space E with $\sigma-$algebra σ_E and acting on the Hilbert space H is a map : $P : \sigma_E \to \mathcal{L}(H;H)$ such that (Maths.1240) :

i) $P(\varpi)$ is a projection

ii) $P(E) = Id$

iii) $\forall \psi \in H$ the map: $\varpi \rightarrow \langle P(\varpi)\psi, \psi\rangle_H = \|P(\varpi)\psi\|^2$ is a finite positive measure on (E, σ_E).

One can show (Maths.1242) that there is a bijective correspondence between the spectral measures on H and the maps : $\chi : \sigma_E \rightarrow H$ such that :

i) $\chi(\varpi)$ is a closed vector subspace of H

ii) $\chi(E) = H$

iii) $\forall \varpi, \varpi' \in \sigma_E, \varpi \cap \varpi' = \varnothing : \chi(\varpi) \cap \chi(\varpi') = \{0\}$

then $P(\varpi)$ is the orthogonal projection on $\chi(\varpi)$, denoted : $\widehat{\pi}_{\chi(\varpi)}$

Thus, for any fixed $\psi \neq 0 \in H$ the function $\widehat{\chi}_\psi : \sigma_E \rightarrow \mathbb{R} :: \widehat{\chi}_\psi(\varpi) = \frac{\langle \widehat{\pi}_{\chi(\varpi)}\psi, \psi\rangle}{\|\psi\|^2} = \frac{\|\widehat{\pi}_{\chi(\varpi)}\psi\|^2}{\|\psi\|^2}$ is a probability law on (E, σ_E).

2. An application of standard theorems on spectral measures (Maths.1243, 1245) tells that, for any bounded measurable function $f : E \rightarrow \mathbb{R}$, the spectral integral : $\int_E f(\xi)\widehat{\pi}_{\chi(\xi)}$ defines a continuous operator $\widehat{\Phi}_f$ on H. $\widehat{\Phi}_f$ is such that :

$\forall \psi, \psi' \in H : \left\langle \widehat{\Phi}_f(\psi), \psi' \right\rangle = \int_E f(\xi) \left\langle \widehat{\pi}_{\chi(\xi)}(\psi), \psi' \right\rangle$

And conversely (Math.1252), for any continuous normal operator $\widehat{\Phi}$ on H, that is such that :

$\widehat{\Phi} \in \mathcal{L}(H; H) : \widehat{\Phi} \circ \widehat{\Phi}^* = \widehat{\Phi}^* \circ \widehat{\Phi}$ with the adjoint $\widehat{\Phi}^*$

there is a unique spectral measure P on $(\mathbb{R}, \sigma_\mathbb{R})$ such that $\widehat{\Phi} = \int_{Sp(\widehat{\Phi})} sP(s)$ where $Sp(\widehat{\Phi}) \subset \mathbb{R}$ is the spectrum of $\widehat{\Phi}$.

So there is a map $\chi : \sigma_\mathbb{R} \rightarrow H$ where $\sigma_\mathbb{R}$ is the Borel algebra of $\mathbb{R}$ such that :

$\chi(\varpi)$ is a closed vector subspace of H

$\chi(\mathbb{R}) = Id$

$\forall \varpi, \varpi' \in \sigma_\mathbb{R}, \varpi \cap \varpi' = \varnothing \Rightarrow \chi(\varpi) \cap \chi(\varpi') = \{0\}$

and $\widehat{\Phi} = \int_{Sp(\widehat{\Phi})} s\widehat{\pi}_{\chi(s)}$

The spectrum $Sp(\widehat{\Phi})$ is a non empty compact subset of $\mathbb{R}$. If $\widehat{\Phi}$ is normal then $\lambda \in Sp(\widehat{\Phi}) \Leftrightarrow \overline{\lambda} \in Sp(\widehat{\Phi}^*)$.

For any fixed $\psi \neq 0 \in H$ the function $\widehat{\mu}_\psi : \sigma_\mathbb{R} \rightarrow \mathbb{R} :: \widehat{\mu}_\psi(\varpi) = \frac{\langle \widehat{\pi}_{\chi(\varpi)}\psi, \psi\rangle}{\|\psi\|^2} = \frac{\|\widehat{\pi}_{\chi(\varpi)}\psi\|^2}{\|\psi\|^2}$ is a probability law on $(\mathbb{R}, \sigma_\mathbb{R})$.

3. We will define :

Definition 12 *A* ***secondary observable*** *is a linear map $\Phi \in L(V; V)$ valued in a finite dimensional vector subspace of V, such that $\widehat{\Phi} = \Upsilon \circ \Phi \circ \Upsilon^{-1}$ is a normal operator : $\widehat{\Phi} \circ \widehat{\Phi}^* = \widehat{\Phi}^* \circ \widehat{\Phi}$ with the adjoint $\widehat{\Phi}^*$*

Theorem 13 *Any secondary observable Φ is a compact, continuous map, its associated map $\widehat{\Phi} = \Upsilon \circ \Phi \circ \Upsilon^{-1}$ is a compact, self-adjoint, Hilbert-Schmidt and trace class operator.*

$\Phi = \sum_{p=1}^n \lambda_p Y_{J_p}$ where $(Y_{J_p})_{p=1}^N$ are primary observables associated to a basis $(e_i)_{i \in I}$ of V and $(J_p)_{p=1}^n$ are disjoint finite subsets of I

Proof. i) $\widehat{\Phi}(H)$ is a finite dimensional vector subspace of H. So :

$\widehat{\Phi}$ has 0 for eigen value, with an infinite dimensional eigen space H_c.

$\Phi, \widehat{\Phi}$ are compact and thus continuous (Maths.912).

ii) As $\widehat{\Phi}$ is continuous and normal, there is a unique spectral measure P on $(\mathbb{R}, \sigma_{\mathbb{R}})$ such that $\widehat{\Phi} = \int_{Sp(\widehat{\Phi})} sP(s)$ where $Sp(\widehat{\Phi}) \subset \mathbb{R}$ is the spectrum of $\widehat{\Phi}$. As $\widehat{\Phi}$ is compact, by the Riesz theorem (Maths.1142) its spectrum is either finite or is a countable sequence converging to 0 (which may or not be an eigen value) and, except possibly for 0, is identical to the set $(\lambda_p)_{p\in\mathbb{N}}$ of its eigen values (Maths.1020). For each distinct eigen value the eigen spaces H_p are orthogonal and H is the direct sum $H = \oplus_{p\in\mathbb{N}} H_p$. For each non null eigen value λ_p the eigen space H_p is finite dimensional.

Let λ_0 be the eigen value 0 of $\widehat{\Phi}$. So : $\widehat{\Phi} = \sum_{p\in\mathbb{N}} \lambda_p \widehat{\pi}_{H_p}$and any vector of H reads : $\psi = \sum_{p\in\mathbb{N}} \psi_p$ with $\psi_p = \widehat{\pi}_{H_p}(\psi)$

Because $\widehat{\Phi}(H)$ is finite dimensional, the spectrum is finite and the non null eigen values are $(\lambda_p)_{p=1}^n$, the eigen space corresponding to 0 is $H_c = \left(\oplus_{p=1}^n H_p\right)^{\perp}$

$\forall \psi \in H : \psi = \psi_c + \sum_{p=1}^n \psi_p$ with $\psi_p = \widehat{\pi}_{H_p}(\psi), \psi_c = \widehat{\pi}_{H_c}(\psi)$

$\widehat{\Phi} = \sum_{p=1}^n \lambda_p \widehat{\pi}_{H_p}$

Its adjoint reads : $\widehat{\Phi}^* = \sum_{p\in\mathbb{N}} \overline{\lambda}_p \widehat{\pi}_{H_p} = \sum_{p\in\mathbb{N}} \lambda_p \widehat{\pi}_{H_p}$ because H is a real Hilbert space

$\widehat{\Phi}$ is then self-adjoint, Hilbert-Schmidt and trace class, as the sum of the trace class operators $\widehat{\pi}_{H_p}$.

iii) The observable reads :

$\Phi = \sum_{p=1}^n \lambda_p \pi_p$ where $\pi_p = \Upsilon^{-1} \circ \widehat{\pi}_{H_p} \circ \Upsilon$ is the projection on a finite dimensional vector subspace of V :

$\pi_p \circ \pi_q = \Upsilon^{-1} \circ \widehat{\pi}_{H_p} \circ \Upsilon \circ \Upsilon^{-1} \circ \widehat{\pi}_{H_q} \circ \Upsilon = \Upsilon^{-1} \circ \widehat{\pi}_{H_p} \circ \widehat{\pi}_{H_q} \circ \Upsilon = \delta_{pq} \Upsilon^{-1} \circ \widehat{\pi}_{H_p} \circ \Upsilon = \delta_{pq} \pi_p$

$\Phi \circ \pi_p = \lambda_p \pi_p$ so $\pi_p(V) = V_p$ is the eigen space of Φ for the eigen value λ_p and the subspaces $(V_p)_{p=1}^n$ are linearly independent.

By choosing any basis $(e_i)_{i\in J_p}$ of V_p, and $(e_i)_{i\in J^c}$ with $J^c = \complement_I \left(\oplus_{p=1}^n J_n\right)$ for the basis of $V_c = Span\left((e_i)_{i\in J^c}\right)$

$X = Y_{J^c}(X) + \sum_{p=1}^n Y_{J_p}(X)$

the observable Φ reads : $\Phi = \sum_{p=1}^n \lambda_p Y_{J_p}$ ■

We have :

$Y_{J_p}(X) = \sum_{i\in J_p} \left\langle \phi_i, \widehat{Y}_{Jp}(\Upsilon(X)) \right\rangle_H e_i$

$\Phi(X) = \sum_{p=1}^n \lambda_p \sum_{i\in J_p} \left\langle \phi_i, \widehat{Y}_{Jp}(\Upsilon(X)) \right\rangle_H e_i$

$= \sum_{i\in I} \left\langle \phi_i, \sum_{p=1}^n \lambda_p \widehat{Y}_{Jp}(\Upsilon(X)) \right\rangle_H e_i$

$= \sum_{i\in I} \left\langle \phi_i, \widehat{\Phi}(\Upsilon(X)) \right\rangle_H e_i$

$\Phi, \widehat{\Phi}$ have invariant vector spaces, which correspond to the direct sum of the eigen spaces.

The probability law $\widehat{\mu}_\psi : \sigma_{\mathbb{R}} \to \mathbb{R}$ reads :

$\widehat{\mu}_\psi(\varpi) = \Pr(\lambda_p \in \varpi) = \frac{\left\|\widehat{\pi}_{H_p}(\psi)\right\|^2}{\|\psi\|^2}$

To sum up :

Theorem 14 *For any primary or secondary observable Φ, there is a basis $(e_i)_{i\in I}$ of V, a compact, self-adjoint, Hilbert-Schmidt and trace class operator $\widehat{\Phi}$ on the associated Hilbert space H such that :*

$\widehat{\Phi} = \Upsilon \circ \Phi \circ \Upsilon^{-1}$

if the system is in the state $X = \sum_{i \in I} \langle \phi_i, \Upsilon(X) \rangle_H e_i$ *the value of the observable is :* $\Phi(X) = \sum_{i \in I} \left\langle \phi_i, \widehat{\Phi}(\Upsilon(X)) \right\rangle_H e_i$

$\widehat{\Phi}$ has a finite set of eigen values, whose eigen spaces (except possibly for 0) are finite dimensional and orthogonal. The vectors corresponding to the eigen value 0 are never observed, so it is convenient to represent the Hilbert space H through a basis of eigen vectors, each of them corresponding to a definite state, which usually can be identified. This is a method commonly used in Quantum Mechanics, however the vector has also a component in the eigen space corresponding to the null eigen value, which is not observed but exists. Conversely any observable (on V) can be defined through an operator on H with the required properties (compact, normal, it is then self-adjoint). We will come back on this point in the following, when a group is involved.

2.2.4 Efficiency of an observable

A crucial factor for the quality and the cost of the estimation procedure is the number of parameters to be estimated, which is closely related to the dimension of the vector space $\Phi(V)$, which is finite. The error made by the choice of $\Phi(X)$ when the system is in the state X is : $o_\Phi(X) = X - \Phi(X)$. If two observables Φ, Φ' are such that $\Phi(V), \Phi'(V)$ have the same dimension, one can say that Φ is more efficient than Φ' if : $\forall X : \|o_\Phi(X)\|_V \leq \|o_{\Phi'}(X)\|_V$

To assess the efficiency of a secondary observable Φ it is legitimate to compare Φ to the primary observable Y_J with a set J which has the same cardinality as the dimension of $\oplus_{p=1}^n H_p$.

The error with the choice of Φ is :

$o_\Phi(X) = X - \Phi(X) = Y_c(\psi) + \sum_{p=1}^n (1 - \lambda_p) Y_p(\psi)$

$\|o_\Phi(X)\|_V^2 = \|Y_c(\psi)\|_V^2 + \sum_{p=1}^n (1 - \lambda_p)^2 \|Y_p(\psi)\|^2$

$\widehat{o}_\Phi(\Upsilon(X)) = \Upsilon(X) - \widehat{\Phi}(\Upsilon(X)) = \widehat{\pi}_{H_c}(\psi) + \sum_{p=1}^n (1 - \lambda_p) \widehat{\pi}_{H_p}(\psi)$

$\|\widehat{o}_\Phi(\Upsilon(X))\|^2 = \|\widehat{\pi}_{H_c}(\psi)\|^2 + \sum_{p=1}^n (1 - \lambda_p)^2 \left\|\widehat{\pi}_{H_p}(\psi)\right\|^2 = \|o_\Phi(X)\|_V^2$

And for Y_J : $\|\widehat{o}_{Y_J}(\Upsilon(X))\|^2 = \|\widehat{\pi}_{H_c}(\psi)\|^2$ because $\lambda_p = 1$

So :

Theorem 15 *For any secondary observable there is always a primary observable which is at least as efficient.*

This result justifies the restriction, in the usual formalism, of observables to operators belonging to a von Neumann algebra.

2.2.5 Statistical estimation and primary observables

At first the definition of a primary observable seems naive, and the previous results will seem obvious. After all the definition of a primary observable requires only the choice of a finite number of independent vectors of V. A primary observable is always better than a, more sophisticated, secondary observable. But we have also to compare a primary observable to what is practically done in an experiment, where we have to estimate a map from a batch of data.

Consider a model with variables X, maps, belonging to a Hilbert space H (to keep it simple), from a set M to a normed vector space E, endowed with a scalar product $\langle\rangle_E$. The physicist has a batch of data, that is a finite set $\{x_p \in E, p = 1...N\}$ of N measures of X done at different points $\Omega = \{m_p \in M, p = 1...N\}$: of M : $x_p = X(m_p)$. The estimated map $\widehat{X}$ should be a solution of the collection of equations : $x_p = X(m_p)$ where x_p, m_p are known.

The **evaluation maps**, that we will encounter several times, is the collection of maps $\mathcal{E}(m)$ on H :

$\mathcal{E}(m) : H \to E :: \mathcal{E}(m) Y = Y(m)$

Because H and E are vector spaces $\mathcal{E}(m)$ is a linear map : $\mathcal{E}(m) \in L(H; E)$, depending on both H and E. It can be continuous or not.

The set of solutions of the equations, that is of maps Y of H such that $\forall m_p \in \Omega : Y(m_p) = x_p$ is :

$A = \cap_{m_p \in \Omega} \mathcal{E}(m_p)^{-1}(x_p)$

$Y \in A \Leftrightarrow \forall m \in \Omega : Y(m) = X(m)$

It is not empty because it contains at least X. Its closed convex hull is the set B in H (Maths.361) :

$\forall Z \in B : \exists \alpha \in [0.1], Y, Y' \in A : Z = \alpha Y + (1-\alpha) Y'$

$\Rightarrow \forall m \in \Omega : Z(m) = x_p$

B is the smallest closed set of H such that all its elements Z are solutions of the equations : $\forall p = 1..N : Z(m_p) = x_p$.

If we specify an observable, we restrict X to a finite dimensional subspace $H_J \subset H$. With the evaluation map $\mathcal{E}_J$ on H_J we can consider the same procedure, but then usually $A_J = \varnothing$. The simplification of the map to be estimated as for consequence that there is no solution to the equations. So the physicist uses a statistical method, that is a map which associates to each batch of data $X(\Omega)$ a map $\varphi(X(\Omega)) = \widehat{X} \in H_J$. Usually $\widehat{X}$ is such that it minimizes the sum of the distance between points in E : $\sum_{m \in \Omega} \left\| \widehat{X}(m) - x_p \right\|_E$ (other additional conditions can be required).

The primary observable Φ gives another solution : $\Phi(X)$ is the orthogonal projection of X on the Hilbert space H_J, it is such that it minimizes the distance between maps :

$\forall Z \in H_J : \|X - Z\|_H \geq \|X - \Phi(X)\|_H$.

$\Phi(X)$ always exist, and does not depend on the choice of an estimation procedure φ. $\Phi(X)$ minimizes the distance between maps in H, meanwhile $\varphi(X(\Omega))$ minimizes distance between points in E. Usually $\varphi(X(\Omega))$ is different from $\Phi(X)$ and $\Phi(X)$ is a better estimate than $\widehat{X}$: *a primary observable is actually the best statistical estimator* for a given size of the sample. But it requires the explicit knowledge of the scalar product and H_J. This can be practically done in some significant cases (see for an example J.C.Dutailly *Estimation of the probability of transitions between phases).*

Knowing the estimate $\widehat{X}$ provided by a statistical method φ, we can implement the previous procedure to the set $\widehat{X}(\Omega)$ and compute the set of solutions : $\widehat{A} = \cap_{m_p \in \Omega} \mathcal{E}_J(m_p)^{-1}\left(\widehat{X}(m)\right)$. It is not empty. Its closed convex hull $\widehat{B}$ in H_J can be considered as the domain of confidence of $\widehat{X}$: they are maps which take the same values as $\widehat{X}$ in Ω and as a consequence give the same value to $\sum_{m \in \Omega} \left\| \widehat{X}(m) - x_p \right\|_E$.

Because $\widehat{B}$ is closed and convex there is a unique orthogonal projection Y of X on $\widehat{B}$ (Maths.1107) and :

$\forall Z \in \widehat{B} : \|X - Z\|_H \geq \|X - Y\|_H \Rightarrow \left\| X - \widehat{X} \right\|_H \geq \|X - Y\|_H$

so Y is a better estimate than $\varphi(X(\Omega))$, and can be computed if we know the scalar product on H.

We see clearly the crucial role played by the choice of a specification. But it leads to a more surprising result, of deep physical meaning.

2.2.6 Quantization of singularities

A classic problem in Physics is to prove the existence of a singular phenomenon, appearing only for some values of the parameters m. To study this problem we use a model similar to the previous one, with the same notations. But here the variable X is comprised of two maps, X_1, X_2 with unknown, disconnected, domains $M_1, M_2 : M = M_1 + M_2$. The first problem is to estimate X_1, X_2.

With a statistical process $\varphi(X(\Omega))$ it is always possible to find estimations $\widehat{X}_1, \widehat{X}_2$ of X_1, X_2. The key point is to distinguish in the set Ω the points which belong to M_1 and M_2. There are $\frac{1}{2}\left(2^N - 2\right) = 2^{N-1} - 1$ distinct partitions of Ω in two subsets $\Omega_1 + \Omega_2$, on each subset the statistical method φ gives the estimates :

$\widehat{Y}_1 = \varphi(X(\Omega_1)), \widehat{Y}_2 = \varphi(X(\Omega_2))$

Denote : $\rho(\Omega_1, \Omega_2)$

$= \sum_{m_p \in \Omega_1} \|X(m_p) - \varphi(X(\Omega_1))(m_p)\| + \sum_{m_p \in \Omega_2} \|X(m_p) - \varphi(X(\Omega_2))(m_p)\|$

A partition (Ω_1, Ω_2) is said to be a better fit than (Ω'_1, Ω'_2) if :

$\rho(\Omega_1, \Omega_2) \leq \rho(\Omega'_1, \Omega'_2)$

Then $\widehat{X}_1 = \varphi(X(\Omega_1)), \widehat{X}_2 = \varphi(X(\Omega_2))$ is the solution for the best partition.

So there is a procedure, which provides always the best solution given the data and φ, but it does not give M_1, M_2 precisely, their estimation depends on the structure of M.

However it is a bit frustrating, if we want to test a law, because the procedure provides always a solution, even if actually there is no such partition of X. And this can happen. If we define the sets as above with the evaluation map : $\mathcal{E}_J(m) : H_J \to E :: \mathcal{E}(m) Y = Y(m)$

$A_k = \cap_{m_p \in \Omega_k} \mathcal{E}(m_p)^{-1}\left(\widehat{X}_k(m_p)\right) \subset H_J$ for $k = 1, 2$. It is not empty because it contains at least $\widehat{X}_k$.

B_k the closed convex hull of A_k in H_J

Then : $\forall Y \in B_k, m \in \Omega_k : Y(m) = \widehat{X}_k(m)$

If $B_1 \cap B_2 \neq \varnothing$ there is at least one map, which can be defined uniquely on M, belongs to H_J and is equivalent to $\widehat{X}_1, \widehat{X}_2$.

This issue is of importance because many experiments aim at proving the existence of a special behavior. We need, in addition, a test of the hypothesis (denoted H_0) : there is a partition (and then the best solution would be $\widehat{X}_1, \widehat{X}_2$) against the hypothesis (denoted H_1) there is no partition : there is a unique map $\widehat{X} \in H_J$ for the domain Ω. The simplest test is to compare $\sum_{m_p \in \Omega} \|X(m_p) - \varphi(\Omega)(m_p)\|$ to $\rho(\Omega_1, \Omega_2)$. If $\varphi(\Omega)$ gives results as good as $\widehat{X}_1, \widehat{X}_2$ we can reject the hypothesis. Notice that it accounts for the properties assumed for the maps in H_J. For instance if H_J is comprised uniquely of continuous maps, then $\varphi(X(\Omega))$ is continuous, and clearly distinct from the maps $\widehat{X}_1, \widehat{X}_2$ continuous only on M_1, M_2.

It is quite obvious that the efficiency of this test decreases with N : the smaller N, the greater the chance to accept H_0. Is there a way to control the validity of an experiment ? The Theory of Tests, a branch of Statistics, studies this kind of problems.

The problem is, given a sample of points $\Omega = (m_p)_{p=1}^N$ and the corresponding values $x = (x_p)_{p=1}^N$, decide if they obey to a simple (X, Hypothesis H_1) or a double (X_1, X_2, Hypothesis H_0) distribution law.

The choice of the points $(m_p)_{p=1}^N$in a sample is assumed to be random : all the points m of M have the same probability to be in Ω, but the size of M_1, M_2 can be different, so it could give a different chance for a point of M_1 or M_2 to be in the sample. Let us say that :

$\Pr(m \in M_1|H_0) = 1-\lambda, \Pr(m \in M_2|H_0) = \lambda, \Pr(m \in M|H_1) = 1$

(all the probabilities are for a sample of a given size N)

Then the probability for any vector of E to have a given value x depends only on the map X : this is the number of points m of M for which $X(m) = x$. For instance if there are two points m with $X(m) = x$ then x has two times the probability to appear, and if X is more concentrated in an area of E, this area has more probability to appear. Let us denote this value $\rho(x) \in [0,1]$.

Rigorously (Maths.869), with a measure dx on E, μ on M, $\rho(x)\,dx$ is the pull-back of the measure μ on M. For any ϖ belonging to the Borel algebra σE of E :

$\int_\varpi \rho(x)\,dx = \int_{\mathcal{E}(m)^{-1}(\varpi)} \mu(m) \Leftrightarrow \rho(x)\,dx = X^*\mu$

If H_1 is true, the probability $\Pr(x|H_1) = \rho(x)$ depends only on the value x, that is of the map X.

If H_0 is true the probability depends on the maps and if $m \in M_1$ or $m \in M_2$ $(M = M_1 + M_2)$

$\Pr(x|H_0 \wedge m \in M_1) = \rho_1(x)$

$\Pr(x|H_0 \wedge m \in M_2) = \rho_2(x)$

$\Rightarrow \Pr(x|H_0) = (1-\lambda)\rho_1(x) + \lambda\rho_2(x)$

Moreover we have with some measure dx on E :

$\int_E \rho(x)\,dx = \int_E \rho_1(x)\,dx = \int_E \rho_2(x)\,dx = 1$

The likehood function is the probability of a given batch of data. It depends on the hypothesis :

$$L(x|H_0) = \Pr(x_1, x_2, ...x_N|H_0) = \prod_{p=1}^N ((1-\lambda)\rho_1(x_p) + \lambda\rho_2(x_p))$$

$$L(x|H_1) = \Pr(x_1, x_2, ...x_N|H_1) = \prod_{p=1}^N \rho(x_p)$$

The Theory of Tests gives us some rules (see Kendall t.II). A critical region is an area $w \subset E^N$ such that H_0 is rejected if $x \in w$.

One considers two risks :

- the risk of type I is to wrongly reject H_0. It has the probability : $\alpha = \Pr(x \in w|H_0)$

- the risk of type II is to wrongly accept H_0. It has the probability : $1-\beta = \Pr(x \in E^N - w|H_0)$ called the power of the test thus :

$\beta = \Pr(x \in w|H_1)$

A simple rule, proved by Neyman and Pearson, says that the best critical region w is defined by :

$w = \left\{ x : \frac{L(x|H_0)}{L(x|H_1)} \leq k \right\}$

the scalar k being defined by : $\alpha = \Pr(x \in w|H_0)$. So we are left with a single parameter α, which can be seen as the rigor of the test.

The critical area $w \subset E^N$ is then :

$$w = \left\{ x \in E^N : \prod_{p=1}^N \frac{((1-\lambda)\rho_1(x_p) + \lambda\rho_2(x_p))}{\rho(x_p)} \leq k \right\}$$

with :

$$\alpha = \int_w \prod_{p=1}^N ((1-\lambda)\rho_1(\xi_p) + \lambda\rho_2(\xi_p))\,(d\xi)^N$$

It provides a reliable method to build a test, but requires to know, or to estimate, $\rho, \rho_1, \rho_2, \lambda$.

In most of the cases encountered, actually one looks for an anomaly.
H_1 is unchanged, there is only one map X,defined over M. Then : $\Pr(x|H_1) = \rho(x)$
H_0 becomes :
$M = M_1 + M_2$
$\Pr(m \in M_1|H_0) = 1 - \lambda, \Pr(m \in M_2|H_0) = \lambda$
On M_1 the variable is X :
$\Pr(x_p|H_0 \wedge m_p \in M_1) = \rho(x) \Rightarrow \Pr(x_p|H_0) = (1 - \lambda)\rho(x)$
On M_2 the variable becomes X_2
$\Pr(x_p|H_0 \wedge m_p \in M_2) = \rho_2(x) \Rightarrow \Pr(x_p|H_0) = \lambda\rho_2(x)$
And w is :
$$w = \left\{ x \in E^N : \prod_{p=1}^{N} \frac{((1-\lambda)\rho(x_p)+\lambda\rho_2(x_p))}{\rho(x_p)} \leq k \right\}$$
$$w = \left\{ x \in E^N : \prod_{p=1}^{N} \left(1 - \lambda + \lambda\frac{\rho_2(x_p)}{\rho(x_p)}\right) \leq k \right\}$$
$$\alpha = \int_w \prod_{p=1}^{N} ((1-\lambda)\rho(x_p) + \lambda\rho_2(x_p))(dx)^N$$
$$\beta = \Pr(x \in w|H_1) = \int_w \left(\prod_{p=1}^{N} \rho(x_p)\right)(dx)^N$$
If there is one observed value such that $\rho(x_p) = 0$ then H_0 should be accepted. But, because ρ, ρ_2 are not well known, and the imprecision of the experiments, H_0 would be proven if $\frac{L(x|H_0)}{L(x|H_1)} > k$ for a great number of experiments. So we can say that H_0 is scientifically proven if :
$$\forall(x_1, x_2, ...x_N) : \prod_{p=1}^{N} \left((1-\lambda) + \lambda\frac{\rho_2(x_p)}{\rho(x_p)}\right) > k$$
By taking $x_1 = x_2 = ... = x_N = x$:
$\forall x : (1-\lambda) + \lambda\frac{\rho_2(x)}{\rho(x)} > k^{1/N}$
$\frac{\rho_2(x)}{\rho(x)} > \left(k^{1/N} + \lambda - 1\right)/\lambda$
When $N \to \infty : k^{1/N} \to 1 \Rightarrow \frac{\rho_2(x)}{\rho(x)} > 1$

So a necessary condition to have a chance to say that a singularity has been reliably proven is that : $\forall x : \frac{\rho_2(x)}{\rho(x)} > 1$.

The function $\frac{\rho_2(x)}{\rho(x)}$ can be called the Signal to Noise Ratio, by similarity with the Signal Theory. Notice that we have used very few assumptions about the variables. And we can state :

Theorem 16 *In a system represented by variables X which are maps defined on a set M and valued in a vector space E, a necessary condition for a singularity to be detected is that the Signal to Noise Ratio is greater than 1 for all values of the variables in E.*

This result can be seen the other way around : if a signal is acknowledged, then necessarily it is such that $\frac{\rho_2(x)}{\rho(x)} > 1$. Any other signal would be interpreted as related to the imprecision of the measure. So there is a threshold under which phenomena are not acknowledged, and their value is necessarily above this threshold. The singular phenomena are quantized. We will see an application of this result about the Planck's law.

2.3 PROBABILITY

One of the main purposes of the model is to know the state X, represented by some vector $\psi \in H$. The model is fully determinist, in that the values of the variables X are not assumed to depend on a specific event : there is no probability law involved in its definition. However the value of X which will be acknowledged at the end of the experiment, when all the data have been collected and analyzed, differs from its actual value. The discrepancy stems from the usual imprecision of any measure, but also more fundamentally from the fact that we estimate a vector in an infinite dimensional vector space from a batch of data, which is necessarily finite. We will focus on this later aspect, that is on the discrepancy between an observable $\Phi(X)$ and X.

In any practical physical experiment the estimation of X requires the choice of an observable. The most efficient solution is to choose a primary observable which, furthermore, provides the best statistical estimator. However usually neither the map Φ nor the basis $(e_i)_{i \in I}$ are explicit, even if they do exist. An observable Φ can be defined simply by choosing a finite number of independent vectors, and it is useful to assess the consequences of the choice of these vectors. So we can look at the discrepancy $X - \Phi(X)$ from a different point of view : for a given, fixed, value of the state X, what is the uncertainty which stems from the choice of Φ among a large class of observables ? This sums up to assess the risk linked to the choice of a specification for the estimation of X.

2.3.1 Primary observables

Let us start with primary observables : the observable Φ is some projection on a finite dimensional vector subspace of V.

The bases of the vector space V_0 (such that $O \subset V_0$) have the same cardinality, so we can consider that the set I does not depend on a choice of a basis (actually one can take $I = \mathbb{N}$). The set 2^I is the largest $\sigma-$algebra on I. The set $\left(I, 2^I\right)$ is measurable (Maths.802).

For any fixed $\psi \neq 0 \in H$ the function

$\widehat{\mu}_\psi : 2^I \to \mathbb{R} :: \widehat{\mu}_\psi(J) = \frac{\langle \widehat{Y}_J \psi, \psi \rangle}{\|\psi\|^2} = \frac{\left\|\widehat{Y}_J \psi\right\|^2}{\|\psi\|^2}$

is a probability law on $\left(I, 2^I\right)$: it is positive, countably additive and $\widehat{\mu}_\psi(I) = 1$ (Maths.11.4.1).

The choice of a finite subset $J \in 2^I$ can be seen as an event from a probabilist point of view. For a given $\psi \neq 0 \in H$ the quantity $\widehat{Y}_J(\psi)$ is a random variable, with a distribution law $\widehat{\mu}_\psi$

The operator $\widehat{Y}_J$ has two eigen values : 1 with eigen space $\widehat{Y}_J(H)$ and 0 with eigen space $\widehat{Y}_{J^c}(H)$. Whatever the primary observable, the value of $\Phi(X)$ will be $Y_J(X)$ for some J, that is an eigen vector of the operator $\Phi = Y_J$, and the probability to observe $\Phi(X)$, if the system is in the state X, is :

$\Pr(\Phi(X) = Y_J(X)) = \Pr(J|\psi) = \widehat{\mu}_\psi(J) = \frac{\left\|\widehat{Y}_J \psi\right\|^2}{\|\psi\|^2} = \frac{\left\|\widehat{\Phi}(\Upsilon(X))\right\|_H^2}{\|\Upsilon(X)\|_H^2}$

This result still holds if another basis had been chosen : $\Phi(X)$ will be $Y_J(X)$ for some J, expressed in the new basis, but with a set J of same cardinality. And some specification must always be chosen. So we have :

Theorem 17 *For any primary observable Φ, the value $\Phi(X)$ which is measured is an eigen vector of the operator Φ, and the probability to measure a value $\Phi(X)$ if the system is in the state X is :*

$\Pr(\Phi(X)|X) = \frac{\left\|\widehat{\Phi}(\Upsilon(X))\right\|_H^2}{\|\Upsilon(X)\|_H^2}$

2.3.2 Secondary observables

For a secondary observable, as defined previously :

$\Phi = \sum_{p=1}^{n} \lambda_p Y_{J_p}$

$\widehat{\Phi} = \sum_{p=1}^{n} \lambda_p \widehat{\pi}_{H_p}$

The vectors decompose as :

$X = Y_{J^c}(X) + \sum_{p=1}^{n} X_p$

with $X_p = Y_{J_p}(X) = \sum_{i \in J_p} \left\langle \phi_i, \widehat{Y}_{Jp}(\Upsilon(X)) \right\rangle_H e_i \in V_p$

$\Upsilon(X) = \psi = \psi_c + \sum_{p=1}^{n} \psi_p$ with $\psi_p = \widehat{\pi}_{H_p}(\psi), \psi_c = \widehat{\pi}_{H_c}(\psi)$

where $\psi_p$is an eigen vector of $\widehat{\Phi}, X_p$ is an eigen vector of Φ both for the eigen value λ_p and

$\Phi(X) = \sum_{p=1}^{n} \lambda_p X_p$

$\widehat{\Phi}(\psi) = \sum_{p=1}^{n} \lambda_p \psi_p$

If, as above, we see the choice of a finite subset $J \in 2^I$ as an event in a probabilist point of view then the probability that $\Phi(X) = \lambda_p X_p$ if the system is in the state X, is given by $\Pr(J_p|X) = \frac{\|\widehat{Y}_p\psi\|^2}{\|\psi\|^2} = \frac{\|\psi_p\|^2}{\|\psi\|^2}$

And we have :

Theorem 18 *For any secondary observable Φ, the value $\Phi(X)$ which is observed if the system is in the state X is a linear combination of eigen vectors X_p of Φ for the eigen value λ_p: $\Phi(X) = \sum_{p=1}^{n} \lambda_p X_p$*

The probability that $\Phi(X) = \lambda_p X_p$ is:

$\Pr(\Phi(X) = \lambda_p X_p | X) = \frac{\|\Upsilon(X_p)\|^2}{\|\Upsilon(X)\|^2}$

Which can also be stated as : $\Phi(X)$ can take the values $\lambda_p X_p$, each with the probability $\frac{\|\psi_p\|^2}{\|\psi\|^2}$, then $\Phi(X)$ reads as an expected value. This is the usual way it is expressed in QM.

The interest of these results comes from the fact that we do not need to explicit any basis, or even the set I. And we do not involve any specific property of the estimator of X, other than Φ is an observable. The operator $\widehat{\Phi}$ sums up the probability law.

Of course this result can be seen in another way : as only $\Phi(X)$ can be accessed, one can say that the system takes only the states $\Phi(\lambda_p X_p)$, with a probability $\frac{\|\psi_p\|^2}{\|\psi\|^2}$. This gives a probabilistic behavior to the system (X becoming a random variable) which is not present in its definition, but is closer to the usual interpretation of QM.

This result can be illustrated by a simple example. Let us take a model where a function x is assumed to be continuous and take its values in $\mathbb{R}$. It is clear that any physical measure will at best give a rational number $Y(x) \in \mathbb{Q}$ up to some scale. There are only countably many rational numbers for unaccountably many real scalars. So the probability to get $Y(x) \in \mathbb{Q}$ should be zero. The simple fact of the measure gives the paradox that rational numbers have an incommensurable weight, implying that each of them has some small, but non null, probability to appear. In this case I can be assimilated to $\mathbb{Q}$, the subsets J are any finite collection of rational numbers.

2.3.3 Wave function

The wave function is a central object in QM, but it has no general definition and is deemed non physical (except in the Bohm's interpretation). Usually this is a complex valued function, defined over the space of configuration of the system : the set of all possible values of the variables representing the system. If it is square integrable, then it belongs to a Hilbert space, and can be assimilated to the vector representing the state. Because its arguments comprise the coordinates of objects such as particles, it has a value at each point, and the square of the module of the function is proportional to the probability that the measure of the variable takes the values of the arguments at this point. Its meaning is relatively clear for systems comprised of particles, but less so for systems which include force fields, because the space of configuration is not defined. But it can be precisely defined in our framework.

Theorem 19 *In a system modelled by N variables, collectively denoted X, which are maps : $X : M \to F$ from a common measured set M to a finite dimensional normed vector space F and belonging to an open subset of an infinite dimensional, separable, real Fréchet vector space V, such that the evaluation map : $\mathcal{E}(m) : V \to F :: \mathcal{E}(m)(X) = X(m)$ which assigns at any X its value in a fixed point m in M is measurable : then for any state X of the system there is a function : $W : M \times F \to \mathbb{R}$ such that $W(m,y) = \Pr(\Phi(X)(m) = y|X)$ is the probability that the measure of the value of any primary observable $\Phi(X)$ at m is y.*

Proof. The conditions 1 apply, there is a Hilbert space H and an isometry $\Upsilon : V \to H$.

To the primary observable $\Phi : V \to V_J$ is associated the self-adjoint operator $\widehat{\Phi} = \Upsilon \circ \Phi \circ \Upsilon^{-1}$

We can apply the theorem 17: the probability to measure a value $\Phi(X) = Y$ if the system is in the state X is :

$\Pr(\Phi(X) = Y|X) = \frac{\left\|\widehat{\Phi}(\Upsilon(Y))\right\|_H^2}{\|\Upsilon(X)\|_H^2} = \pi(Y)$

Because only the maps belonging to V_J are observed it provides a probability law π on the set V_J : $\pi : V_\sigma \to [0,1]$ where V_σ is the Borel algebra of V_J.

The evaluation map : $\mathcal{E}_J(m) : V_J \to F :: \mathcal{E}_J(m)(Y) = Y(m)$ assigns at any $Y \in V_J$ its value in the fixed point m in M.

If $y \in F$ is a given vector of F, the set of maps in V_J which gives the value y in m is : $\varpi(m,y) = \mathcal{E}_J(m)^{-1}(y) \subset V_J$.

The probability that the observable takes the value y at m $\Phi(X)(m) = y$ is

$\pi(\varpi(m,y)) = \pi\left(\mathcal{E}_J(m)^{-1}(x)\right)$

$= \frac{1}{\|\Upsilon(X)\|_H^2} \int_{Y \in \varpi(m,y)} \left\|\widehat{\Phi}(\Upsilon(Y))\right\|_H^2 \pi(Y) = W(m,y)$ ■

If M is endowed with a positive measure μ and X is a scalar function, the space V of square integrable maps $\int_\Omega |X(m)|^2 \mu(m) < \infty$ is a separable Hilbert space H, then the conditions 1 are met and H can be identified with the space of the states.

$W(m,y) = \frac{1}{\|X\|_H^2} \int_{Y \in \varpi(m,y)} |Y|_H^2 = \left(\int_\Omega |X|^2 \mu\right)^{-1} \mu\left(Y^{-1}(m,y)\right)$

No structure, other than the existence of the measure μ, is required on M. But of course if the variables X include derivatives M must be at least a differentiable manifold.

W can be identified with the square of the wave function of QM.

2.4 CHANGE OF VARIABLES

In the conditions 1 the variables can be defined over different connected domains. Actually one can go further and consider the change of variables, which leads to a theorem similar to the well known Wigner's theorem. The problem appears in Physics in two different ways, which reflect the interpretations of Scientific laws.

2.4.1 Two ways to define the same state of a system

The first way : from a theoretical model

In the first way the scientist has built a theoretical model, using known concepts and their usual representation by mathematical objects. A change of variables appears notably when :

i) The variables are the components of a geometric quantity (a vector, a tensor,...) expressed in some basis. According to the general Principle of Relativity, the state of the system shall not depend on the observers (those measuring the coordinates). For instance it should not matter if the state of a system is measured in different units. The data change, but according to rules which depend on the mathematical representation which is used, and not on the system itself. In a change of basis coordinates change but they represent the same vectorial quantity.

ii) The variables are maps, depending on arguments which are themselves coordinates of some event : $X_k = X_k\left(\xi_1, ...\xi_{p_k}\right)$. Similarly these coordinates ξ can change according to some rules, while the variable X_k represents the same event.

By definition in both cases there is a continuous bijective map $U : V \rightarrow V'$ such that X and $X' = U(X)$ represent the same state of the system. This is the way mathematicians see a change of variables, and is usually called the passive way by physicists.

Any primary or secondary observable Φ is a linear map $\Phi \in L\left(V; W\right)$ into a finite dimensional vector subspace W. For the new variable the observable is $\Phi' \in L\left(V; W'\right)$. Both $W, W' \subset V$ but W' is not necessarily identical to W. However the assumption that $X' = U(X)$ and X represents the same state of the system implies that for any measure of the state we have a similar relation : $\Phi' \circ U\left(X\right) = U \circ \Phi\left(X\right) \Leftrightarrow \Phi' \circ U = U \circ \Phi$. This is actually the true meaning of "represent the same state". This means that actually one makes the measures according to a fixed procedure, given by Φ, on variables which vary with U. Because U is a bijection on V : $\Phi' = U \circ \Phi \circ U^{-1}$.

The second way : from experimental measures

In the second way the scientist makes measures with a device that can be adjusted according to different values of a parameter, say θ : the simplest example is using different units, but often it is the orientation of the device which can be changed. And the measures $Y\left(\theta\right)$ which are taken are related to the choice of parameter for the device. If the results of experiments show that $Y\left(\theta\right) = Q\left(\theta\right)Y\left(\theta_0\right)$ with a bijective map $Q\left(\theta\right)$ and θ_0 some fixed value of the parameter one can assume that this experimental relation is a feature of the system itself.

Physicists distinguish a passive transformation, when only the device changes, and an active transformation, when actually the experiment involves a physical change on the system. In a passive transformation we come back to the first way and it is legitimate to assume that we have actually the same state, represented by different data, reflecting some mathematical change in their expression, even if the observable, which is valued in a finite dimensional space, does not account for all the possible values of the variables. In an active transformation (for instance in the Stern-Gerlach experiment one changes the orientation of a magnetic field to which the particles are submitted) one can say that there is some map U acting on the space V of the states of the system, such that the measure is done by a unique procedure $\widetilde{\Phi}$ on a state X which is changed by a map $U(\theta)$. So that the measures are $Y(\theta) = \widetilde{\Phi} \circ U(\theta) X$ and the relation $Y(\theta) = Q(\theta) Y(\theta_0)$ reads : $\widetilde{\Phi} \circ U(\theta)(X) = U(\theta) \circ \widetilde{\Phi}(X)$. So this is very similar to the first case, where θ represents the choice of a frame.

In both cases there is the general idea that the state of the system is represented by some fixed quantity, which can be measured in different procedures, so that there is a relation, given by the way one goes from one procedure to the others, between the measures of the state. In the first way the conclusion comes from the mathematical definition in a theoretical model : this is a simple mathematical deduction using the Principle of Relativity. In the second way there is an assumption : that one can extend the experimental facts, necessarily limited to a finite number of data, to the whole set of possible values of the variable.

The Theorem 2 is based on the existence of a Fréchet manifold structure on the set of possible values of the maps X. The same manifold structure can be defined by different, compatible, atlas. So the choice of other variables can lead to the same structure, and the fixed quantity that we identify with a state is just a point on the manifold, and the change of variables is a change of charts between compatible atlas. The variables must be related by transition maps, that is continuous bijections, but additional conditions are required, depending on the manifold structure considered. For instance for differentiable manifolds the transition maps must be differentiable. We will request that the transition maps preserve the positive kernel, which plays a crucial role in Fréchet manifolds.

2.4.2 Fundamental theorem for a change of variables

We will summarize these features in the following :

Condition 20

i) The same system is represented by the variables $X = (X_1, ...X_N)$ and $X' = (X'_1, ...X'_{N'})$ which belong to open subsets O, O' of the infinite dimensional, separable, Fréchet vector space V.

ii) There is a continuous map $U : V \to V$, bijective on (O, O'), such that X and $X' = U(X)$ represent the same state of the system

iii) U preserves the positive kernel on V[2]

iv) For any observable Φ of X, and Φ' of X' : $\Phi' \circ U = U \circ \Phi$

The map U shall be considered as part of the model, as it is directly related to the definition of the variables, and is assumed to be known. There is no hypothesis that it is linear.

Theorem 21 *Whenever a change of variables on a system meets the conditions 20 above,*

i) there is a unitary, linear, bijective map $\widehat{U} \in \mathcal{L}(H; H)$ such that : $\forall X \in O : \widehat{U}(\Upsilon(X)) = \Upsilon(U(X))$ where H is the Hilbert space and Υ is the linear map : $\Upsilon : V \to H$ associated to X, X'

ii) U is necessarily a bijective linear map.

For any observables Φ, Φ':

iii) $W' = \Phi'(V)$ is a finite dimensional vector subspace of V, isomorphic to $W = \Phi(V)$: $W' = U(W)$

iv) the associated operators $\widehat{\Phi} = \Upsilon \circ \Phi \circ \Upsilon^{-1}, \widehat{\Phi}' = \Upsilon \circ \Phi' \circ \Upsilon^{-1}$ are such that : $\widehat{\Phi}' = \widehat{U} \circ \widehat{\Phi} \circ \widehat{U}^{-1}$ and $H'_{\Phi'} = \widehat{\Phi}'(H)$ is a vector subspace of H isomorphic to $H_\Phi = \widehat{\Phi}(H)$

[2] The positive kernel plays a role similar to the probability of transition between states of the Wigner's Theorem.

Proof. i) Let $V_0 = O \cup O'$. This is an open set and we can apply the theorem 2. There is a homeomorphism $\Xi : V_0 \to H_0$ where H_0 is an open subset of a Hilbert space H. For a basis $(e_i)_{i \in I}$ of Span(V_0) there is an isometry Υ such that :

$\Upsilon : V_0 \to H_0 :: \Upsilon(Y) = \sum_{i \in I} \langle \phi_i, \Upsilon(Y) \rangle_H \varepsilon_i$

such that :

$\forall i \in I : \varepsilon_i = \Upsilon(e_i)$;

$\forall i, j \in I : \langle \phi_i, \varepsilon_j \rangle_H = \delta_{ij}$;

ii) Υdefines a positive kernel on $V_0 : K_V(Y_1, Y_2) = \langle \Upsilon Y_1, \Upsilon Y_2 \rangle_H$

The sets (V_0, Υ, H) and $(V_0, \Upsilon U, H)$ are two realizations triple of K_V. Then there is an isometry φ such that :

$\Upsilon U = \varphi \circ \Upsilon$ (Maths.1200).

$\langle UX_1, UX_2 \rangle_V = \langle \Upsilon UX_1, \Upsilon UX_2 \rangle_H = \langle \varphi \circ \Upsilon X_1, \varphi \circ \Upsilon X_2 \rangle_H$

$= \langle \Upsilon X_1, \Upsilon X_2 \rangle_H = \langle X_1, X_2 \rangle$

So U preserves the scalar product on V

Let be : $\widehat{U} = \Upsilon \circ U \circ \Upsilon^{-1}$

$\left\langle \widehat{U}\psi_1, \widehat{U}\psi_2 \right\rangle_H = \left\langle \Upsilon \circ U \circ \left(\Upsilon^{-1}\psi_1\right), \Upsilon \circ U \circ \left(\Upsilon^{-1}\psi_2\right) \right\rangle_H$

$= \left\langle U \circ \left(\Upsilon^{-1}\psi_1\right), U \circ \left(\Upsilon^{-1}\psi_2\right) \right\rangle_V = \left\langle \left(\Upsilon^{-1}\psi_1\right), \left(\Upsilon^{-1}\psi_2\right) \right\rangle_V$

$= \langle \psi_1, \psi_2 \rangle_H$

So $\widehat{U}$ preserves the scalar product on H

iii) As seen in Theorem 2 starting from the basis $(\varepsilon_i)_{i \in I}$ of H one can define a Hermitian basis $(\widetilde{\varepsilon}_i)_{i \in I}$ of H, an orthonormal basis $(\widetilde{e}_i)_{i \in I}$ of V for the scalar product $K_V = \langle \rangle_V$ with $\widetilde{e}_i = \Upsilon^{-1}(\widetilde{\varepsilon}_i)$

U is defined for any vector of V, so for $(\widetilde{e}_i)_{i \in I}$ of V.

Define : $\widehat{U}(\widetilde{\varepsilon}_i) = \widehat{U}(\Upsilon(\widetilde{e}_i)) = \Upsilon(U(\widetilde{e}_i)) = \widetilde{\varepsilon}'_i$

The set of vectors $(\widetilde{\varepsilon}'_i)_{i \in I}$ is an orthonormal basis of H:

$\left\langle \widetilde{\varepsilon}'_i, \widetilde{\varepsilon}'_j \right\rangle_H = \left\langle \widehat{U}(\Upsilon(\widetilde{e}_i)), \widehat{U}(\Upsilon(\widetilde{e}_j)) \right\rangle_H = \langle \widetilde{e}_i, \widetilde{e}_j \rangle_V = \delta_{ij}$

The map : $\chi : \ell^2(I) \to H :: \chi(y) = \sum_{i \in I} y_i \widetilde{\varepsilon}'_i$ is an isomorphism (same as in Theorem 2) and $(\widetilde{\varepsilon}'_i)_{i \in I}$ is a Hilbertian basis of H. So we can write :

$\forall \psi \in H : \psi = \sum_{i \in I} \psi^i \widetilde{\varepsilon}_i, \widehat{U}(\psi) = \sum_{i \in I} \psi'^i \widetilde{\varepsilon}'_i$

and : $\psi^i = \langle \widetilde{\varepsilon}_i, \psi \rangle_H = \left\langle \widehat{U}(\widetilde{\varepsilon}_i), \widehat{U}(\psi) \right\rangle_H = \left\langle \widetilde{\varepsilon}'_i, \sum_{j \in I} \psi'^j \widetilde{\varepsilon}'_j \right\rangle_H = \psi'^i$

Thus the map $\widehat{U}$ reads : $\widehat{U} : H \to H :: \widehat{U}\left(\sum_{i \in I} \psi^i \widetilde{\varepsilon}_i\right) = \sum_{i \in I} \psi^i \widetilde{\varepsilon}'_i$

It is linear, continuous and unitary : $\left\langle \widehat{U}(\psi_1), \widehat{U}(\psi_2) \right\rangle = \langle \psi_1, \psi_2 \rangle$ and $\widehat{U}$ is invertible

$U = \Upsilon^{-1} \circ \widehat{U} \circ \Upsilon$ is linear and bijective

iv) For any primary or secondary observable Φ there is a self-adjoint, Hilbert-Schmidt and trace class operator $\widehat{\Phi}$ on the associated Hilbert space H such that : $\widehat{\Phi} = \Upsilon \circ \Phi \circ \Upsilon^{-1}$. For the new variable the observable is $\Phi' \in L(V; W')$ and $W\prime \subset V$ is not necessarily identical to W. It is associated to the operator : $\widehat{\Phi}' = \Upsilon \circ \Phi' \circ \Upsilon^{-1}$. W and W' are finite dimensional vector subspaces of V.

$$\begin{array}{ccccccccccccc}
 & & \Phi & & & & U & & & & \Phi' & & \\
W & \leftarrow & \leftarrow & \leftarrow & V & \rightarrow & \rightarrow & \rightarrow & V & \rightarrow & \rightarrow & \rightarrow & W' \\
\downarrow & & & & \downarrow & & & & \downarrow & & & & \downarrow \\
\downarrow & \Upsilon & & & \downarrow & \Upsilon & & \Upsilon & \downarrow & & & \Upsilon & \downarrow \\
\downarrow & & \widehat{\Phi} & & \downarrow & & \widehat{U} & & \downarrow & & \widehat{\Phi'} & & \downarrow \\
H_{\Phi} & \leftarrow & \leftarrow & \leftarrow & H & \rightarrow & \rightarrow & \rightarrow & H & \rightarrow & \rightarrow & \rightarrow & H_{\Phi'}
\end{array}$$

Because U is a bijection on V : $\Phi' \circ U = U \circ \Phi \Rightarrow \Phi' = U \circ \Phi \circ U^{-1}$ and V is globally invariant by U

$\Phi'(V) = W' = U \circ \Phi \circ U^{-1}(V) = U \circ \Phi(V) = U(W)$

thus W' is a vector subspace of V isomorphic to W

$\widehat{\Phi}' = \Upsilon \circ \Phi' \circ \Upsilon^{-1} = \Upsilon \circ U \circ \Phi \circ U^{-1} \circ \Upsilon^{-1} = \widehat{U} \circ \Upsilon \circ \Phi \circ \Upsilon^{-1} \circ \widehat{U}^{-1} = \widehat{U} \circ \widehat{\Phi} \circ \widehat{U}^{-1}$

Let us denote : $\widehat{\Phi}(H) = H_{\Phi}, \widehat{\Phi}'(H) = H_{\Phi'}$

$\widehat{U}(H) = H$ because it is a unitary map

$\widehat{\Phi}'(H) = \widehat{U} \circ \widehat{\Phi} \circ \widehat{U}^{-1}(H) = \widehat{U} \circ \widehat{\Phi}(H) = \widehat{U}(H_{\Phi}) = H_{\Phi'}$

thus $H_{\Phi'}$ is a vector subspace of H isomorphic to H_{Φ} ■

As a consequence the map U is necessarily linear, even if this was not assumed in the conditions 20 : variables which are not linearly related (in the conditions 20) cannot represent the same state.

As $\widehat{U}$ is unitary, it cannot be self adjoint or trace class (except if $U = Id$). So it differs from an observable.

2.4.3 Change of units

A special case of this theorem is the choice of units to measure the variables. A change of units is a map : $X'_k = \alpha_k X_k$ with fixed scalars $(\alpha_k)_{k=1}^N$. As we must have :

$\langle U(X_1), U(X_2)\rangle_V = \langle X_1, X_2\rangle_V = \sum_{k=1}^N \alpha_k^2 \langle X_1, X_2\rangle_V = \langle X_1, X_2\rangle_V \Rightarrow \sum_{k=1}^N \alpha_k^2 = 1$

which implies for any single variable $X_k : \alpha_k = 1$. So the variables in the model should be dimensionless quantities. This is in agreement with the elementary rule that any formal theory should not depend on the units which are used.

More generally whenever one has a law which relates quantities which are not expressed in the same units, there should be some fundamental constant involved, to absorb the discrepancy between the units. For instance some Physicals laws involve an exponential, such as the wave equation for a plane wave :

$\psi = \exp i \left(\left\langle \overrightarrow{k}, \overrightarrow{r} \right\rangle - \varpi t \right)$

They require that the argument in the exponential is dimensionless, and because $\overrightarrow{r}$ is a length and t a time we should have a fundamental constant with the dimension of a speed (in this case c).

But also it implies that there should be some "universal system of units" (based on a single quantity) in which all quantities of the theory can be measured. In Physics this is the Planck's system which relate the units of different quantities through the values of the fundamental constants c, G (gravity), R (Boltzmann constant), $\hbar$, and the charge of the electron (see Wikipedia for more).

Usually the variables are defined with respect to some frame, then the rules for a change of frame have a special importance and are a defining feature of the model. When the rules involve a group, the previous theorem can help to precise the nature of the abstract Hilbert space H and from there the choice of the maps X.

2.4.4 Group representation

The theory of group representation is a key tool in Physics. We will remind some basic results here, see Maths.23 for a comprehensive study of this topic.

The left action of a group G on a set E is a map :$\lambda : G \times E \rightarrow E :: \lambda(g,x)$ such that $\lambda(gg',x) = \lambda(g,\lambda(g',x)), \lambda(1,x) = x$. And similarly for a right action $\rho(x,g)$.

The representation of a group G is a couple (E,f) of a vector space E and a continuous map $f : G \rightarrow G\mathcal{L}(E;E)$ (the set of linear invertible maps from E to E) such that :

$\forall g,g' \in G : f(g \cdot g') = f(g) \circ f(g') ; f(1) = Id \Rightarrow f\left(g^{-1}\right) = f(g)^{-1}$

A representation is **faithful** if f is bijective.

A vector subspace F is **invariant** if $\forall u \in F, g \in G : f(g)u \in F$

A representation is **irreducible** if there is no other invariant subspace than $E, 0$.

A representation is not unique : from a given representation one can build many others. The sum of the representations $(E_1,f_1),(E_2,f_2)$ is $(E_1 \oplus E_2, f_1 + f_2)$.

A representation is **unitary** if there is a scalar product on E and $f(g)$ is unitary : $\forall u,v \in F, g \in G : \langle f(g)u, f(g)v\rangle = \langle u,v\rangle$

If two groups $G, G\prime$ are isomorphic by ϕ, then a representation (E,f) of G provides a representation of $G\prime$:

$\phi : G' \rightarrow G :: \forall g,g' \in G' : \phi(g \cdot g') = \phi(g) \cdot \phi(g')$;

$\phi(1_{G'}) = 1_G \Rightarrow \phi\left(g^{-1}\right) = \phi(g)^{-1}$

$f : G \rightarrow G\mathcal{L}(E;E)$

Define $f' : G' \rightarrow G\mathcal{L}(E;E) :: f'(g') = f(\phi(g'))$

$f'(g_1' \cdot g_2') = f(\phi(g_1' \cdot g_2')) = f(\phi(g_2')) \circ f(\phi(g_1')) = f'(g_1') \circ f'(g_2')$

A **Lie group** is a group endowed with the structure of a manifold. On the tangent space T_1G at its unity (that we will denote 1) there is an algebraic structure of **Lie algebra**, that we will also denote generally T_1G, endowed with a bracket [] which is a bilinear antisymmetric map on T_1G.

If G is a Lie group with Lie algebra T_1G and (E,f) a representation of G, then $(E, f'(1))$ is a representation of the Lie algebra T_1G :

$f'(1) \in \mathcal{L}(T_1G; \mathcal{L}(E;E))$

$\forall X,Y \in T_1G : f'(1)([X,Y]) = f'(1)(X) \circ f'(1)(Y) - f'(1)(Y) \circ f'(1)(X)$

The converse, from the Lie algebra to the group, holds if G is simply connected, otherwise a representation of the Lie algebra provides usually multiple valued representations of the group (we will see important examples later).

Any Lie group G has the **adjoint representation** (T_1G, Ad) over its Lie algebra.

Any irreducible representation of a commutative (abelian) group is unidimensional.

Any unitary representation of a compact or finite group is reducible in the sum of orthogonal, finite dimensional, irreducible unitary representations.

Any representation of a group on a finite dimensional vector space becomes a representation on a set of matrices by choosing a basis. The representations of the common groups of matrices are tabulated. In the standard representation $(K^n, \imath)$ of a group G of $n \times n$ matrices on a field K the map $\imath$ is the usual action of matrices on column vectors in the space K^n. If G is a Lie group then the standard representation of its Lie algebra is the representation $(K^n, \imath)$ by matrices, deduced by derivation.

Two representations $(E,f),(F,\rho)$ of the same group G are **equivalent** if there is an isomorphism : $\phi : E \rightarrow F$ such that :

$\forall g \in G : f(g) = \phi^{-1} \circ \rho(g) \circ \phi$

Then from a basis $(e_i)_{i \in I}$ of E one deduces a basis $|e_i>$ of F by : $|e_i >= \phi(e_i)$. Because ϕ is an isomorphism $|e_i>$ is a basis of F. Moreover the matrix of the action of G is in this basis the same as for (E, f) :

$\rho(g)|e_i >= \sum_{j \in J} [\rho(g)]^i_j |e_j >= \rho(g) \phi(e_i) = \phi \circ f(g)(e_i)$

$= \phi\left(\sum_{j \in I} [f(g)]^j_i e_j\right) = \sum_{p \in I} [f(g)]^j_i \phi(e_j) = \sum_{p \in I} [f(g)]^j_i |e_j >$

$[\rho(g)] = [f(g)]$

If K is a subgroup of G, and (E, f) a representation of G, then (E, f) is a subrepresentation of K.

The vector subspaces F of E which are invariant by K provide representations (F, f) of K.

2.4.5 Application to groups of transformations

Change of variable parametrized by a group

This is the usual case in Physics. The second point of view that we have noticed above is clear when U is defined by a group. The system is represented by fixed variables, and the measures are taken according to procedures which change with g and we have :

$\Phi(g)(X) = U(g) \circ \Phi(1)(X)$

$\Phi \in L(V; W)$ and $U(g)$ is a bijection so X and $\Phi(1)(X)$ are in bijective correspondence and X must belong to $W \subset V$: we reduce the definition of the states at what can be observed. And to assume that this is true for any observable leads to redefine X as in the first way, but this requires and additional assumption.

Theorem 22 *If the conditions 20 are met, and (V, U) is a representation of the group G, then:*

i) $\left(H, \widehat{U}\right)$ is a unitary representation of the group G with $\widehat{U}(g) = \Upsilon \circ U(g) \circ \Upsilon^{-1}$

ii) For any observable $\Phi \in L(V; W)$ the vector space $W \subset V$ is invariant by U and (W, U) is a representation of G, and for the associated operator $\widehat{\Phi} = \widehat{U}(g) \circ \widehat{\Phi} \circ \widehat{U}(g)^{-1} \in L(H; H_\Phi)$, $\left(H_\Phi, \widehat{U}\right)$ is a finite dimensional unitary representation of the group G.

If G is a Lie group, and U continuous, then :

iii) U is smooth, $\widehat{U}$ is differentiable and $\left(\widehat{U}'(1), H\right)$ is an anti-symmetric representation of the Lie algebra T_1G of G

iv) For any observable $\Phi \in L(V; W)$ $\left(H_\Phi, \widehat{U}'(1)\right)$ is an anti-symmetric representation of the Lie algebra T_1G of G

If (F, f) is a unitary representation of G, equivalent to $\left(H_\Phi, \widehat{U}\right)$, and Φ a primary or secondary observable, then :

v) The results of measures of Φ for two values $1, g$ and the same state of the system are related by :

$\Phi \circ U(1)(X) = \sum_{j \in J} X^j(1) e_j, \Phi \circ U(g)(X) = \sum_{j \in J} X^j(g) e_j$ *for some basis $(e_i)_{i \in I}$ of V*

$X^j(g) = \sum_{k \in J} [f(g)]^j_k X^k(1)$ *where $[f(g)]$ is the matrix of $f(g)$ in orthonormal bases of F*

vi) If moreover G is a Lie group and U, f continuous, then the action $U'(1)(\kappa_a)$ of $U'(1)$ for vectors κ_a of T_1G are expressed by the same matrices $[K_a]$ of the action $f'(1)(\kappa_a)$:

$f'(1)(\kappa_a)(f_j) = \sum_{k \in J} [K_a]^k_j f_k \to U'(1)(\kappa_a)(e_j) = \sum_{k \in J} [K_a]^k_j e_k$

and similarly for the observable $\Phi : \Phi \circ U'(1)(\kappa_a)(e_j) = \sum_{k \in J} [K_a]^k_j e_k$

Proof. i) The map : $U : G \to G\mathcal{L}(V;V)$ is such that : $U(g \cdot g') = U(g) \circ U(g'); U(1) = Id$ where G is a group and 1 is the unit in G.

Then $U(g)$ is necessarily invertible, because $U(g^{-1}) = U(g)^{-1}$

$\widehat{U} : G \to \mathcal{L}(H;H) :: \widehat{U} = \Upsilon \circ U \circ \Upsilon^{-1}$ is such that :

$\widehat{U}(g \cdot g') = \Upsilon \circ U(g \cdot g') \circ \Upsilon^{-1} = \Upsilon \circ U(g) \circ U(g') \circ \Upsilon^{-1} = \Upsilon \circ U(g) \circ \Upsilon^{-1} \circ \Upsilon \circ U(g') \circ \Upsilon^{-1} = \widehat{U}(g) \circ \widehat{U}(g')$

$\widehat{U}(1) = \Upsilon \circ U(1) \circ \Upsilon^{-1} = Id$

So $\left(H, \widehat{U}\right)$ is a unitary representation of the group G ($\widehat{U}(g)$ is bijective, thus invertible).

ii) For any observable : $\Phi \circ U(g) = U(g) \circ \Phi, \widehat{\Phi} = \widehat{U}(g) \circ \widehat{\Phi} \circ \widehat{U}(g)^{-1}$

Let us take $Y \in W = \Phi(V) : \exists X \in V : Y = \Phi(X)$

$U(g) Y = U(g)(\Phi(X)) = \Phi(U(g) X) \in \Phi(V)$

And similarly

$\widehat{Y} \in \widehat{\Phi}(H) : \exists \psi \in H : \widehat{Y} = \widehat{\Phi}(\psi)$

$\widehat{U}(g) \widehat{Y} = \widehat{U}(g)\left(\widehat{\Phi}(\psi)\right) = \widehat{\Phi}\left(\widehat{U}(g)\psi\right) \in \widehat{\Phi}(H)$

thus $W, H_\Phi = \widehat{\Phi}(H)$ are invariant by U,$\widehat{U}$

The scalar product on H holds on the finite dimensional subspace $\widehat{\Phi}(H)$, which is a Hilbert space.

iii) If G is a Lie group and the map $U : G \to \mathcal{L}(V;V)$ continuous, then it is smooth (Maths.1789), $\widehat{U}$ is differentiable and $\left(\widehat{U}'(1), H\right)$ is an anti-symmetric representation of the Lie algebra T_1G of G :

$\forall \kappa \in T_1G : \left(\widehat{U}'(1)\kappa\right)^* = -\left(\widehat{U}'(1)\kappa\right)$

$\widehat{U}(\exp \kappa) = \exp \widehat{U}'(1)\kappa$ where the first exponential is taken on T_1G and the second on $\mathcal{L}$(H;H) (Maths.1886).

iv) Φ is a primary or secondary observable, and so is $\Phi \circ U(g)$, then $\widehat{\Phi} \circ \widehat{U}(g) = \widehat{U}(g) \circ \widehat{\Phi}$ is a self-adjoint, compact operator, and by the Riesz theorem (Math.1142) its spectrum is either finite or is a countable sequence converging to 0 (which may or not be an eigen value) and, except possibly for 0, is identical to the set $(\lambda_p(g))_{p \in \mathbb{N}}$ of its eigen values (Maths.1020). For each distinct eigen value the eigen spaces $H_p(g)$ are orthogonal and H is the direct sum $H = \oplus_{p \in \mathbb{N}} H_p(g)$. For each non null eigen value $\lambda_p(g)$ the eigen space $H_p(g)$ is finite dimensional. For a primary observable the eigen values are either 1 or 0.

Because H_Φ is finite dimensional, for each value of g there is an orthonormal basis $(\widetilde{\varepsilon}_i(g))_{i \in J}$ of H_Φ comprised of a finite number of vectors which are eigen vectors of $\widehat{\Phi} \circ \widehat{U}(g) : \widehat{\Phi} \circ \widehat{U}(g)(\widetilde{\varepsilon}_j(g)) = \lambda_j(g)\widetilde{\varepsilon}_j(g)$

Any vector of H_Φ reads :

$\psi = \sum_{j \in J} \psi^j(g)\widetilde{\varepsilon}_j(g)$ and

$\widehat{\Phi} \circ \widehat{U}(g) = \sum_{p \in \mathbb{N}} \lambda_p(g)\widehat{\pi}_{H_p(g)}$ with the orthogonal projection $\widehat{\pi}_{H_p(g)}$ on $H_p(g)$.

And, because any measure belongs to H_Φ it is a linear combination of eigen vectors

$\Phi \circ U(g)(X) = \Upsilon^{-1} \circ \widehat{\Phi} \circ \widehat{U}(g) \circ \Upsilon(X) = \Upsilon^{-1}\left(\sum_{j \in J} \lambda_j(g)\psi^j(g)\widetilde{\varepsilon}_j(g)\right)$

$= \sum_{j \in J} \lambda_j(g)\psi^j \Upsilon^{-1}(\widetilde{\varepsilon}_j(g)) = \sum_{j \in J} \lambda_j(g)\psi^j e_j(g)$

for some basis $(e_i)_{i \in I}$ of V : $e_j(g) = \Upsilon^{-1}(\widetilde{\varepsilon}_j(g))$ and $\Phi \circ U(g)(e_j(g)) = \lambda_j e_j(g)$

That we can write :

$\Phi \circ U(g)(X) = \sum_{j \in J} \lambda_j \psi^j(g) e_j(g) = \sum_{j \in J} X^j(g) e_j(g) = U(g) \circ \Phi(X)$

$\Phi(X) = U(g^{-1})\left(\sum_{j \in J} X^j(g) e_j(g)\right)$

v) If the representations $\left(H_\Phi, \widehat{U}\right), (F, f)$ are equivalent (which happens if they have the same finite dimension) there is an isomorphism $\phi : H_\Phi \to F$ which can be defined by taking an orthonormal basis $(\widetilde{\varepsilon}_i (g_0))_{i\in J}, (f_j (g_0))_{j\in J}$ in each vector space, for some fixed $g_0 \in G$ that we can take $g_0 = 1 : \phi \left(\sum_{i\in J} \psi^j \widetilde{\varepsilon}_j (1)\right) = \sum_{i\in J} \psi^j f_j (1) \Leftrightarrow \phi (\widetilde{\varepsilon}_j (1)) = f_j (1)$

To a change of g corresponds a change of orthonormal basis, both in H_Φ and F, given by the known unitary map $f(g) :\ f_j (g) = f (g) (f_j (1)) = \sum_{k\in J} [f (g)]^k_j f_k (1)$ and thus we have the same matrix for $\widehat{U} (g) :$

$\widetilde{\varepsilon}_j (g) = \widehat{U} (g) (\widetilde{\varepsilon}_j (1)) = \phi^{-1} \circ f (g) \circ \phi (\widetilde{\varepsilon}_j (1)) = \phi^{-1} \circ f (g) (f_j (1)) = \sum_{k\in J} [f (g)]^k_j \widetilde{\varepsilon}_k (1)$

$$\begin{array}{ccccccc}
V & \xrightarrow{U(g)} & V & \xrightarrow{\Phi} & W & & \\
\downarrow \Upsilon & & \downarrow \Upsilon & & \downarrow \Upsilon & & \\
H & \xrightarrow{\widehat{U}(g)} & H & \xrightarrow{\widehat{\Phi}} & H_\Phi & \xrightarrow{\widehat{U}(g)} & H_\Phi \\
 & & & & \downarrow \phi & & \downarrow \phi \\
 & & & & F & \xrightarrow{f(g)} & F
\end{array}$$

$\widetilde{\varepsilon}_j (g) = \widehat{U} (g) (\widetilde{\varepsilon}_j (1)) = \sum_{k\in J} [f (g)]^k_j \widetilde{\varepsilon}_k (1)$

$e_j (g) = \Upsilon^{-1} (\widetilde{\varepsilon}_j (g)) = \Upsilon^{-1} \left(\sum_{k\in J} [f (g)]^k_j \widetilde{\varepsilon}_k (1)\right)$

$= \sum_{k\in J} [f (g)]^k_j \Upsilon^{-1} (\widetilde{\varepsilon}_k (1)) = \sum_{k\in J} [f (g)]^k_j e_k (1)$

$e_j (g) = \Upsilon^{-1} \circ \widehat{U} (g) \circ \Upsilon (e_j (1)) = U (g) (e_j (1))$

Thus the matrix of $U(g)$ to go from 1 to g is $[f (g)]$

$\Phi (X) = U \left(g^{-1}\right) \left(\sum_{j\in J} X^j (g) e_j (g)\right)$

$\Phi \circ U (g) (X) = \sum_{j\in J} X^j (g) e_j (g) = \sum_{j\in J} X^j (g) \sum_{k\in J} \left[f \left(g^{-1}\right)\right]^k_j e_k (1)$

$\Phi \circ U (1) (X) = \sum_{k\in J} X^k (1) e_k (1) \Rightarrow \sum_{j\in J} X^j (g) \left[f \left(g^{-1}\right)\right]^k_j = X^k (1)$

$X^j (g) = \sum_{k\in J} [f (g)]^k_j X^j (1)$

The measures $\Phi \circ U (g) (X)$ transform with the known matrix $f (g)$.

vi) $\left(H_\Phi, \widehat{U}'(1)\right), (F, f' (1))$ are equivalent, anti-symmetric (or anti-hermitian for complex vector spaces) representations of the Lie algebra T_1G. If $(\kappa_a)_{a=1}^m$ is a basis of T_1G then $f' (1)$, which is a linear map, is defined by the values of $f'(1) (\kappa_a) \in L (F; F)$.

$$\begin{array}{ccc}
H_\Phi & \xrightarrow{\widehat{U}'(1)(\kappa)} & H_\Phi \\
\phi \downarrow & & \phi \downarrow \\
F & \xrightarrow{f'(1)(\kappa)} & F
\end{array}$$

$\widehat{U}' (1) (\kappa) (\psi) = \phi^{-1} \circ f' (1) (\kappa) \circ \phi (\psi)$

If we know the values of the action of $f' (1) (\kappa_a)$ on any orthonormal basis $(f_j)_{j\in J}$ of F :

$f' (1) (\kappa_a) (f_j) = \sum_{k\in J} [K_a]^k_j f_k$

we have the value of $\widehat{U}' (1) (\kappa_a)$ for the corresponding orthonormal basis $(\widehat{\varepsilon}_j)_{j\in J}$ of H_Φ

$\widehat{U}' (1) (\kappa_a) (\widehat{\varepsilon}_j) = \widehat{U}' (1) (\kappa_a) \phi^{-1} (f_j) = \phi^{-1} \circ f' (1) (\kappa_a) (f_j)$

$= \phi^{-1}\left(\sum_{k\in J}[K_a]_j^k f_k\right) = \sum_{k\in J}[K_a]_j^k \widehat{\varepsilon}_k$

So $\widehat{U}'(1)$ is represented in an orthonormal basis of H_Φ by the same matrices $[K_a]$

And similarly :

$\widehat{U}(g) = \Upsilon \circ U(g) \circ \Upsilon^{-1} \Rightarrow \widehat{U}'(1)(\kappa) = \Upsilon \circ U'(1)(\kappa) \circ \Upsilon^{-1}$

$U'(1)(\kappa_a)(e_j) = \Upsilon \circ U'(1)(\kappa_a) \circ \Upsilon^{-1}(e_j) = \Upsilon \circ U'(1)(\kappa_a)(\widehat{\varepsilon}_j)$

$= \Upsilon\left(\sum_{k\in J}[K_a]_j^k \widehat{\varepsilon}_k\right) = \sum_{k\in J}[K_a]_j^k e_k$

vii) Because $\Phi \circ U(g) = U(g) \circ \Phi \Rightarrow \Phi \circ U'(1)(\kappa_a) = U'(1)(\kappa_a) \circ \Phi$:

$\Phi \circ U'(1)(\kappa_a)(e_j) = \sum_{k\in J}[K_a]_j^k \Phi(e_k)$ ■

This result is specially important in Physics. Any unitary representation of a compact or finite group is reducible in the sum of orthogonal, finite dimensional, irreducible unitary representations. As a consequence the space V of the variables X has the same structure. If, as it can be assumed, the state of the system stays in the same irreducible representation, it can belong only to some specific finite dimensional spaces, defined through the representation or an equivalent representation of G. X depends only on a finite number of parameters, This is the starting point of quantization.

Notice that the nature of the space E does not matter, only the matrices $[f(g)], [K]$.

Usually in Physics the changes are not parametrized by the group, but by a vector of the Lie algebra (for instance rotations are not parametrized by a matrix but by a vector representing the rotation), which gives a special interest to the two last results.

The usual geometric representations, based on frames defined through a point and a set of vectors, such as in Galilean Geometry and Special Relativity, have been generalized by the formalism of fiber bundles, which encompasses also General Relativity, and is the foundation of gauge theories. Gauge theories use abundantly group transformations, so they are a domain of choice to implement the previous results.

Fourier transform

If G is an abelian group we have more. Irreducible representations of abelian groups are unidimensional, and any unitary representation of an abelian group is the sum of projections on unidimensional vector subspaces which, for infinite dimensional representations, takes the form of spectral integrals. More precisely, there is a bijective correspondence between the unitary representation of an abelian group G and the spectral measures on the Pontryagin dual $\widehat{G}$, which is the space of continuous maps : $\vartheta : G \to T$ where T is the set of complex numbers of module 1 (Maths.1932). This can be made less abstract if G is a topological, locally compact group. Then it has a Haar measure μ and the representation $\left(H, \widehat{U}\right)$ is equivalent to $\left(L^2(G,\mu,\mathbb{C}), \mathcal{F}\right)$ that is to the Fourier transform $\mathcal{F}$ on complex valued, square integrable, functions on G (Maths.2421).

If $\varphi \in L^2(G,\mu,\mathbb{C}) \cap L^1(G,\mu,\mathbb{C})$:

$\mathcal{F}(\varphi)(\vartheta) = \int_G \varphi(g)\overline{\vartheta(g)}\mu(g)$

$\mathcal{F}^*(h)(g) = \int_{\widehat{G}} h(\vartheta)\vartheta(g)\nu(\vartheta)$ for a unique Haar measure ν on $\widehat{G}$ and $\mathcal{F}^* = \mathcal{F}^{-1}$

If G is a compact group then we have Fourier series on a space of periodic functions, and if G is a non compact, finite dimensional Lie group, G is isomorphic to some vector space E and we have the usual Fourier transform on functions on E.

These cases are important from a practical point of view as it is possible to replace the abstract Hilbert space H by more familiar spaces of functions, and usually one can assume that the space V is itself some Hilbert space. The previous tools (observables,...) are then directly available.

The most usual application is about periodic phenomena : whenever a system is inclosed in some box, it can be usually assumed that the variables are periodic (and null out of the box). Then the representation is naturally through Fourier series and we have convenient Hilbert bases.

One parameter groups

An important case, related to the previous one, is when the variables X depend on a scalar real argument, and the model is such that $X(t), X'(t') = X(t+\theta)$, with any fixed θ, represent the same state. The associated operator is parametrized by a scalar and we have a map :
$\widehat{U} : \mathbb{R}_+ \to G\mathcal{L}(H,H)$ such that :
$\widehat{U}(t+t') = \widehat{U}(t) \circ \widehat{U}(t')$
$\widehat{U}(0) = Id$

Then we have a one parameter semi-group. If moreover the map $\widehat{U}$ is strongly continuous (that is $\lim_{\theta\to 0}\left\|\widehat{U}(\theta) - Id\right\| = 0$), it can be extended to $\mathbb{R}$. $\left(\widehat{U}, H\right)$ is a unitary representation of the abelian group $(\mathbb{R}, +)$. We have a one parameter group, and because $\widehat{U}$ is a continuous Lie group morphism it is differentiable with respect to θ (Maths.1784).

Any strongly continuous one parameter group of operators on a Banach vector space admits an infinitesimal generator $S \in \mathcal{L}(H;H)$ such that : $\widehat{U}(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} S^n = \exp tS$ (Maths.1033). By derivation with respect to t we get : $\frac{d}{ds}\widehat{U}(s)|_{t=s} = (\exp tS) \circ S \Rightarrow S = \frac{d}{ds}\widehat{U}(s)|_{t=0}$

Because $\widehat{U}(t)$ is unitary S is anti-hermitian :
$\left\langle \widehat{U}(t)\psi, \widehat{U}(t)\psi' \right\rangle_H = \langle \psi, \psi' \rangle_H$
$\Rightarrow \left\langle \frac{d}{dt}\widehat{U}(t)\psi, \widehat{U}(t)\psi' \right\rangle_H + \left\langle \widehat{U}(t)\psi, \frac{d}{dt}\widehat{U}(t)\psi' \right\rangle_H = 0 \Rightarrow S = -S^*$
S is normal and has a spectral resolution P :
$S = \int_{Sp(S)} sP(s)$

S is anti-hermitian so its eigen-values are pure imaginary : $\lambda = -\overline{\lambda}$. $\widehat{U}(t)$ is not compact and S is not compact, usually its spectrum is continuous, so it is not associated to any observable.

2.4.6 Extension to manifolds

Several extensions of the theorem 2 can be considered. On frequent case is the following. In a model variables X are maps defined on a manifold M, valued in a fixed vector space, and belong to a space V of maps with the required properties. But a variable Y is defined through X : $Y(m) = f(X(m))$ and belongs to a manifold $N(X)$ depending on X. So the conditions 1 do not apply.

To address this kind of problem we need to adapt our point of view. We have seen the full mathematical definition of a manifold in the first section. A manifold M is a class of equivalence : the same point m of M can be defined by several charts, maps $\varphi : E \to M$ from a vector space E to M, with different coordinates : $m = \varphi_a(\xi_a) = \varphi_b(\xi_b)$ so that it defines classes of equivalence between sets of coordinates : $\xi_a \sim \xi_b \Leftrightarrow \varphi_a(\xi_a) = \varphi_b(\xi_b)$. These classes of equivalence are made clear by the transitions maps $\chi_{ba} : E \to E$, which are bijective : $\xi_a \sim \xi_b \Leftrightarrow \xi_b = \chi_{ba}(\xi_a)$. And these transitions maps are the key characteristic of the manifold. To a point m of M corresponds a class of equivalence of coordinates and one can conceive that to each value of Y is associated a specific class of equivalence.

So let us consider a system represented by a model which meets the following general properties :

Condition 23 *The model is comprised of :*

i) A finite number of variables, collectively denoted X, which are maps valued in a vector space E and meeting the conditions 1 : they belong to an open subset O of a separable, infinite dimensional Fréchet space V.

ii) A variable Y, valued in a set F, defined by a map :

$f : O \to F :: Y = f(X)$

iii) A collection of linear continuous bijective maps $\mathfrak{U} = (U_a \in G\mathcal{L}(V;V))_{a \in A}$, comprising the identity, closed under composition : $\forall a, b \in A : U_a \circ U_b \in \mathfrak{U}$

iv) On V and F the equivalence relation :

$R : X \sim X' \Leftrightarrow \exists a \in A : X' = U_a(X) : f(X) = f(X')$

The conditions iii) will be usually met by the action of a group : $U_a(X) = \lambda(a, X)$.

Denote the set $N = \{Y = f(X), X \in O\}$. The quotient set : N/R is comprised of classes of equivalence of points Y which can be defined by related coordinates. This is a manifold, which can be discrete and comprising only a finite number of points. One can also see the classes of equivalence of N/R as representing states of the system, defined equivalently by the variable $X, X' = U_a(X)$.

Notice that f is unique, no condition is required on E other than to be a vector space, and nothing on F. Usually the maps U_a are defined by : $U_a(X) = \chi_a \circ X$ where the maps $\chi_a \in GL(E;E)$ are bijective on E (not F or V) but only the continuity of U_a can be defined.

We have the following result :

Theorem 24 *For a system represented by a model meeting the conditions 23 :*

i) V can be embedded as an open of a Hilbert space H with a linear isometry $\Upsilon : V \to H$, to each U_a is associated the unitary operator $\widehat{U}_a = \Upsilon \circ U_a \circ \Upsilon^{-1}$ on H, each class of equivalence $[V]_y$ of R on V is associated to a class of equivalence $[H]_y$ in H of :

$\widehat{R} : \psi \sim \psi' \Leftrightarrow \exists a \in A : \psi' = \widehat{U}_a(\psi)$. $[V]_y$ is a partition of V and $[H]_y$ of H.

ii) If (V, U) is a representation of a Lie group G, then $\left(H, \widehat{U}\right)$ is a unitary representation of G and each $[H]_y$ is invariant by the action of G.

Proof. i) R defines a partition of V, we can label each class of equivalence by the value of Y, and pick one element X_y in each class :

$[V]_y = \{X \in O : f(X) \sim f(X_y) = y\} \equiv \{X \in O : \exists a \in A : X = U_a(X_y)\}$

$\equiv \{X \in O : X = U_a(X_y), a \in A\}$

The variables X meet the conditions 1, O can be embedded as an open of a Hilbert space H and there is linear isomorphism : $\Upsilon : V \to H$

In $[V]_y$ the variables $X, X' = U_a(X)$ define the same state and we can implement the theorem 21. $\widehat{U}_a = \Upsilon \circ U_a \circ \Upsilon^{-1}$ is an unitary operator on H

$\forall X \in [V]_y : \widehat{U}_a \circ \Upsilon(X_y) = \Upsilon \circ U_a(X_y) = \Upsilon(X)$

The set $[H]_y = \Upsilon\left([V]_y\right) = \left\{\psi \in H : \psi = \widehat{U}_a(\Upsilon(X_y)), a \in A\right\}$ is the class of equivalence of :

$\widehat{R} : \psi \sim \psi' \Leftrightarrow \exists a \in A : \psi' = \widehat{U}_a(\psi)$

R defines a partition of V : $V = \cup_y [V]_y$ and $\widehat{R}$ defines a partition of H : $H = \cup_y [H]_y$

ii) If (V, U) is a representation of a Lie group G then $[V]_y$ is the orbit of X_y, $\left(H, \widehat{U}\right)$ is a unitary representation of G

Each $[H]_y$ is invariant by G. The vector subspace $[F]_y$ spanned by $[H]_y$ is invariant by G, so $\left([F]_y, \widehat{U}\right)$ is a representation of G. ■

As a consequence of the last result : if U is a compact group, then the representation $\left(H, \widehat{U}\right)$ is the sum of irreducible, orthogonal, finite dimensional representations. For each value of Y the subset $[H]_y$ is invariant by the action of G, so it must belong to one of the irreducible representations, as well as $[F]_y$. The maps X, for a given value of Y, belong to a finite dimensional vector space, and depend on a finite number of parameters. This is the usual meaning of the quantization of X.

2.5 THE EVOLUTION OF THE SYSTEM

In many models involving maps, the variables X_k are functions of the time t, which represents the evolution of the system. So this is a privileged argument of the functions. So far we have not made any additional assumption about the model : the open Ω of the Hilbert space contains all the possible values but, due to the laws to which it is subjected, only some solutions will emerge, depending on the initial conditions. They are fixed by the value $X(0)$ of the variables at some origin 0 of time. They are specific to each realization of the system, but we should expect that the model and the laws provide a general solution, that is a map : $X(0) \to X$ which determines X for each specific occurrence of $X(0)$. It will happen if the laws are determinist. One says that the problem is well posed if for any initial conditions there is a unique solution X, and that X depends continuously on $X(0)$. We give a more precise meaning of determinism by enlarging the conditions 1 as follows :

Condition 25 *: The model representing the system meets the conditions 1. Moreover :*
i) V is an infinite dimensional separable Fréchet space V of maps : $X = (X_k)_{k=1}^N :: R \to E$ where R is an open subset of $\mathbb{R}$ and E a normed vector space
ii) $\forall t \in R$ the evaluation map : $\mathcal{E}(t) : V \to E : \mathcal{E}(t) X = X(t)$ is continuous
The laws for the evolution of the system are such that the variables $(X_k)_{k=1}^N$, which define the possible states considered for the system (that we call the admissible states) meet the conditions :
iii) The initial state of the system, defined at $t = 0 \in R$, belongs to an open subset A of E
iv) For any solutions X, X' belonging to O if the set $\varpi = \{t \in R : X(t) = X'(t)\}$ has a non null Lebesgue measure then $X = X'$.

The last condition iv) means that the system is semi determinist : to the same initial conditions can correspond several different solutions, but if two solutions are equal on some interval then they are equal almost everywhere.

The condition ii) is rather technical and should be usually met. Practically it involves some relation between the semi-norms on V and the norm on E (this is why we need a norm on E) : when two variables X, X' are close in V, then their values $X(t), X'(t)$ must be close for almost all t. More precisely, because $\mathcal{E}(t)$ is linear, the continuity can be checked at $X = 0$ and reads:
$\forall t \in R, \forall X \in O : \forall \varepsilon > 0, \exists \eta : d(X,0)_V < \eta \Rightarrow \|X(t)\|_E < \varepsilon$ where d is the metric on V
In all usual cases (such as L^p spaces or spaces of differentiable functions, Maths.2282) $d(X,0)_V \to 0 \Rightarrow \forall t \in R : \|X(t)\|_E \to 0$ and the condition ii) is met, but this is not a general result.
This condition is met if : $\forall t \neq t' \in R, \exists X \in V : X(t) \neq X(t')$
Proof. The family of maps X is separating, the weak topology (also called initial topology) on V induced by the family of maps X is Hausdorff. Then $d(X,0)_V = 0 \Rightarrow \|X(t)\|_E = 0$. (Maths.10.2.3). ■
Notice that :
- the variables X can depend on any other arguments besides t as previously
- E can be infinite dimensional but must be normed
- no continuity condition is imposed on X.

2.5.1 Fundamental theorems for the evolution of a system

If the model meets the conditions 25 then it meets the conditions 1 : there is a separable, infinite dimensional, Hilbert space H, defined up to isomorphism, such that the states (admissible or not) $\mathcal{S}$ belonging to O can be embedded as an open subset $\Omega \subset H$ which contains 0 and a convex subset. Moreover to any basis of V is associated a bijective linear map $\Upsilon : V \to H$.

Theorem 26 *If the conditions 25 are met, then there are :*

i) a Hilbert space F*, an open subset* $\widetilde{A} \subset F$

ii) a map : $\Theta : R \to \mathcal{L}(F;F)$ *such that* $\Theta(t)$ *is unitary and, for the admissible states* $X \in O \subset V$ *:*

$X(0) \in \widetilde{A} \subset F$

$\forall t : X(t) = \Theta(t)(X(0)) \in F$

iii) for each value of t *an isometry :* $\widehat{\mathcal{E}}(t) \in \mathcal{L}(H;F)$ *such that for the admissible states* $X \in O \subset V$ *:*

$\forall X \in O : \widehat{\mathcal{E}}(t)\Upsilon(X) = X(t)$

where H *is the Hilbert space and* Υ *is the linear chart associated to* X *and any basis of* V

Proof. i) Define the equivalence relation on V :

$\mathcal{R} : X \sim X' \Leftrightarrow X(t) = X'(t)$ for almost every $t \in R$

and take the quotient space $V/\mathcal{R}$, then the set of admissible states is a set $\widetilde{O}$ such that :

$\widetilde{O} \in O \subset V$

$\forall X \in \widetilde{O} : X(0) \in A$

$\forall X, X' \in \widetilde{O}, \forall t \in R : X(t) = X'(t) \Rightarrow X = X'$

ii) Define :

$\forall t \in R : \widetilde{F}(t) = \left\{X(t), X \in \widetilde{O}\right\}$ thus $\widetilde{F}(0) = A$

A is a subset of E. There are families of independent vectors belonging to A, and a largest family $(f_j)_{j\in J}$ of independent vectors. It generates a vector space $F(0)$ which is a vector subspace of E, containing A.

$\forall u \in F(0) : \exists (x_j)_{j\in J} \in \mathbb{R}_0^J : u = \sum_{j\in J} x_j f_j$

The map :

$\widetilde{\Theta}(t) : \widetilde{F}(0) \to \widetilde{F}(t) :: \widetilde{\Theta}(t)u = \mathcal{E}(t) \circ \mathcal{E}(0)^{-1} u$

is bijective and continuous

The set $F(t) = \widetilde{\Theta}(t) F(0) \subset E$ is well defined by linearity :

$\widetilde{\Theta}(t)\left(\sum_{j\in J} x_j f_j\right) = \sum_{j\in J} x_j \widetilde{\Theta}(t)(f_j)$

The map : $\widetilde{\Theta}(t) : F(0) \to F(t)$ is linear, bijective, continuous on an open subset A, thus continuous, and the spaces $F(t)$ are isomorphic, vector subspaces of E, containing $\widetilde{F}(t)$.

Define : $(\varphi_j)_{j\in J}$ the largest family of independent vectors of

$\left\{\widetilde{\Theta}(t)(f_j), t \in R\right\}$. This is a family of independent vectors of E, which generates a subspace $\widetilde{F}$ of E, containing each of the $F(t)$ and thus each of the $\widetilde{F}(t)$. Moreover each of the φ_j is the image of a unique vector f_j for some $t_j \in R$.

The map $\widetilde{\Theta}(t)$ is then a continuous linear map $\widetilde{\Theta}(t) \in \mathcal{L}\left(\widetilde{F};\widetilde{F}\right)$

iii) The conditions of proposition 1 are met for O and V, so there are a Hilbert space H and a linear map : $\Upsilon : O \to \Omega$

Each of the φ_j is the image of a unique vector f_j for some $t \in R$,and thus there is a uniquely defined family $(X_j)_{j\in J}$ of $\widetilde{O}$ such that $X_j(t_j) = \varphi_j$.

Define on $\widetilde{F}$ the bilinear symmetric definite positive form with coefficients :

$\langle \varphi_j, \varphi_k \rangle_{\widetilde{F}} = K_V\left(\mathcal{E}(t_j)^{-1}\varphi_j, \mathcal{E}(t_k)^{-1}\varphi_k\right)$

$= \left\langle \Upsilon\mathcal{E}(t_j)^{-1}\varphi_j, \Upsilon\mathcal{E}(t_k)^{-1}\varphi_k\right\rangle_H = \langle X_j, X_k\rangle_H$

By the Gram-Schmidt procedure we can build an orthonormal basis $(\widetilde{\varphi}_j)_{j\in J}$ of $\widetilde{F}$: $\widetilde{F} = Span(\widetilde{\varphi}_j)_{j\in J}$ and the Hilbert vector space :

$F=\left\{\sum_{j\in J}\widetilde{x}_j\widetilde{\varphi}_j,(\widetilde{x}_j)_{j\in J}\in\ell^2(J)\right\}$ which is a vector space containing $\widetilde{F}$ (but is not necessarily contained in E).

iv) The map : $\widetilde{\Theta}(t)\in\mathcal{L}\left(\widetilde{F};\widetilde{F}\right)$ is a linear homomorphism, $\widetilde{F}$ is dense in F, thus $\widetilde{\Theta}(t)$ can be extended to a continuous operator $\Theta(t)\in\mathcal{L}(F;F)$ (Math.1003).

$\widetilde{\Theta}(t)$ is unitary on $\widetilde{F}$: $\langle u,v\rangle_{\widetilde{F}}=K_V\left(\mathcal{E}(0)^{-1}u,\mathcal{E}(0)^{-1}v\right)$ so $\Theta(t)$ is unitary on F.

iv) Define the map :

$\widehat{\mathcal{E}}(t):\Omega\to F::\widehat{\mathcal{E}}(t)\Upsilon(X)=X(t)$

where $\Omega\subset H$ is the open associated to V and O.

For $X\in\widetilde{O}$:

$\widehat{\mathcal{E}}(t)\Upsilon(X)=X(t)=\widetilde{\Theta}(t)X=\mathcal{E}(t)\circ\mathcal{E}(0)^{-1}X$

$\widehat{\mathcal{E}}(t)=\mathcal{E}(t)\circ\mathcal{E}(0)^{-1}\circ\Upsilon^{-1}$

$\widehat{\mathcal{E}}(t)$ is linear, continuous, bijective on Ω, it is an isometry :

$\left\langle\widehat{\mathcal{E}}(t)\psi,\widehat{\mathcal{E}}(t)\psi'\right\rangle_F=\langle X(t),X'(t)\rangle_F=\langle\Upsilon X,\Upsilon X'\rangle_H=\langle\psi,\psi'\rangle_H$

v) $A=\widetilde{F}(0)$ is an open subset of $F(0)$, which is itself an open vector subspace of F. Thus A can be embedded as an open subset $\widetilde{A}$ of F. ■

The key point in the proof is the property :

"The map : $\widetilde{\Theta}(t):\widetilde{F}(0)\to\widetilde{F}(t)::\widetilde{\Theta}(t)u=\mathcal{E}(t)\circ\mathcal{E}(0)^{-1}u$ is bijective and continuous"

which is easily understood when t is the only variable, then it means that laws for the evolution of the system are such that the initial value $X(0)$ defines, up to a negligible set of points, uniquely $X(t)$.

But it is more difficult to meet when other arguments than t are involved. Let $X(t,x)$ with x other (possible multiple) arguments. Then the sets $\widetilde{F}(t)=\left\{X(t),X\in\widetilde{O}\right\}$ of the values taken by X depend on t, but also x, and shall be interpreted as $\widetilde{F}(t,x)=\left\{X(t,x),X\in\widetilde{O}\right\}$ for a fixed, value of x. Then the evaluation map is bijective, for a given, fixed, value of x. And the operator $\Theta(t)$ acts on the map $X_x:R\to X_x(t)=X(t,x)$ that is : $X_x(t)=\Theta(t)X_x(0)$.

As a consequence the model is determinist, up to the equivalence between maps almost everywhere equal. But the operator $\Theta(t)$ depends on t and not necessarily continuously, so the problem is not necessarily well posed. Notice that each solution $X(t)$ belong to E, but the Hilbert space F can be larger than E. Moreover the result holds if the conditions apply to some variables only.

But we have a stronger result.

Theorem 27 *If the model representing the system meets the conditions 1 and moreover :*

i) V is an infinite dimensional separable Fréchet space V of maps : $X=(X_k)_{k=1}^N::R\to E$ *where E is a normed vector space*

ii) $\forall t\in\mathbb{R}$ *the evaluation map :* $\mathcal{E}(t):V\to E:\mathcal{E}(t)X=X(t)$ *is continuous*

iii) the variables $X_k'(t)=X_k(t+\theta)$ *and* $X_k(t)$ *represent the same state of the system, for any* $t'=t+\theta$ *with a fixed* $\theta\in\mathbb{R}$

then :

i) there is a continuous map $S\in\mathcal{L}(V;V)$ *such that :*

$\mathcal{E}(t)=\mathcal{E}(0)\circ\exp tS$

$\forall t\in\mathbb{R}:X(t)=(\exp tS\circ X)(0)=\left(\sum_{n=0}^{\infty}\frac{t^n}{n!}S^nX\right)(0)$

and the operator $\widehat{S}=\Upsilon\circ S\circ\Upsilon^{-1}$ *associated to S is anti-hermitian* *ii) there are a Hilbert space F, an open $\widetilde{A} \subset F$, a continuous anti-hermitian map $\widetilde{S} \in \mathcal{L}(F;F)$ such that :*

$\forall X \in O \subset V : X(0) \in \widetilde{A} \subset F$

$\forall t : X(t) = \left(\exp t\widetilde{S}\right)(X(0)) \in F$

iii) The maps X are smooth and :

$\frac{d}{ds}X(s)|_{s=t} = \widetilde{S}X(t)$

Proof. i) We have a change of variables U depending on a parameter $\theta \in \mathbb{R}$ which reads with the evaluation map : $\mathcal{E} : \mathbb{R} \times V \to F :: \mathcal{E}(t)X = X(t)$:

$\forall t, \theta \in \mathbb{R} : \mathcal{E}(t)(U(\theta)X) = \mathcal{E}(t+\theta)(X)$

$\Leftrightarrow \mathcal{E}(t)U(\theta) = \mathcal{E}(t+\theta) = \mathcal{E}(\theta)U(t)$:

U defines a one parameter group of linear operators:

$U(\theta+\theta')X(t) = X(t+\theta+\theta') = U(\theta)\circ U(\theta')X(t)$

$U(0)X(t) = X(t)$

It is obviously continuous at $\theta = 0$ so it is continuous.

ii) The conditions 1 are met, so there are a Hilbert space H, a linear chart Υ, and $\widehat{U} : \mathbb{R} \to \mathcal{L}(H;H)$ such that $\widehat{U}(\theta)$ is linear, bijective, unitary :

$\forall X \in O : \widehat{U}(\theta)(\Upsilon(X)) = \Upsilon(U(\theta)(X))$

$\widehat{U}(\theta+\theta') = \Upsilon \circ U(\theta+\theta') \circ \Upsilon^{-1} = \Upsilon \circ U(\theta) \circ U(\theta') \circ \Upsilon^{-1} = \Upsilon \circ U(\theta) \circ \Upsilon^{-1} \circ \Upsilon \circ U(\theta') \circ \Upsilon^{-1} = \widehat{U}(\theta) \circ \widehat{U}(\theta')$

$\widehat{U}(0) = \Upsilon \circ U(0) \circ \Upsilon^{-1} = Id$

The map : $\widehat{U} : \mathbb{R} \to \mathcal{L}(H;H)$ is uniformly continuous with respect to θ, it defines a one parameter group of unitary operators. So there is an anti-hermitian operator $\widehat{S}$ with spectral resolution P such that :

$\widehat{U}(\theta) = \sum_{n=0}^{\infty} \frac{\theta^n}{n!}\widehat{S}^n = \exp\theta\widehat{S}$

$\frac{d}{ds}\widehat{U}(s)|_{\theta=s} = \left(\exp\theta\widehat{S}\right) \circ \widehat{S}$

$\widehat{S} = \int_{Sp(S)} sP(s)$

$\left\|\widehat{U}(\theta)\right\| = 1 \leq \exp\left\|\theta\widehat{S}\right\|$

iii) $S = \Upsilon^{-1} \circ \widehat{S} \circ \Upsilon$ is a continuous map on the largest vector subspace V_0 of V which contains O, which is a normed vector space with the norm induced by the positive kernel.

$\|S\| \leq \left\|\Upsilon^{-1}\right\| \left\|\widehat{S}\right\| \|\Upsilon\| = \left\|\widehat{S}\right\|$ because Υ is an isometry.

So the series $\sum_{n=0}^{\infty} \frac{\theta^n}{n!}S^n$ converges in V_0 and :

$U(\theta) = \Upsilon^{-1} \circ \widehat{U}(\theta) \circ \Upsilon = \sum_{n=0}^{\infty} \frac{\theta^n}{n!}S^n = \exp\theta S$

$\forall \theta, t \in \mathbb{R} : U(\theta)X(t) = X(t+\theta) = (\exp\theta S)X(t)$

$\mathcal{E}(t)\exp\theta S = \mathcal{E}(t+\theta)$

Exchange θ, t and take $\theta = 0$:

$\mathcal{E}(\theta)\exp tS = \mathcal{E}(t+\theta)$

$\mathcal{E}(0)\exp tS = \mathcal{E}(t) \in \mathcal{L}(V;E)$

which reads :

$\forall t \in \mathbb{R} : U(t)X(0) = X(t) = (\exp tS)X(0)$

(U, V_0) is a continuous representation of $(\mathbb{R}, +)$, U is smooth and X is smooth :

$\frac{d}{ds}U(s)X(0)|_{s=t} = \frac{d}{ds}X(s)|_{s=t} = SX(t)$

$\Leftrightarrow \frac{d}{ds}\mathcal{E}(s)|_{s=t} = S\mathcal{E}(t)$

The same result holds whatever the size of O in V, so S is defined over V.

iv) The set : $F(t) = \{X(t), X \in V\}$ is a vector subspace of E.

Each map is fully defined by its value at one point :
$\forall t\in\mathbb{R}: X(t)=(\exp tS\circ X)(0)$
$X(t)=X'(t)\Rightarrow\forall\theta: X(t+\theta)=X'(t+\theta)\Leftrightarrow X=X'$
So the conditions 4 are met.
$\Theta(t): F(0)\to F(t) :: \Theta(t)u=\mathcal{E}(t)\circ\mathcal{E}(0)^{-1}u=\mathcal{E}(0)\circ\exp tS\circ\mathcal{E}(0)^{-1}u$
The map $\Theta(\theta): F\to F$ defines a one parameter group, so it has an infinitesimal generator $\widetilde{S}\in\mathcal{L}(F;F):\Theta(\theta)=\exp\theta\widetilde{S}$ and because $\Theta(\theta)$ is unitary $\widetilde{S}$ is anti-hermitian.
$\frac{d}{ds}\Theta(s)X(0)|_{s=t}=\frac{d}{ds}X(s)|_{s=t}=\widetilde{S}X(t)$ ∎

As a consequence such a model is necessarily determinist, and the system is represented by smooth maps whose evolution is given by a unique operator. It is clear that the conditions 25 are then met, so this case is actually a special case of the previous one. Notice that, even if X was not assumed to be continuous, smoothness is a necessary result. This result can seem surprising, but actually the basic assumption about a translation in time means that the laws of evolution are smooth, and as a consequence the variables depend smoothly on the time. And conversely this implies that, whenever there is some discontinuity in the evolution of the system, the conditions above cannot hold : time has a specific meaning, related to a change in the environment.

Comments

The conditions above depend deeply on how the time is understood in the model. We have roughly two cases :

A) t is a parameter used only to identify a temporal location. In Galilean Geometry the time is independent from the spatial coordinates for any observer and one can consider a change of coordinates such as : $t'=t+\theta$ with any constant θ. The variables X,X' such that $X'(t')=X(t+\theta)$ represent the same system. Similarly in Relativist Geometry the universe can be modelled as a manifold, and a change of coordinates with affine parameters, $\xi'=\xi+\theta$ with a fixed 4 vector θ, is a change of charts. The components of any quantity defined on the tensorial tangent bundle change according to the jacobian $\left[\frac{\partial\xi'}{\partial\xi}\right]$ which is the identity, so the corresponding variables represent the same system. Then we are usually in the conditions of the Theorem 27, and this is the basis of the Schrödinger equation.

B) t is a parameter used to measure the duration of a phenomenon, usually the time elapsed since some specific event, and it is clear that the origin of time matters and the variables X,X' such that $X'(t')=X(t+\theta)$ do not represent the same system. This is the case in more specific models, such as in Engineering. The proposition 27 does not hold, but the proposition 26 holds if the model is determinist.

The conditions 25 require at least that all the variables which are deemed significant are accounted for. Usually probabilist laws appear because some of them are missing. The Theorem 26 precise this issue : by denoting the missing variables Y, one needs to enlarge the vector space E, and similarly F. The map $\Theta(t)$ still exists, but it encompasses the couples $(X(t),Y(t))$. The dispersion of the observed values of $X(t)$ are then imputed to the distribution of the unknown values $Y(t)$.

It seems strange that a law for the evolution of the system can appear without any hypotheses about the mechanisms at play in this evolution. Actually the theorems do not provide the laws of evolution - they assume that they exist, in the form of semi-determinism - they only precise their specification. The existence of laws (in the form of the maps X) encompassing the whole of the period under review has the effect that going from one state of the system at a given time to the state at another time is like a change of observer, and this is obvious in the second theorem. Then the change of the time parameter is an operation which is done on a given set of states, which are assumed to exist. But of course this assumption is critical.

2.5.2 Observables

When a system is studied through its evolution, the observables can be considered from two different points of view :

- in the movie way : the estimation of the parameters is done at the end of the period considered, from a batch of data corresponding to several times (which are not necessarily the same for all variables). So this is the map X which is estimated through an observable $X \to \Phi(X)$.

- in the picture way : the estimation is done at different times (the same for all the variables which are measured). So there are the values $X(t)$ which are estimated. Then the estimation of $X(t)$ is given by $\varphi(X(t)) = \varphi(\mathcal{E}(t) X)$, with φ a linear map from E to a finite dimensional vector space, which usually does not depend on t (the specification stays the same).

In the best scenario the two methods should give the same result, which reads :

$\varphi(\mathcal{E}(t) X) = \mathcal{E}(t)(\Phi X) \Leftrightarrow \varphi = \mathcal{E}(t) \circ \Phi \circ \mathcal{E}(t)^{-1}$

But usually, when it is possible, the first way gives a better statistical estimation.

2.5.3 Phases Transitions

There is a large class of problems which involve transitions in the evolution of a system. They do not involve the maps X, which belong to the same family as above, but the values $X(t)$ which are taken over a period of time in some vector space E. There are distinct subsets of E, that we will call **phases** (to avoid any confusion with states which involves the map X), between which the state of the system goes during its evolution, such as the transition solid / gas or between magnetic states. The questions which arise are then : what are the conditions, about the initial conditions or the maps X, for the occurrence of such an event ? Can we forecast the time at which such event takes place ?

Staying in the general model meeting the conditions 25, the first issue is the definition of the phases. The general idea is that they are significantly different states, and it can be formalized by : the set $\{X(t), t \in R, X \in O\}$ is disconnected, it comprises two disjoint subsets E_1, E_2 closed in E.

If the maps $X : R \to F$ are continuous and R is an interval of $\mathbb{R}$ (as we will assume) then the image $X(R)$ is connected, the maps X cannot be continuous, and we cannot be in the conditions of proposition 27 (a fact which is interesting in itself), but we can be in the case of proposition 26. This is a difficult but also very common issue : in the real life such discontinuous evolutions are the rule. However, as we have seen, in the physical world discontinuities happen only at isolated points : the existence of a singularity is what makes interesting a change of phase. If the transition points are isolated, there is an open subset of R which contains each of them, a finite number of them in each compact subset of R, and at most a countable number of transition points. A given map X is then continuous (with respect to t) except in a set of points $(\theta_\alpha)_{\alpha \in A}$, $A \subset \mathbb{N}$. If $X(0) \in E_1$ then the odd transition points $\theta_{2\alpha+1}$ mark a transition $E_1 \to E_2$ and the opposite for the even points $\theta_{2\alpha}$.

If the conditions 25 are met then Θ is continuous except in $(\theta_\alpha)_{\alpha \in A}$, the transition points do not depend on the initial state $X(0)$, but the phase on each segment does. Then it is legitimate to assume that there is some probability law which rules the occurrence of a transition. We will consider two cases.

The simplest assumption is that the probability of the occurrence of a transition at any time t is constant. Then it depends only on the cumulated lengths of the periods $T_1 = \sum_{\alpha=0} [\theta_{2\alpha}, \theta_{2\alpha+1}], T_2 = \sum_{\alpha=0} [\theta_{2\alpha+1}, \theta_{2\alpha+2}]$ respectively.

Let us assume that $X(0) \in E_1$ then the changes $E_1 \to E_2$ occur for $t = \theta_{2\alpha+1}$, the probability of transitions read :

$\Pr(X(t+\varepsilon) \in E_2 | X(t) \in E_1) = \Pr(\exists \alpha \in \mathbb{N} : t + \varepsilon \in [\theta_{2\alpha+1}, \theta_{2\alpha+2}])$

$= T_2 / (T_1 + T_2)$

$\Pr(X(t+\varepsilon) \in E_1 | X(t) \in E_2) = \Pr(\exists \alpha \in \mathbb{N} : t + \varepsilon \in [\theta_{2\alpha}, \theta_{2\alpha+1}])$

$= T_1 / (T_1 + T_2)$

$\Pr(X(t) \in E_1) = T_1 / [R] ; \Pr(X(t) \in E_2) = T_2 / [R]$

The probability of a transition at t is : $T_2/(T_1+T_2) \times T_1/(T_1+T_2) + T_1/(T_1+T_2) \times T_2/(T_1+T_2) = 2T_1T_2/(T_1+T_2)^2$. It does not depend of the initial phase, and depends only on Θ. This probability law can be checked from a batch of data about the values of T_1, T_2 for each observed transition.

However usually the probability of a transition depends on the values of the variables. The phases are themselves characterized by the value of $X(t)$, so a sensible assumption is that the probability of a transition increases with the proximity of the other phase . Using the Hilbert space structure of F it is possible to address practically this case.

If E_1, E_2 are *closed convex subsets* of F, which is a Hilbert space, there is a unique map : $\pi_1 : F \to E_1$. The vector $\pi_1(x)$ is the unique $y \in E_1$ such that $\|x - y\|_F$ is minimum. The map π_1 is continuous and $\pi_1^2 = \pi_1$. And similarly for E_2.

The quantity $r = \|X(t) - \pi_1(X(t))\|_F + \|X(t) - \pi_2(X(t))\|_F =$ the distance to the other subset than where $X(t)$ lies, so one can assume that the probability of a transition at t is : $f(r)$ where $f : \mathbb{R} \to [0,1]$ is a probability density. The probability of a transition depends only on the state at t, but one cannot assume that the transitions points θ_α do not depend on X.

The result holds if E_1, E_2 are closed *vector subspaces* of F such that $E_1 \cap E_2 = \{0\}$. Then

$X(t) = \pi_1(X(t)) + \pi_2(X(t))$

and $\|X(t)\|^2 = \|\pi_1(X(t))\|^2 + \|\pi_2(X(t))\|^2$

$\frac{\|\pi_1(X(t))\|^2}{\|X(t)\|^2}$ can be interpreted as the probability that the system at t is in the phase E_1.

One important application is forecasting a transition for a given map X. From the measure of $X(t)$ one can compute for each t the quantity $r(t) = \|X(t) - \pi_1(X(t))\|_F + \|X(t) - \pi_2(X(t))\|_F$ and, if we know f, we have the probability of a transition at t. The practical problem is then to estimate f from the measure of r over a past period $[0, T]$. A very simple, non parametric, estimator can be built when X are maps depending only of t (see J.C.Dutailly *Estimation of the probability of transitions between phases*). It can be used to forecast the occurrence of events such as earth quakes.

2.6 INTERACTING SYSTEMS

2.6.1 Representation of interacting systems

In the propositions above no assumption has been done about the interaction with exterior variables. If the values of some variables are given (for instance to study the impact of external factors with the system) then they shall be fully integrated into the set of variables, at the same footing as the others.

A special case occurs when one considers two systems S_1, S_2, which are similarly represented, meaning that that we have the same kind of variables, defined as identical mathematical objects and related significance. To account for the interactions between the two systems the models are of the form :

$$\begin{array}{ccc} \ulcorner & S_1 & \urcorner \\ X_1 & & Z_1 \\ V_1 & \times & W_1 \\ & \downarrow \Upsilon_1 & \\ & \psi_1 & \\ & H_1 & \end{array} \qquad \begin{array}{ccc} \ulcorner & S_2 & \urcorner \\ X_2 & & Z_2 \\ V_2 & \times & W_2 \\ & \downarrow \Upsilon_2 & \\ & \psi_2 & \\ & H_2 & \end{array}$$

$$\begin{array}{ccc} \ulcorner & S_{1+2} & \urcorner \\ X_1 & & X_2 \\ V_1 & \times & V_2 \\ & & \\ \psi_1 & & \psi_2 \\ H_1 & \times & H_2 \end{array}$$

X_1, X_2 are the variables (as above X denotes collectively a set of variables) characteristic of the systems S_1, S_2,and Z_1, Z_2 are variables representing the interactions. Usually these variables are difficult to measure and to handle. One can consider the system S_{1+2} with the direct product $X_1 \times X_2$, but doing so we obviously miss the interactions Z_1, Z_2.

We see now how it is possible to build a simpler model which keeps the features of S_1, S_2 and accounts for their interactions.

We consider the models without interactions (so with only X_1, X_2) and we assume that they meet the conditions 1. For each model $S_k, k = 1, 2$ there are

a linear map : $\Upsilon_k : V_k \to H_k :: \Upsilon_k (X_k) = \psi_k = \sum_{i \in I_k} \langle \phi_{ki}, \psi_k \rangle e_{ki}$

a positive kernel : $K_k : V_k \times V_k \to \mathbb{R}$

Let us denote S the new model. Its variables will be collectively denoted Y, valued in a Fréchet vector space V'. There will be another Hilbert space H', and a linear map $\Upsilon' : V' \to H'$ similarly defined. As we have the choice of the model, we will impose some properties to Y and V' in order to underline both that they come from S_1, S_2 and that they are interacting.

Condition 28 *i) The variable Y can be deduced from the value of X_1, X_2 : there must be a bilinear map : $\Phi : V_1 \times V_2 \to V'$*

ii) Φ must be such that whenever the systems S_1, S_2 are in the states ψ_1, ψ_2 then S is in the state ψ' and

$\Upsilon'^{-1} (\psi') = \Phi \left(\Upsilon_1^{-1} (\psi_1), \Upsilon_2^{-1} (\psi_2) \right)$

iii) The positive kernel is a defining feature of the models, so we want a positive kernel K' of (V', Υ') such that :

$\forall X_1, X_1' \in V_1, \forall X_2, X_2' \in V_2$:
$K'\left(\Phi\left(X_1, X_2\right), \Phi\left(X_1', X_2'\right)\right) = K_1\left(X_1, X_1'\right) \times K_2\left(X_2, X_2'\right)$

We will prove the following :

Theorem 29 *Whenever two systems S_1, S_2 interact, there is a model S encompassing the two systems and meeting the conditions 28 above. It is obtained by taking the tensor product of the variables specific to S_1, S_2. Then the Hilbert space of S is the tensorial product of the Hilbert spaces associated to each system.*

Proof. First let us see the consequences of the conditions if they are met.

The map : $\varphi : H_1 \times H_2 \to H' :: \varphi\left(\psi_1, \psi_2\right) = \Phi\left(\Upsilon_1^{-1}\left(\psi_1\right), \Upsilon_2^{-1}\left(\psi_2\right)\right)$ is bilinear. So, by the universal property of the tensorial product, there is a unique map $\widehat{\varphi} : H_1 \otimes H_2 \to H'$ such that : $\varphi = \widehat{\varphi} \circ \imath$ where $\imath : H_1 \times H_2 \to H_1 \otimes H_2$ is the tensorial product (Maths.369).

The condition iii) reads :
$\left\langle \Upsilon_1\left(X_1\right), \Upsilon_1\left(X_1'\right)\right\rangle_{H_1} \times \left\langle \Upsilon_2\left(X_2\right), \Upsilon_2\left(X_2'\right)\right\rangle_{H_2}$
$= \left\langle \left(\Upsilon' \circ \Phi\left(\Upsilon_1\left(X_1\right), \Upsilon_2\left(X_2\right)\right), \Upsilon' \circ \Phi\left(\Upsilon_1\left(X_1'\right), \Upsilon_2\left(X_2'\right)\right)\right)\right\rangle_{H'}$
$\left\langle \psi_1, \psi_1'\right\rangle_{H_1} \times \left\langle \psi_2, \psi_2'\right\rangle_{H_2} = \left\langle \varphi\left(\psi_1, \psi_2\right), \varphi\left(\psi_1', \psi_2'\right)\right\rangle_{H'}$
$= \left\langle \widehat{\varphi}\left(\psi_1 \otimes \psi_2\right), \widehat{\varphi}\left(\psi_1' \otimes \psi_2'\right)\right\rangle_{H'}$

The scalar products on H_1, H_2 extend in a scalar product on $H_1 \otimes H_2$, endowing the latter with the structure of a Hilbert space with :
$\left\langle \left(\psi_1 \otimes \psi_2\right), \left(\psi_1' \otimes \psi_2'\right)\right\rangle_{H_1 \otimes H_2} = \left\langle \psi_1, \psi_1'\right\rangle_{H_1} \left\langle \psi_2, \psi_2'\right\rangle_{H_2}$
and then the reproducing kernel is the product of the reproducing kernels (Maths.1208).

So we must have : $\left\langle \widehat{\varphi}\left(\psi_1 \otimes \psi_2\right), \widehat{\varphi}\left(\psi_1' \otimes \psi_2'\right)\right\rangle_{H'} = \left\langle \psi_1 \otimes \psi_2, \psi_1' \otimes \psi_2'\right\rangle_{H_1 \otimes H_2}$ and $\widehat{\varphi}$ must be an isometry : $H_1 \otimes H_2 \to H'$

So by taking $H' = H_1 \otimes H_2$ and $V' = V_1 \otimes V_2$ we meet the conditions. ■

The conditions above are a bit abstract, but are logical and legitimate in the view of the Hilbert spaces. They lead to a natural solution, which is not unique and makes sense only if the systems are defined by similar variables. The measure of the tensor S can be addressed as before, the observables being linear maps defined in the tensorial products $V_1 \otimes V_2$, $H_1 \otimes H_2$ and valued in finite dimensional vector subspaces of these tensor products.

Entanglement

A key point in this representation is the difference between the simple direct product : $V_1 \times V_2$ and the tensorial product $V_1 \otimes V_2$, an issue about which there is much confusion.

The knowledge of the states (X_1, X_2) of both systems requires two vectors of I components each, that is $2 \times I$ scalars, and the knowledge of the state S requires a vector of I^2 components. So the measure of S requires more data, and brings more information, because it encompasses all the interactions. Moreover *a tensor is not necessarily the tensorial product of vectors* (if it is so it is said to be **decomposable**), it is the sum of such tensors. There is no canonical map : $V_1 \otimes V_2 \to V_1 \times V_2$. So *there is no simple and unique way to associate two vectors (X_1, X_2) to one tensor S*. This seems paradoxical, as one could imagine that both systems can always be studied, and their states measured, even if they are interacting. But the simple fact that we consider interactions means that the measure of the state of one of the system shall account for the conditions in which the measure is done, so it shall precise the value of the state of the other system and of the interactions Z_1, Z_2.

If a model is arbitrary, its use must be consistent : if the scientist assumes that there are interactions, they must be present somewhere in the model, as variables for the computations as well as data to be collected. They can be dealt with in two ways. Either we opt for the two systems model, and we have to introduce the variables Z_1, Z_2 representing the interactions, then we have two separate models as in the first section. The study of their interactions can be a topic of the models, but this is done in another picture and requires additional hypotheses about the laws of the interactions. Or, if we intend to account for both systems and their interactions in a single model, we need a representation which supports more information that can bring $V_1 \times V_2$. The tensorial product is one way to enrich the model, this is the most economical and, as far as one follows the guidelines i),ii),iii) above, the only one. The complication in introducing general tensors is the price that we have to pay to account for the interactions. This representation does not, in any way, imply anything about *how* the systems interact, or even if they interact at all (in this case S is always decomposable). As usual the choice is up to the scientist, based upon how he envisions the problem at hand. But he has to live with his choice.

This issue is at the root of the paradoxes of entanglement. With many variants it is an experiment which involves two objects, which interact at the beginning, then are kept separated and non interacting, and eventually one measures the state of one of the two objects, from which the state of the other can be deduced with some probability. If we have two objects which interact at some point, with a significant result because it defines a new state, and we compare their states, then we must either incorporate the interactions, or consider that they constitute a single system and use the tensorial product. The fact that the objects cease to interact at some point does not matter : they are considered together if we compare their states. The interactions must be accounted for, one way or another and, when an evolution is considered, this is the map which represents the whole of the evolution which is significant, not its value at some time.[3]

A common interpretation of this representation is to single out decomposable tensors $\Psi = \psi_1 \otimes \psi_2$, called "pure states", so that actual states would be a superposition of pure states (a concept popularized by the famous Schrödinger's cat). It is clear that in an interacting system the pure states are an abstraction, which actually would represent two non interacting systems, so their superposition is an artificial construction. It can be convenient in simple cases, where the states of each system can be clearly identified, or in complicated models to represent quantities which are defined over the whole system as we will see later. But it does not imply any mysterious feature, notably any probabilist behavior, for the real systems. A state of the two interacting systems is represented by a single tensor, and a tensor is not necessarily decomposable, but it is a sum of decomposable tensors.

2.6.2 Homogeneous systems

The previous result can be extended to N (a number that we will assumed to be fixed) similar systems (that we will call **microsystems**), represented by the same model, interacting together. For each microsystem, identified by a label s, the Hilbert space H and the linear map Υ are the same, the state S of the total system can be represented as a vector belonging to the tensorial product $\mathbf{V}_N = \otimes_{s=1}^N V$, associated to a tensor Ψ belonging to the tensorial product $\mathbf{H}_N = \otimes_{s=1}^N H$. The linear maps $\Upsilon \in \mathcal{L}(V;H)$ can be uniquely extended as maps $\Upsilon_N \in \mathcal{L}(\mathbf{V}_N;\mathbf{H}_N)$ such that (Maths.423) :

$\Upsilon_N(X_1 \otimes ... \otimes X_N) = \Upsilon(X_1) \otimes ... \otimes \Upsilon(X_N)$

The state of the system is then totally defined by the value of tensors S, Ψ, with I^N components.

We have general properties on these tensorial products (Maths.1208).

If $(\widetilde{\varepsilon}_i)_{i \in I}$ is a Hilbertian basis of H then $E_{i_1 \dots i_N} = \widetilde{\varepsilon}_{i_1} \otimes ... \otimes \widetilde{\varepsilon}_{i_N}$ is a Hilbertian basis of $\otimes_{s=1}^N H$. The scalar product is defined by linear extension of

[3] On this point see Haag p.106

$\langle \Psi, \Psi' \rangle_{\mathbf{H}_N} = \langle \psi_1, \psi_1' \rangle_H \times ... \times \langle \psi_N, \psi_N' \rangle_H$
for decomposable tensors : $\Psi = \psi_1 \otimes ... \otimes \psi_N, \Psi' = \psi_1' \otimes ... \otimes \psi_N'$.
The subspaces $\otimes_{s=1}^p H \otimes \widetilde{\varepsilon}_i \otimes_{s=p+2}^N H$ are orthogonal and $\otimes_{s=1}^N H \simeq \ell^2 \left(I^N\right)$
Any operator on H can be extended on $\otimes_{s=1}^N H$ with similar properties : a self adjoint, unitary or compact operator extends uniquely as a self adjoint, unitary or compact operator (Maths.1211).

In the general case the label matters : the state $S = X_1 \otimes ... \otimes X_N$ is deemed different from $S = X_{\sigma(1)} \otimes ... \otimes X_{\sigma(N)}$ where $\left(X_{\sigma(p)}\right)_{p=1}^N$ is a permutation of $(X_s)_{s=1}^N$. If the microsystems have all the same behavior they are, for the observer, indistinguishable. Usually the behavior is related to a parameter analogous to a size, so in such cases the microsystems are assumed to have the same size. We will say that these interacting systems are homogeneous :

Definition 30 *A* ***homogeneous system*** *is a system comprised of a fixed number N of microsystems, represented in the same model, such that any permutation of the N microsystems gives the same state of the total system.*

We have the following result :

Proposition 31 *The states Ψ of homogeneous systems belong to an open subset of a subspace $\boldsymbol{h}$ of the Hilbert space $\otimes_{s=1}^N H$, defined by :*
i) a class of conjugacy $\mathfrak{S}(\lambda)$ of the group of permutations $\mathfrak{S}(N)$,defined itself by a decomposition of N in p parts :
$\lambda = \{0 \leq n_p \leq ... \leq n_1 \leq N, n_1 + ...n_p = N\}$.
ii) p distinct vectors $(\widetilde{\varepsilon}_j)_{j=1}^p$ of a Hermitian basis of H which together define a subspace H_J
iii) The space $\boldsymbol{h}$ of tensors representing the states of the system is then :
either the symmetric tensors belonging to : $\odot_{n_1} H_J \otimes \odot_{n_2} H_J ... \otimes \odot_{n_p} H_J$
or the antisymmetric tensors belonging to : $\Lambda_{n_1} H_J \otimes \wedge_{n_2} H_J ... \otimes \wedge_{n_p} H_J$

Proof. i) In the representation of the general system the microsystems are identified by some label s = 1 ... N. An exchange of labels $U(\sigma)$ is a change of variables, represented by an action of the group of permutations $\mathfrak{S}(N)$: U is defined uniquely by linear extension of $U(\sigma)(X_1 \otimes ... \otimes X_N) = X_{\sigma(1)} \otimes ... \otimes X_{\sigma(N)}$ on decomposable tensors.
We can implement the Theorem 22 proven previously. The tensors ψ representing the states of the system belong to a Hilbert space $\mathbf{H}_N \subset \otimes_{s=1}^N H$ such that $\left(\mathbf{H}_N, \widehat{U}\right)$ is a unitary representation of $\mathfrak{S}(N)$. Which implies that $\mathbf{H}_N$ is invariant by $\widehat{U}$. The action of $\widehat{U}$ on $\otimes_{s=1}^N H$ is defined uniquely by linear extension of
$\widehat{U}(\sigma)(\psi_1 \otimes ... \otimes \psi_N) = \psi_{\sigma(1)} \otimes ... \otimes \psi_{\sigma(N)}$ on decomposable tensors.
$\Psi \in \otimes_{s=1}^N H$ reads in a Hilbert basis $(\widetilde{\varepsilon}_i)_{i \in I}$ of H :
$\Psi = \sum_{i_1...i_N \in I} \Psi^{i_1...i_N} \widetilde{\varepsilon}_{i_1} \otimes ... \widetilde{\varepsilon}_{i_N}$ and :
$\widehat{U}(\sigma)\Psi = \sum_{i_1...i_N \in I} \Psi^{i_1...i_N} \widehat{U}(\sigma)(\widetilde{\varepsilon}_{i_1} \otimes ... \widetilde{\varepsilon}_{i_N}) = \sum_{i_1...i_N \in I} \Psi^{i_1...i_N} \widetilde{\varepsilon}_{\sigma(i_1)} \otimes ... \widetilde{\varepsilon}_{\sigma(i_N)}$
$= \sum_{i_1...i_N \in I} \Psi^{\sigma(i_1)...\sigma(i_N)} \widetilde{\varepsilon}_{i_1} \otimes ... \widetilde{\varepsilon}_{i_N}$
$\left\langle \widehat{U}(\sigma)\Psi, \widehat{U}(\sigma)\Psi' \right\rangle = \langle \Psi, \Psi' \rangle$
$\Leftrightarrow \sum_{i_1...i_N \in I} \Psi^{\sigma(i_1)...\sigma(i_N)} \Psi'^{\sigma(i_1)...\sigma(i_N)} = \sum_{i_1...i_N \in I} \Psi^{i_1...i_N} \Psi'^{i_1...i_N}$
The only vector subspaces of $\otimes_{s=1}^N H$ which are invariant by $\widehat{U}$ and on which $\widehat{U}$ is unitary are spaces of symmetric or antisymmetric tensors :
symmetric : $\Psi^{\sigma(i_1)...\sigma(i_N)} = \Psi^{i_1...i_N}$
antisymmetric : $\Psi^{\sigma(i_1)...\sigma(i_N)} = \epsilon(\sigma) \Psi^{i_1...i_N}$

ii) $\mathfrak{S}(N)$ is a finite, compact group. Its unitary representations are the sum of orthogonal, finite dimensional, unitary, irreducible representations (Maths.1948). Let $\mathbf{h} \subset \otimes_{s=1}^{N} H$ be an irreducible, finite dimensional, representation of $\widehat{U}$. Then $\forall \sigma \in \mathfrak{S}(N) : \widehat{U}(\sigma)\mathbf{h} \subset \mathbf{h}$

iii) Let J a finite subset of I with $card(J) \geq N$, H_J the associated Hilbert space, $\widehat{Y}_J : H \rightarrow H_J$ the projection, and $\widehat{Y}_{J_N} = \otimes_N \widehat{Y}_J$ be the extension of $\widehat{Y}_J$ to $\otimes_{s=1}^{N} H$:

$\widehat{Y}_{J_N} \left(\sum_{i_1 ... i_N \in I} \Psi^{i_1 ... i_N} \widetilde{\varepsilon}_{i_1} \otimes ... \widetilde{\varepsilon}_{i_N} \right) = \sum_{i_1 ... i_N \in J} \Psi^{i_1 ... i_N} \widetilde{\varepsilon}_{i_1} \otimes ... \widetilde{\varepsilon}_{i_N}$

Then :

$\forall \sigma \in \mathfrak{S}(N) : \widehat{U}(\sigma) \widehat{Y}_{J_N} \left(\sum_{i_1 ... i_N \in I} \Psi^{i_1 ... i_N} \widetilde{\varepsilon}_{i_1} \otimes ... \widetilde{\varepsilon}_{i_N} \right)$

$= \sum_{i_1 ... i_N \in J} \Psi^{\sigma(i_1) ... \sigma(i_N)} \widetilde{\varepsilon}_{i_1} \otimes ... \widetilde{\varepsilon}_{i_N} = \widehat{Y}_{J_N} \widehat{U}(\sigma) \Psi$

So if $\mathbf{h}$ is invariant by $\widehat{U}$ then $\widehat{Y}_{J_N}\mathbf{h}$ is invariant by $\widehat{U}$. If $\left(\mathbf{h}, \widehat{U}\right)$ is an irreducible representation then the only invariant subspace are 0 and $\mathbf{h}$ itself, so necessarily $\mathbf{h} \subset \widehat{Y}_{J_N} \left(\otimes_{s=1}^{N} H \right)$ for $card(J) = N$. Which implies : $\mathbf{h} \subset \otimes_N H_J$ with $H_J = \widehat{Y}_J H$ and $card(J) = N$.

iv) There is a partition of $\mathfrak{S}(N)$ in conjugacy classes $\mathfrak{S}(\lambda)$ which are subgroups defined by a decomposition of N in p parts :

$\lambda = \{0 \leq n_p \leq ... \leq n_1 \leq N, n_1 + ... n_p = N\}$. Notice that there is an order on the sets $\{\lambda\}$. Each element of a conjugacy class is then defined by a repartition of the integers $\{1, 2, ...N\}$ in p subsets of n_k items (this is a Young Tableau) (Maths. 5.2.2). A class of conjugacy is an abelian subgroup of $\mathfrak{S}(N)$: its irreducible representations are unidimensional.

The irreducible representations of $\mathfrak{S}(N)$ are then defined by a class of conjugacy, and the choice of a vector.

$\mathbf{h}$ is a Hilbert space, thus it has a Hilbertian basis, composed of decomposable tensors which are of the kind $\widetilde{\varepsilon}_{j_1} \otimes ... \otimes \widetilde{\varepsilon}_{j_N}$ where $\widetilde{\varepsilon}_{j_k}$ are chosen among the vectors of a Hermitian basis $(\widetilde{\varepsilon}_j)_{j \in J}$ of H_J

If $\widetilde{\varepsilon}_{j_1} \otimes ... \otimes \widetilde{\varepsilon}_{j_N} \in H, \forall \sigma \in \mathfrak{S}(N) : \widehat{U}(\sigma) \widetilde{\varepsilon}_{j_1} \otimes ... \otimes \widetilde{\varepsilon}_{j_N} = \widetilde{\varepsilon}_{j_{\sigma(1)}} \otimes ... \otimes \widetilde{\varepsilon}_{j_{\sigma(N)}} \in \mathbf{h}$

and because the representation is irreducible the basis of $\mathbf{h}$ is necessarily composed from a set of $p \leq N$ vectors $\widetilde{\varepsilon}_j$ by action of $\widehat{U}(\sigma)$

Conversely : for any Hermitian basis $(\widetilde{\varepsilon}_i)_{i \in I}$ of H, any subset J of cardinality N of I, any conjugacy class λ, any family of vectors $(\widetilde{\varepsilon}_{j_k})_{k=1}^{p}$ chosen in $(\widetilde{\varepsilon}_i)_{i \in J}$, the action of $\widehat{U}$ on the tensor :

$\Psi_\lambda = \otimes_{n_1} \widetilde{\varepsilon}_{j_1} \otimes_{n_2} \widetilde{\varepsilon}_{j_2} ... \otimes_{n_p} \widetilde{\varepsilon}_{j_p}, j_1 \leq j_2 .. \leq j_p$

gives the same tensor if $\sigma \in \mathfrak{S}(\lambda) : \widehat{U}(\sigma) \Psi_\lambda = \Psi_\lambda$

gives a different tensor if $\sigma \in \mathfrak{S}(\lambda^c)$ the conjugacy class complementary to $\mathfrak{S}(\lambda) : \mathfrak{S}(\lambda^c) = \complement_{\mathfrak{S}(N)}^{\mathfrak{S}(\lambda)}$

so it provides an irreducible representation by :

$\forall \Psi \in \mathbf{h} : \Psi = \sum_{\sigma \in \mathfrak{S}(\lambda^c)} \Psi^\sigma \widehat{U}(\sigma) \left(\otimes_{n_1} \widetilde{\varepsilon}_{j_1} \otimes_{n_2} \widetilde{\varepsilon}_{j_2} ... \otimes_{n_p} \widetilde{\varepsilon}_{j_p} \right)$

where the components Ψ^σ are labeled by the vectors of a basis of $\mathbf{h}$. The dimension of $\mathbf{h}$ his given by the cardinality of $\mathfrak{S}(\lambda^c)$ that is : $\frac{N!}{n_1! ... n_p!}$. All the vector spaces $\mathbf{h}$ of the same conjugacy class (but different vectors $\widetilde{\varepsilon}_i$) have the same dimension, thus they are isomorphic.

v) A basis of $\mathbf{h}$ is comprised of tensorial products of N vectors of a Hilbert basis of H. So we can give the components of the tensors of $\mathbf{h}$ with respect to $\otimes_{s=1}^{N} H$. We have two non equivalent representation :

By symmetric tensors : $\mathbf{h}$ is then isomorphic to $\odot_{n_1} H_J \otimes \odot_{n_2} H_J ... \otimes \odot_{n_p} H_J$ where the symmetric tensorial product $\odot$ and the space of n order symmetric tensor on H_J is $\odot_n H_J$

By antisymmetric tensors : $\mathbf{h}$ is then isomorphic to $\wedge_{n_1} H_J \otimes \wedge_{n_2} H_J ... \otimes \wedge_{n_p} H_J$ and the space of n order antisymmetric tensor on H_J is $\wedge_n H_J$

The result extends to V_N by : $S = \Upsilon_N^{-1}(\Psi)$ ■

Remarks

i) For each choice of a class of conjugacy, and each choice of the vectors $(\widetilde{\varepsilon}_j)_{j=1}^p$ which defines H_J, we have a different irreducible representation with vector space **h.** Different classes of conjugacy gives non equivalent representations. But different choices of the Hermitian basis $(\widetilde{\varepsilon}_j)_{j\in I}$ and the subset J of I, for a given class of conjugacy, give equivalent representations, and they can be arbitrary. So, for a given system, the set of states is characterized by a subset J of N elements in any basis of H, and by a class of conjugacy.

A change of the state of the system can occur either inside the same vector space **h**, or between irreducible representations: **h**→**h'**. As we will see in the next chapters usually the irreducible representation is fixed by other variables (such that energy) and a change of irreducible representation implies a discontinuous process. The states of the total system are quantized by the interactions.

ii) $\otimes_{n_1}\widetilde{\varepsilon}_{j_1}\otimes_{n_2}\widetilde{\varepsilon}_{j_2}...\otimes_{n_p}\widetilde{\varepsilon}_{j_p}$ can be seen as representing a configuration where n_k microsystems are in the same state $\widetilde{\varepsilon}_{j_k}$.The class of conjugacy, characterized by the integers n_p, correspond to the distribution of the microsystems between fixed states.

iii) If O is a convex subset then S belongs to a convex subset, and the basis can be chosen such that $\forall\Psi\in\mathbf{h}$ is a linear combination $(y_k)_{k=1}^q$ of the generating tensors with $y_k\in[0,1],\sum_{k=1}^q y_k=1$. S can then be identified to the expected value of a random variable which would take one of the value $\otimes_{n_1}X_1\otimes_{n_2}X_2...\otimes_{n_p}X_p$, which corresponds to n_k microsystems having the state X_k. As exposed above the identification with a probabilist model is formal : there is no random behavior assumed for the physical system.

iv) In the probabilist picture one can assume that each microsystem behaves independently, and has a probability π_j to be in the state represented by $\widetilde{\varepsilon}_j$ and $\sum_{j=1}^N\pi_j=1$. Then the probability that we have $(n_k)_{k=1}^p$ microstates in the states $(\widetilde{\varepsilon}_k)_{k=1}^p$ is $\frac{N!}{n_1!...n_p!}(\pi_{j_1})^{n_1}...(\pi_{j_p})^{n_p}$.

v) The set of symmetric tensor $\odot_n H_J$ is a closed vector subspace of $\otimes_n H_J$, this is a Hilbert space, $\dim\otimes_n H_J=C_{p+n-1}^{p-1}$ with Hilbertian basis $\frac{1}{\sqrt{n!}}\odot_{j\in J}\widetilde{\varepsilon}_j=\frac{1}{\sqrt{n!}}S_n(\otimes_{j\in J}\widetilde{\varepsilon}_j)$ where the symmetrizer is :

$S_n\left(\sum_{(i_1...i_n)}\psi^{i_1..i_n}\widetilde{\varepsilon}_{i_1}\otimes..\otimes\widetilde{\varepsilon}_{i_n}\right)=\sum_{(i_1...i_n)}\psi^{i_1..i_n}\sum_{\sigma\in\mathfrak{S}(n)}\widetilde{\varepsilon}_{\sigma(1)}\otimes....\widetilde{\varepsilon}_{\sigma(k)}$

A tensor is symmetric iff : $\Psi\in\odot_n H_J\Leftrightarrow S_n(\Psi)=n!\Psi$ (Maths. 7.2.1,13.5.2).

The set of antisymmetric tensor $\Lambda_n H_J$ is a closed vector subspace of $\otimes_n H_J$, this is a Hilbert space, $\dim\wedge_n H_J=C_p^n$ with Hilbertian basis $\frac{1}{\sqrt{n!}}\wedge_{j\in J}\widetilde{\varepsilon}_j=\frac{1}{\sqrt{n!}}A_n(\otimes_{j\in J}\widetilde{\varepsilon}_j)$ with the antisymmetrizer :

$A_n\left(\sum_{(i_1...i_n)}\psi^{i_1..i_n}\widetilde{\varepsilon}_{i_1}\otimes..\otimes\widetilde{\varepsilon}_{i_n}\right)=\sum_{(i_1...i_n)}\psi^{i_1..i_n}\sum_{\sigma\in\mathfrak{S}(n)}\epsilon(\sigma)\,\widetilde{\varepsilon}_{\sigma(1)}\otimes....\widetilde{\varepsilon}_{\sigma(k)}$

A tensor is antisymmetric iff : $\Psi\in\wedge_n H_J\Leftrightarrow A_n(\Psi)=n!\Psi$ (Maths. 7.2.2,13.5.2)

v) for $\theta\in\mathfrak{S}(N)$: $\widehat{U}(\theta)\Psi$ is usually different from Ψ

2.6.3 Global observables of homogeneous systems

The previous definitions of observables can be extended to homogeneous systems. An observable is defined on the total system, this is a map : $\Phi:\mathbf{V}_N\to W$ where W is a finite dimensional vector subspace of $\mathbf{V}_N$, but not necessarily a tensorial vector product of spaces. To Φ is associated the self-adjoint operator $\widehat{\Phi}=\Upsilon\circ\Phi\circ\Upsilon^{-1}$ and $H_\Phi=\widehat{\Phi}\left(\otimes_{s=1}^N H\right)\subset\otimes_{s=1}^N H$.

Theorem 32 *Any observable of a homogeneous system is of the form :*

$\Phi:\mathbf{V}_N\to W$ where W is generated by vectors Φ_λ associated to each class of conjugacy of $\mathfrak{S}(N)$

The value of $\Phi(X_1\otimes...\otimes X_N)=\varphi(X_1,...,X_N)\Phi_\lambda$ where φ is a scalar linear symmetric map, if the system is in a state corresponding to λ

Proof. The space W must be invariant by U and H_Φ invariant by $\widehat{U}$. If the system is in a state belonging to $\mathbf{h}$ for a class of conjugacy λ, then $H_\Phi = \widehat{\Phi}\mathbf{h}$ and $\left(\widehat{\Phi}\mathbf{h},\widehat{U}\right)$ is an irreducible representation of the abelian subgroup $\mathfrak{S}(\lambda)$ corresponding to λ. It is necessarily unidimensional and $\Phi(X_1 \otimes ... \otimes X_N)$ is proportional to a unique vector. The observable being a linear map, the function φ is a linear map of the components of the tensor. ■

There is no way to estimate the state of each microsystem. From a practical point of view, this is a vector $\gamma = \widehat{\Phi}\left(\otimes_{n_1}\widetilde{\varepsilon}_{j_1} \otimes_{n_2} \widetilde{\varepsilon}_{j2} ... \otimes_{n_p} \widetilde{\varepsilon}_{j_p}\right)$ which is measured, and from it $\lambda, (\widetilde{\varepsilon}_{j_k})_{k=1}^p$ are estimated.

In the probabilist picture the expected value of γ is :
$\langle\gamma\rangle = z(\pi_1, ..., \pi_N)$
with
$z(\pi_1, ..., \pi_N)$
$= \sum_\lambda \frac{N!}{n_1!...n_p!} \sum_{1\le j_1 \le .. \le j_p \le N} (\pi_{j_1})^{n_1} ... (\pi_{j_p})^{n_p}\, \widehat{\Phi}\left(\otimes_{n_1}\varepsilon_{j_1} ... \otimes_{n_p} \varepsilon_{j_p}\right)$

We have a classic statistical problem : estimate the π_i from a statistic given by the measure of γ. If the statistic $\widehat{\Phi}$ is sufficient, meaning that π_i depends only on γ, as F is finite dimensional whatever the number of microsystems, the Pitman-Koopman-Darmois theorem tells us that the probability law is exponential, then an estimation by the maximum likehood gives the principle of Maximum Entropy with entropy :
$E = -\sum_{j=1}^N \pi_j \ln \pi_j$

In the usual interpretation of the probabilist picture, it is assumed that the state of each microsystem can be measured independently. Then the entropy $E = -\sum_{j=1}^N \pi_j \ln \pi_j$ can be seen as a measure of the heterogeneity of the system. And, contrary to a usual idea, the interactions between the micro-systems do not lead to the homogenization of their states, but to their quantization : the states are organized according to the classes of conjugacy.

2.6.4 Evolution of homogeneous systems

The evolution of homogeneous systems raises many interesting issues. The assumptions are a combination of the previous conditions.

Theorem 33 *For a model representing the evolution of a homogeneous system comprised of a fixed number N of microsystems s = 1 ...N which are represented by the same model, with variables* $(X_s)_{s=1}^N$ *such that, for each microsystem :*

i) the variables X_s *are maps :* $X_s :: R \to E$ *where R is an open subset of* $\mathbb{R}$ *and E a normed vector space, belonging to an open subset O of an infinite dimensional Fréchet space V*

ii) $\forall t \in R$ *the evaluation map :* $\mathcal{E}(t) : O \to E : \mathcal{E}(t) X_s = X_s(t)$ *is continuous*

iii) $\forall t \in R : X_s(t) = X'_s(t) \Rightarrow X_s = X'_s$

There is a map : $S : R \to \otimes_N F$ *such that* $S(t)$ *represents the state of the system at t.* $S(t)$ *takes its value in a vector space* $f(t)$ *such that* $\left(\mathbf{f}(t), \widehat{U}_F\right)$ *, where* $\widehat{U}_F$ *is the permutation on* $\otimes_N F$*, is an irreducible representation of* $\mathfrak{S}(N)$

The crucial point is that the homogeneity is understood as the microsystems follow the same laws, but at a given time they do not have necessarily the same state.

Proof. i) Implement the Theorem 2 for each microsystem : there is a common Hilbert space H associated to V and a continuous linear map $\Upsilon : V \to H :: \psi_s = \Upsilon(X_s)$

ii) Implement the Theorem 31 on the homogeneous system, that is for the whole of its evolution. The state of the system is associated to a tensor $\Psi \in \mathbf{h}$ where $\mathbf{h}$ is defined by a

Hilbertian basis $(\widetilde{\varepsilon}_i)_{i\in I}$ of H, a finite subset J of I, a conjugacy class λ and a family of p vectors $(\widetilde{\varepsilon}_{j_k})_{k=1}^p$ belonging to $(\widetilde{\varepsilon}_i)_{i\in J}$. The vector space $\mathbf{h}$ stays the same whatever t.

iii) Implement the Theorem 26 on the evolution of each microsystem : there is a common Hilbert space F, a map : $\widehat{\mathcal{E}} : R \to \mathcal{L}(H;F)$ such that : $\forall X_s \in O : \widehat{\mathcal{E}}(t)\Upsilon(X_s) = X_s(t)$ and $\forall t \in R$, $\widehat{\mathcal{E}}(t)$ is an isometry

Define $\forall i \in I : \varphi_i : R \to F :: \varphi_i(t) = \widehat{\mathcal{E}}(t)\widetilde{\varepsilon}_i$

iv) $\widehat{\mathcal{E}}(t)$ can be uniquely extended in a continuous linear map :

$\widehat{\mathcal{E}}_N(t) : \otimes_N H \to \otimes_N F$ such that : $\widehat{\mathcal{E}}_N(t)(\otimes_N \psi_s) = \otimes_N X_s(t)$

$\widehat{\mathcal{E}}_N(t)\left(\otimes_{s=1}^N \widetilde{\varepsilon}_{i_s}\right) = \otimes_{s=1}^N \varphi_{i_s}(t)$

$\widehat{\mathcal{E}}_N(t)$ is an isometry, so $\forall t \in R : \left\{\otimes_{s=1}^N \varphi_{i_s}(t), i_s \in I\right\}$ is a Hilbertian basis of $\otimes_N F$

v) Define as the state of the system at t : $S(t) = \widehat{\mathcal{E}}_N(t)(\Psi) \in \otimes_N F$

Define : $\forall \sigma \in \mathfrak{S}(N) : \widehat{U}_F(\sigma) \in \mathcal{L}(\otimes_N F; \otimes_N F)$ by linear extension of : $\widehat{U}_F(\sigma)\left(\otimes_{s=1}^N f_s\right) = \otimes_{s=1}^N f_{\sigma(s)}$

$\widehat{U}_F(\sigma)\left(\otimes_{s=1}^N \varphi_{i_s}(t)\right) = \otimes_{s=1}^N \varphi_{\sigma(i_s)}(t) = \widehat{\mathcal{E}}_N(t)\widehat{U}(\sigma)\left(\otimes_{s=1}^N \widetilde{\varepsilon}_{i_s}\right)$

$\forall \Psi \in \mathbf{h} : \Psi = \sum_{\sigma\in\mathfrak{S}(\lambda^c)} \Psi^\sigma \widehat{U}(\sigma)\left(\otimes_{n_1}\widetilde{\varepsilon}_{j_1} \otimes_{n_2} \widetilde{\varepsilon}_{j_2} ... \otimes_{n_p} \widetilde{\varepsilon}_{j_p}\right)$

$S(t) = \sum_{\sigma\in\mathfrak{S}(\lambda^c)} \Psi^\sigma \widehat{\mathcal{E}}_N(t) \circ \widehat{U}(\sigma)\left(\otimes_{n_1}\widetilde{\varepsilon}_{j_1} \otimes_{n_2} \widetilde{\varepsilon}_{j_2} ... \otimes_{n_p} \widetilde{\varepsilon}_{j_p}\right)$

$S(t) = \sum_{\sigma\in\mathfrak{S}(\lambda^c)} \Psi^\sigma \widehat{U}_F(\sigma) \otimes_{n_1} \varphi_{j_1}(t) \otimes_{n_2} \varphi_{j_2}(t) ... \otimes_{n_p} \varphi_{j_p}(t)$

$\forall \theta \in \mathfrak{S}(\lambda) : \widehat{U}_F(\theta)\left(\otimes_{n_1}\varphi_{j_1}(t) \otimes_{n_2} \varphi_{j_2}(t) ... \otimes_{n_p} \varphi_{j_p}(t)\right)$

$= \otimes_{n_1}\varphi_{j_1}(t) \otimes_{n_2} \varphi_{j_2}(t) ... \otimes_{n_p} \varphi_{j_p}(t)$

$\forall \theta \in \mathfrak{S}(\lambda^c) : \widehat{U}_F(\theta)\left(\otimes_{n_1}\varphi_{j_1}(t) \otimes_{n_2} \varphi_{j_2}(t) ... \otimes_{n_p} \varphi_{j_p}(t)\right)$

$\neq \left(\otimes_{n_1}\varphi_{j_1}(t) \otimes_{n_2} \varphi_{j_2}(t) ... \otimes_{n_p} \varphi_{j_p}(t)\right)$

and the tensors are linearly independent

So $\left\{\widehat{U}_F(\sigma)\left(\otimes_{n_1}\varphi_{j_1}(t) \otimes_{n_2} \varphi_{j_2}(t) ... \otimes_{n_p} \varphi_{j_p}(t)\right), \sigma \in \mathfrak{S}(\lambda^c)\right\}$ is an orthonormal basis of

$\mathbf{f}(t) = Span\left\{\widehat{U}_F(\sigma)\left(\otimes_{n_1}\varphi_{j_1}(t) \otimes_{n_2} \varphi_{j_2}(t) ... \otimes_{n_p} \varphi_{j_p}(t)\right), \sigma \in \mathfrak{S}(\lambda^c)\right\}$

$\mathbf{f}(t) = \widehat{\mathcal{E}}_N(t)(\mathbf{h})$

Let $\widetilde{f}(t) \subset \mathbf{f}(t)$ be any subspace globally invariant by $\left\{\widehat{U}_F(\theta), \theta \in \mathfrak{S}(N)\right\} : \widehat{U}_F(\theta)\widetilde{f}(t) \in \widetilde{f}(t)$

$\widehat{\mathcal{E}}_N(t)$ is an isometry, thus a bijective map

$\widetilde{h} = \widehat{\mathcal{E}}_N(t)^{-1}\widetilde{f}(t) \Leftrightarrow \widetilde{f}(t) = \widehat{\mathcal{E}}_N(t)\widetilde{h}$

$\widehat{U}_F(\theta)\widehat{\mathcal{E}}_N(t)\widetilde{h} \in \widehat{\mathcal{E}}_N(t)\widetilde{h}$

$\forall \Psi \in \mathbf{h} : \widehat{U}_F(\theta)\widehat{\mathcal{E}}_N(t)\Psi = \widehat{\mathcal{E}}_N(t)\widehat{U}(\theta)\Psi$

$\Rightarrow \widehat{\mathcal{E}}_N(t)\widehat{U}(\theta)\widetilde{h} \in \widehat{\mathcal{E}}_N(t)\widetilde{h}$

$\Rightarrow \widehat{U}(\theta)\widetilde{h} \in \widetilde{h}$

So $\left(\mathbf{f}(t), \widehat{U}_F\right)$ is an irreducible representation of $\mathfrak{S}(N)$ ∎

For each t the space $\mathbf{f}(t)$ is defined by a Hilbertian basis $(f_i)_{i\in I}$ of F, a finite subset J of I, a conjugacy class $\lambda(t)$ and a family of p vectors $(f_{j_k}(t))_{k=1}^p$ belonging to $(f_i)_{i\in J}$. The set J is arbitrary but defined by $\mathbf{h}$, so it does not depend on t**.** For a given class of conjugacy different families of vectors $(f_{j_k}(t))_{k=1}^p$ generate equivalent representations and isomorphic spaces, by symmetrization or antisymmetrization. So for a given system one can pick up a fixed ordered family $(f_j)_{j=1}^N$ of vectors in $(f_i)_{i\in I}$ such that for each class of conjugacy $\lambda = \{0 \le n_p \le ... \le n_1 \le N, n_1 + ...n_p = N\}$ there is a unique vector space $\mathbf{f}_\lambda$ defined by $\otimes_{n_1} f_1 \otimes_{n_2} f_2 ... \otimes_{n_p} f_p$. Then if $S(t) \in \mathbf{f}_\lambda$:

$S(t) = \sum_{\sigma\in\mathfrak{S}(\lambda^c)} S^\sigma(t)\widehat{U}_F(\sigma)\left(\otimes_{n_1} f_1 \otimes_{n_2} f_2 ... \otimes_{n_p} f_p\right)$

and at all time $S(t) \in \otimes_N F_J$.
The vector spaces $\mathbf{f}_\lambda$ are orthogonal. With the orthogonal projection π_λ on $\mathbf{f}_\lambda$:
$\forall t \in R : S(t) = \sum_\lambda \pi_\lambda S(t)$
$\|S(t)\|^2 = \sum_\lambda \|\pi_\lambda S(t)\|^2$
The distance between S(t) and a given $\mathbf{f}_\lambda$ is well defined and :
$\|S(t) - \pi_\lambda S(t)\|^2 = \|S(t)\|^2 - \|\pi_\lambda S(Ut)\|^2$

Whenever S, and thus Θ, is continuous, the space $\mathbf{f}_\lambda$ stays the same. As we have seen previously one can assume that, in all practical cases, Θ is continuous but for a countable set $\{t_k, k = 1, 2..\}$ of isolated points. Then the different spaces $\mathbf{f}_\lambda$ can be seen as phases, each of them associated with a class of conjugacy λ. And there are as many possible phases as classes of conjugacy. So, in a probabilist picture, one can assume that the probability for the system to be in a phase $\lambda : \Pr(S(t) \in \mathbf{f}_\lambda)$ is a function of $\frac{\|\pi_\lambda S(t)\|^2}{\|S(t)\|^2}$. It can be estimated as seen previously from data on a past period, with the knowledge of both λ and $\frac{\|\pi_\lambda S(t)\|^2}{\|S(t)\|^2}$.

2.7 CORRESPONDENCE WITH QUANTUM MECHANICS

It is useful to compare the results proven in the present paper to the axioms of QM as they are usually expressed.

2.7.1 Hilbert space

QM : 1. *The states of a physical system can be represented by rays in a complex Hilbert space H. Rays meaning that two vectors which differ by the product by a complex number of module 1 shall be considered as representing the same state.*

In Theorem 2 we have proven that in a model meeting precise conditions the states of the system can be represented as vectors in an infinite dimensional, separable, real Hilbert space. We have seen that it is always possible to endow the Hilbert space with a complex structure, but this is not a necessity. Moreover the Hilbert space is defined up to an isometry, so notably up to the product by a fixed complex scalar of module 1. We will see in the following how and why rays appear (this is specific to the representation of particles with electromagnetic fields).

In Quantum Physics a great attention is given to the Principle of Superposition. This Principle is equivalent to the condition that the variables of the system (and then its state) belong to a vector space. There is a distinction between pure states, which correspond to actual measures, and mixed states which are linear combination of pure states, usually not actually observed. There has been a great effort to give a physical meaning to these mixed states. Here the concept of pure states appears only in the tensors representing interacting systems, with the usual, but clear, explanation. In Quantum Mechanics some states of a system cannot be achieved (through a preparation for instance) as a combination of other states, and then super-selection rules are required to sort out these specific states. Here there is a simple explanation : because the set H_0 is not the whole of H it can happen that a linear combination of states is not inside H_0. The remedy is to enlarge the model to account for other physical phenomena, if it appears that these states have a physical meaning.

Actually the main difference comes from the precise conditions of the Theorem 2. The variables must be maps, but also belong to a vector space. Thus for instance it does not apply to the model of a solid body represented by its trajectory $x(t)$ and its speed $v(t)$: the variable $x(t)$ is a map : $x : \mathbb{R} \to M$ valued in a manifold (an affine space in Galilean geometry). So it is necessary to adapt the model, using the fiber bundle formalism, and this leads to a deep redefinition of the concept of motion (including rotation) and to the spinors. And as it has been abundantly said, the state is defined by maps over the evolution of the system, and not pointwise.

2.7.2 Observables

QM : 2. *To any physical measure Φ, called an observable, which can be done on the system, is associated a continuous, linear, self-adjoint operator $\widehat{\Phi}$ on H.*

We have proven that this operator is also compact and trace-class. The main result is that we have here a clear understanding of the concept of observable, rooted in the practical way the data are analyzed and assigned to the value of the variables, with the emphasize given to the procedure of specification, an essential step in any statistical analysis and which is usually overlooked.

There is no assumption about the times at which the measures are taken, when the model represents a process the measures can be taken at the beginning, during the process, or at the end. The variables which are estimated are maps, and the estimation of maps requires more than one value of the arguments. The estimation is done by a statistical method which uses all the available data. From this point of view our picture is closer to what is done in the laboratories, than to the idealized vision of simultaneous measures, which should be taken all together at each time, and would be impossible because of the perturbation caused by the measure.

In QM a great emphasize if given to the commutation of observables, linked to the physical possibility to measure simultaneously two variables. This concept does not play any role here. the product of observables itself has no clear meaning and no use. If a variable is added, we have another model, the variable gets the same status as the others, and it is assumed that it can be measured.

Actually the importance granted to the simultaneity of measures, magnified by Dirac, is somewhat strange. It is also problematic in the Relativist picture. It is clear that some measures cannot be done, at the atomic scale, without disturbing the state of the system that is studied, but this does not preclude to use the corresponding variables in a model, or give them a special status. Before the invention of radar the artillerymen used efficient models even if they were not able to measure the speed of their shells. And in a collider it is assumed that the speed and the location of particles are known when they collide.

From primary observables it is possible to define von Neumann algebras of operators, which are necessarily commutative when a fixed basis has been chosen. As the choice of a privileged basis can always be done, one can say that there is always a commutative von Neumann algebra associated to a system. One can link the choice of a privileged basis to an observer, then, for a given observer, the system can be represented by a commutative von Neumann algebra, and it would be interesting to see what are the consequences for the results already achieved. In particular the existence of a commutative algebra nullifies the emphasize given to the commutation of operators, or at least, it should be understood as the change of observer. But these von Neumann algebras do not play any role in the proofs of the theorems. Their introduction can be useful, but they are not a keystone in our framework.

2.7.3 Measure

QM : 3. *The result of any physical measure is one of the eigen-values* λ *of the associated operator* $\widehat{\Phi}$. *After the measure the system is in the state represented by the corresponding eigen vector* ψ_λ

This is one of the most puzzling axiom. We have here a clear interpretation of this result, with primary observables, and there is always a primary observable which is at least as efficient than a secondary observable.

In our picture there is no assumption about how the measures are done, and particularly if they have or not an impact on the state of the system. If it is assumed that this is the case, a specific variable should be added to the model. Its value can be measured directly or estimated from the value of the other variables, but this does not make a difference : it is a variable as the others.

2.7.4 Probability

QM : 4. *The probability that the measure is* λ *is equal to* $|\langle\psi_\lambda,\psi\rangle|^2$ *(with normalized eigen vectors). If a system is in a state represented by a normalized vector* ψ *, and an experiment is* *done to test whether it is in one of the states* $(\psi_n)_{n=1}^N$ *which constitutes an orthonormal set of vectors, then the probability of finding the system in the state* ψ_n *is* $|\langle\psi_n,\psi\rangle|^2$.

The first part is addressed by the theorem 17. The second part has no direct equivalent in our picture but can be interpreted as follows : a measure of the primary observable has shown that $\psi \in H_J$, then the probability that it belongs to $H_{J'}$ for any subset $J' \subset J$ is $\left\|\widehat{Y}_{J'}(\psi)\right\|^2$. It is a computation of conditional probabilities :

Proof. The probability that $\psi \in H_K$ for any susbset $K \subset I$ is $\left\|\widehat{Y}_K(\psi)\right\|^2$. The probability that $\psi \in H_{J'}$ knowing that $\psi \in H_J$ is :

$\Pr(\psi \in H_{J'}|\psi \in H_J) = \frac{\Pr(\psi\in H_{J'}\wedge\psi\in H_J)}{\Pr(\psi\in H_{J'}|\psi\in H_J)} = \frac{\Pr(\psi\in H_{J'})}{\Pr(\psi\in H_{J'}|\psi\in H_J)} = \frac{\left\|\widehat{Y}_{J'}(\psi)\right\|^2}{\left\|\widehat{Y}_J(\psi)\right\|^2} = \left\|\widehat{Y}_{J'}(\psi)\right\|^2$ because $\widehat{Y}_{J'}(\psi) = \psi$ and $\|\psi\| = 1$ ■

Moreover we have seen how the concept of wave functions can be introduced, and its meaning, for models where the variables are maps defined on the same set. Of course the possibility to define such a function does not imply that it is related to a physical phenomenon.

2.7.5 Interacting systems

QM : 5. *When two systems interacts, the vectors representing the states belong to the tensorial product of the Hilbert states.*

This is the topic of the theorem 28. We have seen how it can be extended to N systems, and the consequences that entails for homogeneous systems. If the number of microsystems is not fixed, the formalism of Fock spaces can be used but would require a mathematical apparatus that is beyond the scope of this book.

There is a fierce debate about the issue of locality in physics, mainly related to the entanglement of states for interacting particles. It should be clear that the formal system that we have built is global : more so, it is its main asset. While most of the physical theories are local, with the tools which have been presented we can deal with variables which are global, and get some strong results without many assumptions regarding the local laws.

2.7.6 Wigner's theorem

QM : 6. *If the same state is represented by two rays* R, R'*, then there is an operator* $\widehat{U}$*, unitary or antiunitary, on the Hilbert space* H *such that if the state* ψ *is in the ray* R *then* $\widehat{U}\psi$ *is in the ray* R'.

This the topic of the theorem 21. The issue unitary / antiunitary exists in the usual presentation of QM because of the rays. In our picture the operator is necessarily unitary, which is actually usually the case.

2.7.7 Schrödinger equation

QM : 7. *The vector representing the state of a system which evolves with time follows the equation :* $i\hbar\frac{\partial\psi}{\partial t} = \widehat{H}\psi$ where $\widehat{H}$ *is the Hamiltonian of the system.*

This is actually the topic of the theorem 27 and the result holds for the variables X in specific conditions, including in the General Relativity context. The imaginary i does not appear because the Hilbert space is real. As for Planck's constant of course it cannot appear in a formal model. However as said before all quantities must be dimensionless, as it is obvious in the equivalent expression $\psi(t) = \exp \frac{t}{i\hbar} \widehat{H} \psi(0)$. Thus it is necessary either to involve some constant, or that all quantities (including the time t) are expressed in a universal system of units. This is commonly done by using the Planck's system of units. Which is more important is that the theorems (and notably the second) precise fairly strong conditions for their validity. In many cases the Schrödinger's equation, because of its linearity, seems "to good to be true". We can see why.

2.7.8 The scale issue

The results presented here hold whenever the model meets the conditions 1. So it is valid whatever the scale. But it is clear that the conditions are not met in many models used in classic physics, notably in Analytic Mechanics (the variables q are not vectorial quantities). Moreover actually in the other cases it can often be assumed that the variables belong themselves to Hilbert spaces. The results about observables and eigen values are then obvious, and those about the evolution of the system, for interacting systems or for gauge theories keep all their interest.

The "Quantic World", with its strange properties does not come from specific physical laws, which would appear below some scale, but from the physical properties of the atomic world themselves. And of course these cannot be addressed in the simple study of formal models : they will be the topic of the rest of this book.

So the results presented here, which are purely mathematical, give a consistent and satisfying explanation of the basic axioms of Quantum Mechanics, without the need for any exotic assumptions. They validate, and in many ways make simpler and safer, the use of techniques used for many years. Moreover, as it is easy to check, most of these results do not involve any physics at all : they hold for any scientific theory which is expressed in a mathematical formalism. From my point of view they bring a definitive answer to the issue of the interpretation of QM : the interpretations were sought in the physical world, but actually there is no such interpretation to be found. There is no physical interpretation because QM is not a physical theory.

The results presented go beyond the usual axioms of QM : on the conditions to detect an anomaly, on the quantization of a variable $Y = f(X)$, on the phases transitions. And other results can probably be found. So the method should give a fresh view of the foundations of QM in Physics.

Chapter 3

GEOMETRY

Almost all, if not all, measures rely on measures of lengths and times. These concepts are expressed in theories about the geometry of the universe, meaning of the container in which live the objects of physics. The issue here is not a model of the Universe, seen in its totality, which is the topic of Cosmology, but a model which tells us how to measure lengths and times, and how to compare measures done by different observers. Such a model is a prerequisite to any physical theory. Geometry, as a branch of Mathematics, is the product of this quest of a theory of the universe, and naturally a physical geometry is formalized with the tools of Mathematical Geometry. There are several Geometries used in Physics : Galilean Geometry, Special Relativity (SR) and General Relativity (GR).

In this first section we will see how such a geometry can be built, from simple observations. We will go directly to the General Relativity model. This is the one which is the most general and will be used in the rest of the book. It is said to be difficult, but actually these difficulties can be overcome with the right formalism. Moreover it forces us to leave usual representations, which are often deceptive.

3.1 MANIFOLD STRUCTURE

3.1.1 The Universe has the structure of a manifold

The first question is how do we measure a location ?

In almost all Physics books the answer will go straight to a an orthonormal frame, or in GR to a map with some $\partial\xi_\alpha$, often with additional provisions for “inertial frames”, before a complicated discourse about light, and quite often trains. Actually, and what is somewhat strange for academics who pride themselves to be respectful of experiments, all these narratives, simply, do not respect the facts.

A small distances it is possible to measure lengths by surveying, and indeed the scientists who established the meter in 1792 based their work on a strict survey along 15 kms. Then it is possible to use an orthonormal frame. But even at small scale, topographers use a set of 3 angles with respect to fixed directions given by staffs, or far enough objects, points in the landscape, or distant stars, combined with one measure of distance. The latter is measured usually by the delay for a signal emitted to rebound on the surface on a distant object. There are small, clever, devices which do that with ultrasound, radars use electromagnetic fields. The speed of the propagation of the signal is taken conventionally fixed and constant. It is assumed to have been measured at small scale, and the results are then extended for larger distances. For not too far away celestial bodies, the distances can be measured using the angles observed at different locations (the parallaxes), the knowledge of the length of the basis of the triangle and some trigonometry. Further away one uses the measure of the luminosity of “standard candles”, and eventually the red shift of some specific light waves. This is the meaning of the “cosmic distance ladder” used in Astrophysics. So, measures of spatial location rely essentially on measures of angles, and one measure of distance, which is established from some phenomena, according to precise protocols based on conventions about the relation between the distance and the phenomenon which is observed. The key is that, on the scale where two methods are applicable, the measures of distances are consistent.

For the temporal location one uses the coincidence with any agreed upon event. For millennia men used the position of celestial bodies for this purpose. Say ”See you at Stonehenge at the spring’s equinox” and you will be understood. Of course one can use a clock, but the purpose of a clock is to measure elapsed time, so one needs a clock and a starting point, which are agreed upon, to locate an event in time. So an observer can locate in time any event which occurs at his place. Are deemed occurring at the time of the observer events that he can see directly, and for events occurring beyond that, the observer accounts for a delay due to the transmission of his perception of the event, based on a convention for the speed of the signal. This speed can be measured itself, for not too far away events, either by a direct communication with a distant observer, or by bouncing a signal on a object at the distant location. But farther away the speed of transmission is set conventionally. Actually the physical support of the signal does not matter much as long as it is efficient, and for the measure of the temporal location can rely on any convention, there is no need for a physical assumption as the constancy of the speed of light.

The measures of location, in time and space, are so based on conventions. This is not an issue, as long as the protocols are precise, and the measures consistent : the purpose of the measures is to be able to identify efficiently an event. One does that with 3 spatial coordinates, and 1 coordinate for the time, organized in charts combining in a consistent way measures done according to different, agreed upon procedures. The key point is that the charts are compatible : when it is possible to proceed to the measures for the same event by different procedures, there is a way to go from one measure to another. And this enables to extend the range of the chart by applying conventions, such as in the cosmic ladder.

These procedures describe a manifold, a mathematical structure seen in the 2nd Chapter. A set of charts covering a domain constitutes an atlas. There are mathematical functions, transition maps, which relate the coordinates of the same point in different charts. A collection of compatible atlas over a set M defines the structure of a manifold. The coordinates represent nothing more than the measures which can be done, and the knowledge of the protocols is sufficient.

This leads to our first proposition :

Proposition 34 *The Universe can be represented as a four dimensional real manifold M*

The charts define over M a topology, deduced from the vector space. The manifold is differentiable (resp. smooth) if the transition maps are differentiable (resp.smooth) (Maths.15.1.1).

In Galilean Geometry the manifold is the product of $\mathbb{R}$ with a 3 dimensional affine space, and in SR this is a 4 dimensional affine space (affine spaces have a manifold structure).

We will limit ourselves to an area Ω of the universe, which can be large, where there is no singularity such as black hole, so that one can assume that one chart suffices. We will represent such a chart as a map :

$\varphi_M : \mathbb{R}^4 \to \Omega :: \varphi_M\left(\xi^0, \xi^1, \xi^2, \xi^3\right) = m$

which is assumed to be bijective and smooth, where $\xi = \left(\xi^0, \xi^1, \xi^2, \xi^3\right)$ are the coordinates of m in the chart φ_M.

We will assume that Ω is a relatively compact open in M, so that the manifold structure on M is the same as on Ω, and Ω is bounded.

A change of chart is represented by a bijective smooth map (the transition map) :

$\chi : \mathbb{R}^4 \to \mathbb{R}^4 :: \eta^\alpha = \chi^\alpha\left(\xi^0, \xi^1, \xi^2, \xi^3\right)$

such that the new map $\widetilde{\varphi}_M$ and the initial map φ_M locate the same point :

$\widetilde{\varphi}_M\left(\chi^\alpha\left(\xi^0, \xi^1, \xi^2, \xi^3\right), \alpha = 0, ..3\right) = \varphi_M\left(\xi^0, \xi^1, \xi^2, \xi^3\right)$

Notice that there is no algebraic structure on M : $am + bm'$ has no meaning. This is illuminating in GR, but still holds in SR or Galilean Geometry. There is a clear distinction between coordinates, which are scalars depending on the choice of a chart, and the point they locate on the manifold (affine space or not).

3.1.2 The tangent vector space

Spatial locations rely heavily on the measures of angles with respect to fixed directions. At any point there is a set of spatial directions, corresponding to small translations in one of the coordinates. And the time direction is just the translation in time for an observer who is spatially immobile. There is the same construct in Mathematics.

Mathematically at any point of a manifold one can define a set which has the structure of a vector space, with the same dimension as M. The best way to see it is to differentiate the map φ_M with respect to the coordinates (this is close to the mathematical construct). To any vector $u \in \mathbb{R}^4$ is associated the vector $u_m = \sum_{\alpha=0}^3 u^\alpha \partial_\alpha \varphi_M\left(\xi^0, \xi^1, \xi^2, \xi^3\right)$ which is denoted $u_m = \sum_{\alpha=0}^3 u^\alpha \partial\xi_\alpha$.

The basis $\left(\partial\xi_\alpha\right)_{\alpha=0}^3$ associated to a chart, called a **holonomic basis**, depends on the chart, but the vector space at m denoted T_mM does not depend on the chart. With this vector space structure one can define a dual space T_mM^* and holonomic dual bases denoted $d\xi^\alpha$ with : $d\xi^\alpha\left(\partial\xi_\beta\right) = \delta_\beta^\alpha$, and any other tensorial structure (see Math.16).

In the definition of the holonomic basis the tangent space is generated by small displacements along one coordinate, around a point m. So, physically, locally the manifold is close to an affine space with a chosen origin m, and locally GR and SR look the same. This is similar to what we see on Earth : locally it looks flat.

However there are essential distinctions between coordinates, used to measure the location of a point in a chart, and components, used to measure a vectorial quantity with respect to a basis. Points and vectors are geometric objects, whose existence does not depend on the way they are measured. However a point on a manifold does not have an algebraic structure attached (the combination $am + bm'$ has no meaning), meanwhile a vector belongs to a vector space : one can combine vectors. Some physical properties of objects can be represented by vectors, other cannot, and the distinction comes from the fundamental assumptions of the theory. It is enshrined in the theory itself. From the construct of the tangent space one sees that any quantity defined as a derivative of another physical quantity with respect to the coordinates is vectorial.

The vector spaces T_mM depend on m, and there is no canonical (meaning independent of the choice of a specific tool) procedure to compare vectors belonging to the tangent spaces at two different points. These vectors u_m can be considered as a couple of a location m and a vector u,which can be defined in a holonomic basis or not, and all together they constitute the tangent bundle TM. Notably there is no physical mean to measure a change in the vectors of a holonomic basis with time : it would require to compare $\partial\xi_\alpha$ at to different locations $m, m' \in M$. But, because there are maps to go from the coordinates in a chart to the coordinates in another chart, there are maps which enable to compute the components of vectors in the holonomic bases of different charts, at the same point.

However because the manifolds are actually affine spaces, in SR and Galilean Geometry the tangent spaces at different points share the same structure (which is the underlying tangent vector space), and only in these cases they can be assimilated to $\mathbb{R}^4$. This is the origin of much confusion on the subject, and the motivation to start in the GR context where the concepts are clearly differentiated.

3.1.3 Vector fields

A vector field on M is a map : $V : M \to TM :: V(m) = \sum_{\alpha=0}^{3} v^\alpha (m) \partial\xi_\alpha$ which associates to any point m a vector of the tangent space T_mM. The vector does not depend on the choice of a basis or a chart, so its components change in a change of chart as (Math.16.1.2) :

$v^\alpha (m) \to \widetilde{v}^\alpha (m) = \sum_{\beta=0}^{3} [J(m)]_\beta^\alpha v^\beta (m)$

where $[J(m)] = \left[\frac{\partial\eta^\alpha}{\partial\xi^\beta} (m) \right]$ is the 4x4 matrix called the jacobian

Similarly a one form on M is a map $\varpi : M \to TM^* :: \varpi(m) = \sum_{\alpha=0}^{3} \varpi_\alpha (m) d\xi^\alpha$ and the components change as :

$\varpi_\alpha (m) \to \widetilde{\varpi}_\alpha (m) = \sum_{\beta=0}^{3} [K(m)]_\alpha^\beta \varpi_\beta (m)$ and $[K(m)] = [J(m)]^{-1}$

The sets of vector fields, denoted $\mathfrak{X}(TM)$,and of one forms, denoted $\mathfrak{X}(TM^*)$ or $\Lambda_1 (M; \mathbb{R})$ are infinite dimensional vector spaces (with pointwise operations).

A **curve** on a manifold is a one dimensional submanifold : this is a geometric structure, and there is a vector space associated to each point of the curve, which is a one dimensional vector subspace of T_mM.

A **path** on a manifold is a map : $p : \mathbb{R} \to M :: m = p(\tau)$ where p is a differentiable map such that $p'(\tau) \neq 0$. Its image is a curve L_p, and p defines a bijection between $\mathbb{R}$ (or any interval of $\mathbb{R}$) and the curve (this is a chart of the curve), the curve is a 1 dimensional submanifold embedded in M. The same curve can be defined by different paths. The tangent is the map : $p'(t) : \mathbb{R} \to T_{p(t)}M :: \frac{dp}{d\tau} \in T_{p(\tau)}L_p$. In a change of parameter in the path : $\widetilde{\tau} = f(\tau)$ (which is a change of chart) for the same point : $m = \widetilde{p}(\widetilde{\tau}) = p(f(\tau))$ the new tangent vector is proportional to the previous one : $\frac{dm}{d\tau} = \frac{d\widetilde{p}}{d\widetilde{\tau}} \frac{d\widetilde{\tau}}{d\tau} \Leftrightarrow \frac{dm}{d\widetilde{\tau}} = \frac{1}{f'} \frac{dm}{d\tau}$

For any smooth vector field there is a collection of smooth paths (the **integrals** of the field) such that the tangent at any point of the curve is the vector field. There is a unique **integral line** which goes through a given point. The **flow** of a vector field V is the map (Math.14.3.5):

$\Phi_V : \mathbb{R} \times M \to M :: \Phi_V(\tau, a)$ such that $\Phi_V(., a) : \mathbb{R} \to M :: m = \Phi_V(\tau, a)$ is the integral path going through a and $\Phi_V(., a)$ is a local diffeomorphism :

$$\left[\begin{array}{l} \forall \theta \in \mathbb{R} : \frac{\partial}{\partial \tau}\Phi_V(\tau, a)|_{\tau=\theta} = V(\Phi_V(\theta, a)) \\ \forall \tau, \tau' \in \mathbb{R} : \Phi_V(\tau + \tau', a) = \Phi_V(\tau, \Phi_V(\tau', a)) \\ \Phi_V(0, a) = a \\ \forall \tau \in \mathbb{R} : \Phi_V(-\tau, \Phi_V(\tau, a)) = a \end{array}\right] \tag{3.1}$$

Notice that, for a given vector field, the parameter τ is defined up to a constant, so *it is uniquely defined* with the condition $\Phi_V(0, a) = a$.

In general the flow is defined only for an interval of the parameter, but this restriction does not exist if Ω is relatively compact.

A map $f : C \to E$ from a curve to a Banach vector space E can be extended to a map $F : \Omega \to E$ (Maths.1467). So any smooth path can be considered as the integral of some vector field (not uniquely defined), and it is convenient to express a path as the flow of a vector field.

3.1.4 Fundamental symmetry breakdown

The idea that the Universe could be 4 dimensional is not new. R.Penrose remarked in his book "The road to reality" that Galileo considered this possibility. The true revolution of Relativity has been to acknowledge that, if the physical universe is 4 dimensional, it becomes necessary to dissociate the abstract representation of the world, the picture given by a mathematical model, from the actual representation of the world as it can be seen through measures. And this dissociation goes through the introduction of a new object in Physics : the observer. Indeed, if the physical Universe is 4 dimensional, the location of a point is absolute : there is a unique material body, in space and time, which can occupy a location. Then, does that mean that past and future exist together ? Can we say that this apple, which is falling, is somewhere in the Universe, still on the tree ? To avoid the conundrum and all the paradoxes that it entails, the solution is to acknowledge that, if there is a unique reality, actually the reality which is scientifically accessible, because it enables experiments and measures, is specific : it depends on the observer. This does not mean that it would be wrong to represent the reality in its entirety, as it can be done with charts, frames or other abstract mathematical objects. They are necessary to give a consistent picture, and more bluntly, to give a picture that is accessible to our mind. But we cannot identify this abstract representation, common to everybody, with the world as it is. This is one of the reasons that motivate the introduction of Geometry in this book through GR : it is common to introduce subtle concepts such as location and velocity through a frame, which is evoked in passing, as if it was obvious, standing somewhere at the disposition of the public. There is nothing like this. I can build my frame, my charts, and from there conceive that it can be extended, and compared to what other Physicists have done. But comparison requires first dissociation, and this is more easily done in a context to which we are less used to, by years of schematic representations.

The four coordinates are not equivalent : the measure of the time ξ^0 cannot be done with the same procedures as the other coordinates, and one cannot move along in time : one cannot survey time. This is the fundamental symmetry breakdown.

The time coordinate of an event can be measured, by conventional procedures which relate the time on the clock (whatever it is) of a given observer to the time at which a distant event has occurred. So one assumes that *a given observer* can tell if two events A, B occur in his present time (they are simultaneous), and that the relation "two events are simultaneous" is a relation of equivalence between events. Then the observer can label each class of equivalence of events by the time of his clock. Which can be expressed by telling that for each observer, there is a function : $f_o : M \to \mathbb{R} :: f_o(m) = t$ which assigns a time t, *with respect to the clock of the observer*, at any point of the universe (or at least Ω). The points : $\Omega(t) = \{m = f_o(t), m \in \Omega\}$ correspond to the **present** of the observer. No assumption is made about the clock, and different clocks can be used, with the condition that, as for any chart, it is possible to convert the time given by a clock to the time given by another clock (both used by the same observer).

In Galilean Geometry instantaneous communication is possible, so it is possible to define a universal time, to which any observer can refer to locate his position, and the present does not depend on the observer. The manifold M can be assimilated to the product $\mathbb{R} \times \mathbb{R}^3$. The usual representation of material bodies moving in the same affine space is a bit misleading, actually one should say that this affine space $\mathbb{R}^3(t)$ changes continuously, in the same way, for everybody. Told this way we see that Galilean Geometry relies on a huge assumption about the physical universe.

In Relativist Geometry instantaneous communication is impossible, so it is impossible to synchronize all the clocks. However a given observer can synchronize the clocks which correspond to his present, this is the meaning of the function f_o.

Whenever there is, on a manifold, a map such that f_o, with $f_o'(m) \neq 0$, it defines on M a folliation : there is a collection of hypersurfaces (3 dimensional submanifolds) $\Omega_3(t)$,and the vectors u of the tangent spaces on $\Omega_3(t)$ are such that $f_o'(m)u = 0$, meanwhile the vectors which are transversal to $\Omega_3(t)$ (corresponding to paths which cross the hypersurface only once) are such that $f_o'(m)u > 0$ for any path with t increasing. So there are two faces on $\Omega_3(t)$: one for the incoming paths, and the other one for the outgoing paths. The hypersurfaces $\Omega_3(t)$ are diffeomorphic : they can be deduced from each other by a differentiable bijection, which is the flow of a vector field. Conversely if there is such a folliation one can define a unique function f_o with these properties (Maths.1502[1]). The successions of present "spaces" for any observer is such a folliation, so our representation is consistent. And we state :

Proposition 35 *For any observer there is a function*

$$f_o : M \to \mathbb{R} :: f_o(m) = t \text{ with } f_o'(m) \neq 0 \tag{3.2}$$

which defines in any area Ω of the Universe a folliation by hypersurfaces

$$\Omega_3(t) = \{m = f_o(t), m \in \Omega\} \tag{3.3}$$

which represents the location of the events occurring at a given time t on his clock.

An observer can then define a chart of M, by taking the time on his clock, and the coordinates of a point x in the 3 dimensional hypersurfaces $\Omega_3(t)$: it would be some map : $\varphi : \mathbb{R} \times \Omega_3(0) \to M :: m = \varphi(t, x)$ however we need a way to build consistently these spatial coordinates, that is to relate $\varphi(t, x)$ to $\varphi(t', x)$.

[1]Actually this theorem, which has far eaching consequences, is new and its proof, quite technical is given in my book.

3.1.5 Trajectories of material bodies

The Universe is a container where physical objects live, and the manifold provides a way to measure a location. This is a 4 dimensional manifold which includes the time, but that does not mean that everything is frozen on the manifold : *the universe does not change, but its content changes.* As bodies move in the universe, their representation are paths on the manifold. And the fundamental symmetry breakdown gives a special meaning to the coordinate with respect to which the changes are measured. *Time is not only a parameter to locate an event, it is also a variable which defines the rates of change in the present of an observer.*

Material bodies and particles

The common definition of a material body in Physics is that of a set of material points which are related. A **material point** is assumed to have a location corresponding to a point of the manifold. According to the relations between material points of the same body we have rigid solids (the distance between two points is constant), deformable solids (the deformation tensor is locally given by the matrix of the transformation of a frame), fluids (the speed of material points are given by a vector field). These relations are formulated by phenomenological laws, they are essential in practical applications, but not in a theoretical study. So we will consider material bodies which have no internal structures, or whose internal structure can be neglected, that we will call **particles**. A particle then can be an electron, a nucleus, a molecule, or even a star system, according to the scale of the study. As in Mechanics a particle is a material point, and its location can be assimilated to a point from a geometrical point of view.

World line and proper time

As required in any scientific theory a particle must be defined by its properties, and the first is that it occupies a precise location at any time. The successive locations of the material body define a curve and the particle travels on this curve according to a specific path called its world line. Any path can be defined by the flow of a vector such that the derivative with respect to the parameter is the tangent to the curve. The parameter called the proper time is then defined uniquely, up to the choice of an origin. The derivative with respect to the proper time is called the velocity. By definition *this is a vector,* defined at each point of the curve, and belonging to the tangent space to M. So the velocity has a definition which is independent of any basis.

Remark : For brevity I will call velocity the 4-vector, also usually called 4-velocity, and spatial speed the common 3 vector.

Observers are assumed to have similarly a world line and a proper time (they have other properties, notably they define a frame).

To sum up :

Definition 36 *Any particle or observer travels in the universe on a curve according to a specific path , $p : \mathbb{R} \to M :: m = p(\tau)$ called the **world line**, parametrized by the **proper time** τ, defined uniquely up to an origin. The derivative of the world line with respect to the proper time is a vector, the **velocity,** u. So that :*

$$\begin{bmatrix} u(\theta) = \frac{dp}{d\tau}|_{\tau=\theta} \in T_{p(\theta)}M \\ p(\tau) = \Phi_u(\tau, a) \;\; with\; a = \Phi_u(0, a) = p(0) \end{bmatrix} \tag{3.4}$$

Observers are assumed to have clocks, that they use to measure their temporal location with respect to some starting point. The basic assumption is the following :

Proposition 37 *For any observer his proper time is the time on his clock.*

So the proper time of a particle can be seen formally as the time on the clock of an observer who would be attached to the particle.

We will strive to denote t the time of an observer (specific to an observer) and τ any other proper time. So for a given observer :

$t = \tau$

$p_o : \mathbb{R} \to M :: m = p_o(t)$

$u(\theta) = \frac{dp}{dt}|_{t=\theta} \in T_{p(\theta)}M$

$p_o(t) = \Phi_u(\tau, a)$ with $a = \Phi_u(0, a) = p(0)$

The observer uses the time on his clock to locate temporally any event : this is the purpose of the function f_o and of the folliation $\Omega_3(t)$. The curve on which any particle travels meets only once each hypersurface $\Omega_3(t)$: it is seen only once. This happens at a time t :

$f_o(p(\tau)) = t = f_o(\Phi_u(\tau, a))$

So there is some relation between t and the proper time τ. It is specific, both to the observer and to the particle. It is bijective and both increases simultaneously, so that : $\frac{d\tau}{dt} > 0$.

The travel of the particle on the curve can be represented by the time of an observer. We will call then this path a **trajectory**.

With this assumption each observer can build a chart. On some hypersurface $\Omega_3(0)$ representing the space of the observer at a time $t = 0$ he chooses a chart identifying each point x of $\Omega_3(0)$ by 3 coordinates ξ^1, ξ^2, ξ^3, using the methods to measure spatial distances described previously, and $m = \varphi_o(t, \xi^1, \xi^2, \xi^3)$ is a chart of the area $\Omega \subset M$ spanned by the $\Omega_3(t)$. Each point $m(t) = \varphi_o(t, \xi^1, \xi^2, \xi^3)$ corresponds to the trajectory of a material body or of an observer which would stand still at x. We will call this kind of chart a **standard chart** for the observer. It relies on the choice of chart of $\Omega_3(0)$, that is a set of procedures to measure a spatial location (so several compatible charts can be used) and a clock or any procedure to identify a time. A standard chart is specific to each observer and is essentially fixed.

Even if two observers can compare the measures of spatial locations, actually so far we cannot go further : the hypersurfaces $\Omega_3(t)$ are defined by the function f_o and, a priori, are specific to each observer. Moreover a clock measures the elapsed time. It seems legitimate to assume that, in the procedure, one chooses clocks which run at the same rate. But, to do this, one needs some way to compare this rate, that is a scalar measure of the velocity $\frac{d}{d\tau}p_o(\tau)$. But, as velocities are 4 dimensional vectors, one needs a special scalar product.

The essential feature of proper time is more striking when one considers particles. They should be located at some point of M : they are not spread over all they world line, their location varies along their world line with respect to the parameter τ, their proper time. So their location is definite, but with respect to a parameter τ which is specific to each particle : there is a priori no universal time[2] and no way to tell where, at some time, are all the particles ! An observer can locate a particle which is in his "present", and so identify specific particles, but this is specific to each observer.

3.1.6 Causal structure

The Principle of causality states that there is some order relations between events. This relation is not total : some events are not related. In the Relativist Geometry it can be stated as a relation between locations in the Universe : a binary relation between two points (A, B).

[2]The popular cosmological models assume the existence of such a universal time, but this is an assumption, and not implemented in the study of specific particles.

The function f_o of an observer provides such a relation : it suffices to compare $f_o(A), f_o(B)$: B follows A if $f_o(B) > f_o(A)$ and is simultaneous to A if $f_o(B) = f_o(A)$. For a relation between points it is natural to look at curves joining the points. For a path $p \in C_1([0,1];M)$ such that $p(0) = A, p(1) = B$ one can compute $f_o(p(\tau))$. If the function is increasing then one can say that B follows A, and this is equivalent to $f'_o(p(\tau))\frac{dp}{d\tau} > 0$. And we can say that the vector $u = \frac{dp}{d\tau} \in T_{p(\tau)}M$ is future oriented for the observer if $f'_o(p(\tau))u > 0$. We have the same conclusion for any vector at a point $m \in M$ which belongs to one of the hypersurfaces $\Omega_3(t)$ of an observer : if it is transversal it can be oriented towards the future by $f'_o(m)u$, and any curve can be similarly oriented at any point, but the orientation is not necessarily constant. The classification of the curves which have a constant orientation is a topic of algebraic geometry, but here there is a more interesting issue : the Principle of Causality should be met for any observer. We can study this issue by looking at vectors u at a given point m. The derivative $f'_o(m)$ is just a covector $\lambda \in T_mM^*$. The function : $B : T_mM^* \times T_mM \to \mathbb{R} :: B(\lambda, u) = \lambda(u)$ is continuous in both variables (T_mM^*, T_mM are finite dimensional vector spaces and have a definite topology). For a given λ if $\lambda(u) > 0$ then $\lambda(-u) < 0$, and we have a partition of T_mM in 3 connected components : future oriented vectors $\lambda(u) > 0$, past oriented vectors $\lambda(u) < 0$, null vectors $\lambda(u) = 0$. This partition of T_mM should hold for any observer. The implementation of the Principle of Causality in Relativist Geometry leads to state that, at each point m, there is a set C_+ of vectors future oriented for all observers, and that vectors which do not belong to C_+ are not future oriented for any observer. The opposite set C_- is the set of past oriented vectors. C_+ is a convex open half cone : if for an observer u, v are future oriented, then $\alpha u + (1-\alpha)v$ is future oriented.

For any observer, there is a hyperplan $H_o(m)$ passing by m, which separates C_+, C_- : take $f'_o(m) \in T_mM^*$

$\forall u \in C_-, v \in C_+ : f'_o(m)(u) < 0 < f'_o(m)(v) \Rightarrow \sup_{u \in C_-} f'_o(m)(u) \leq \inf_{v \in C_+} f'_o(m)(v)$

Moreover this hyperplan is tangent to his hypersurface $\Omega_3(t)$ passing by m.

So any observer can choose a basis of T_mM comprised of 3 vectors $(\varepsilon_i)_{i=1}^3$ belonging to $H_o(m)$, that is his "space". Then $f'_o(m)(\varepsilon_i) = 0, i = 1,2,3$ because the vectors are tangent to $\Omega_3(t)$. With any other vector ε_0 as 4th vector of his basis,

$f'_o(m)(u) = f'_o(m)\left(\sum_{i=0}^3 u^i \varepsilon_i\right) = u^0 f'_o(m)(\varepsilon_0)$

To have a consistent result for this function, that is to be able to distinguish a past from a future oriented vector, the observer must choose $\varepsilon_0 \in C_+$, and this choice is always possible by taking his velocity as ε_0.

And this choice can be done in a consistent manner for any observer. Any "physical" basis chosen by an observer is comprised of 3 spatial vectors, which do not belong to C_+ and the 4th vector belong to C_-. This holds for the holonomic basis induced by a standard chart.

The function $B(\lambda, u)$ is defined all over M, does not depend on the observer, it is a bilinear map, so this is a tensor field $B \in TM^* \otimes TM$. In any basis it is expressed at a point by a 4×4 matrix, and this matrix can be considered as the matrix of a bilinear form, from which a symmetric bilinear form can be computed, and so a metric on TM. However we see that there are vectors such that $B(u,u) = 0$. This metric cannot be definite positive.

A manifold is usually not isotropic : not all directions are equivalent. The fundamental symmetry breakdown introduces a first anisotropy, specific to each observer, and we see that actually it goes deeper, because it is common to all observers and not all vectors representing a translation in time are equivalent : C_+ is a half cone and not a half space.

So the Principle of Causality leads to assume that there is an additional structure in the Universe. This causal structure is usually defined through the propagation of light : a region B

is temporally dependant from a region A if any point of B can be reached from A by a future oriented curve. This is the domain of nice studies (see Wald), but there is no need to involve the light, the causal structure exists at the level of the tangent bundle, its definition does not need the existence of a metric, but clearly leads to assume that there is a metric and that this metric is not definite positive.

3.1.7 Metric on the manifold

Lorentz metric

A scalar product is defined by a bilinear symmetric form g acting on vectors of the tangent space, at each point of the manifold, thus by a tensor field called a **metric**. In a holonomic basis g reads :

$$g(m) = \sum_{\alpha\beta=0}^{3} g_{\alpha\beta}(m)\, d\xi^{\alpha} \otimes d\xi^{\beta} \text{ with } g_{\alpha\beta} = g_{\beta\alpha} \tag{3.5}$$

The matrix of g is symmetric and invertible, if we assume that the scalar product is not degenerate. It is diagonalizable, and its eigen values are real. One wants to account for the symmetry breakdown and the causal structure, so these eigen values cannot have all the same sign (a direction is privileged). One knows that the hypersurface $\Omega_3(t)$ are Riemannian : there is a definite positive scalar product (acting on the 3 dimensional vector space tangent to $\Omega_3(t)$), and that transversal vectors correspond to the velocities of material bodies. So there are only two solutions for the signs of the eigen values of $[g(m)]$: either (-,+,+,+) or (+,-,-,-) which provides both a **Lorentz metric**. The scalar product, in an orthonormal basis $(\varepsilon_i)_{i=0}^{3}$ at m reads :

$$\begin{bmatrix} \text{signature } (3,1) : \langle u,v\rangle = u^1v^1 + u^2v^2 + u^3v^3 - u^0v^0 \\ \text{signature } (1,3) : \langle u,v\rangle = -u^1v^1 - u^2v^2 - u^3v^3 + u^0v^0 \end{bmatrix} \tag{3.6}$$

Such a scalar product defines by restriction on each hypersurface $\Omega_3(t)$ a positive or a negative definite metric, which applies to spatial vectors (tangent to $\Omega_3(t)$) and provides, up to sign, the usual euclidean metric. So that both signatures are acceptable.

Which leads to :

Proposition 38 *The manifold M representing the Universe is endowed with a non degenerate metric, called the* ***Lorentz metric****, with signature either (3,1) of (1,3) defined at each point.*

This reasoning is a legitimate assumption, which is consistent with all the other concepts and assumptions, notably the existence of a causal structure, this is not the proof of the existence of such a metric. Such a proof comes from the formula in a change of frames between observers, which can be checked experimentally.

Notice that on a finite dimensional, connected, Hausdorff manifold, there is always a definite positive metric (Maths.1385). There is no relation between this metric and a Lorentz metric. Not all manifolds can have a Lorentz metric, the conditions are technical (see Giachetta p.224 for more) but one can safely assume that they are met in a limited region Ω.

A metric is represented at each point by a tensor, whose value can change with the location. One essential assumption of General Relativity is that, meanwhile the container M is fixed, and so the chart and its holonomic basis are fixed geometric representations without specific physical meaning, the metric is a physical object and can vary at each point according to specific physical laws. The well known deformation of the space-time with gravity is expressed, not in the structure of the manifold (which is invariant) but in the value of the metric at each point. However the metric conserve always its basic properties - it is a Lorentz metric.

Gauge group

The existence of a metric implies that, at any point, there are orthonormal bases $(\varepsilon_i)_{i=0}^{3}$ with the property :

Definition 39 *$\langle \varepsilon_i, \varepsilon_j \rangle = \eta_{ij}$ for the signature (3,1) and $\langle \varepsilon_i, \varepsilon_j \rangle = -\eta_{ij}$ for the signature (1,3)*

with the matrix $[\eta]$

Notation 40 *In any orthonormal basis ε_0 denotes the time vector.*
$\langle \varepsilon_0, \varepsilon_0 \rangle = -1$ *if the signature is* $(3,1)$
$\langle \varepsilon_0, \varepsilon_0 \rangle = +1$ *if the signature is* $(1,3)$

Notation 41 $[\eta] = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ *whatever the signature*

An orthonormal basis, at each point, is a **gauge**. The choice of an orthonormal basis depends on the observer : he has freedom of gauge. One goes from one gauge to another by a linear map χ which preserves the scalar product. They constitute a group, called the **gauge group**. In any basis these maps are represented by a matrix $[\chi]$ such that :

$$[\chi]^t [\eta] [\chi] = [\eta] \tag{3.7}$$

The group denoted equivalently $O(3,1)$ or $O(1,3)$, does not depend on the signature (replace $[\eta]$ by -$[\eta]$).

$O(3,1)$ is a 6 dimensional Lie group with Lie algebra $o(3,1)$ whose matrices $[h]$ are such that :

$$[h]^t [\eta] + [\eta] [h] = 0 \tag{3.8}$$

(Maths.24.5.3). The Lie algebra is a vector space and we will use the basis :

$$[\kappa_1] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}; [\kappa_2] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}; [\kappa_3] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
$$[\kappa_4] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; [\kappa_5] = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; [\kappa_6] = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

so that any matrix of $o(3,1)$ can be written :
$[\kappa] = [J(r)] + [K(w)]$ with

$$[J(r)] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -r_3 & r_2 \\ 0 & r_3 & 0 & -r_1 \\ 0 & -r_2 & r_1 & 0 \end{bmatrix}; [K(w)] = \begin{bmatrix} 0 & w_1 & w_2 & w_3 \\ w_1 & 0 & 0 & 0 \\ w_2 & 0 & 0 & 0 \\ w_3 & 0 & 0 & 0 \end{bmatrix}$$

The exponential of these matrices read (Maths.493) :
$\exp[K(w)] = I_4 + \frac{\sinh\sqrt{w^t w}}{\sqrt{w^t w}} K(w) + \frac{\cosh\sqrt{w^t w} - 1}{w^t w} K(w)K(w)$

$$\exp[K(w)] = \begin{bmatrix} \cosh\sqrt{w^t w} & w^t \frac{\sinh\sqrt{w^t w}}{\sqrt{w^t w}} \\ w\frac{\sinh\sqrt{w^t w}}{\sqrt{w^t w}} & I_3 + \frac{\cosh\sqrt{w^t w}-1}{w^t w} ww^t \end{bmatrix}$$

$$\exp[J(r)] = I_4 + \frac{\sin\sqrt{r^t r}}{\sqrt{r^t r}} J(r) + \frac{1-\cos\sqrt{r^t r}}{r^r r} J(r)J(r) = \begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix}$$

where R a 3×3 matrix of $O(3)$.

The group $O(3)$ has two connected components : the subgroup $SO(3)$ with determinant = 1, and the subset $O_1(3)$ with determinant -1.

$O(3,1)$ has four connected components which can be distinguished according to the sign of the determinant and their projection under the compact subgroup $SO(3) \times \{I\}$.

Any matrix of $SO(3,1)$ can be written as the product : $[\chi] = \exp[K(w)]\exp[J(r)]$ (or equivalently $\exp[J(r')]\exp[K(w')]$). So we have the 4 cases :

- $SO_0(3,1)$: with determinant 1: $[\chi] = \exp K(w) \times \begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix}$
- $SO_1(3,1)$: with determinant 1: $[\chi] = \exp K(w) \times \begin{bmatrix} -1 & 0 \\ 0 & -R \end{bmatrix}$
- $SO_2(3,1)$ with determinant = -1: $[\chi] = \exp K(w) \times \begin{bmatrix} -1 & 0 \\ 0 & R \end{bmatrix}$
- $SO_3(3,1)$ with determinant = -1: $[\chi] = \exp K(w) \times \begin{bmatrix} 1 & 0 \\ 0 & -R \end{bmatrix}$

where R a 3×3 matrix of $SO(3)$, so that $-R \in O_1(3)$

Orientation and time reversal

Any finite dimensional vector space is orientable. A manifold is orientable if it is possible to define a consistent orientation of its tangent vector spaces, and not all manifolds are orientable. If it is endowed with a metric then the map : $\det g : M \to \mathbb{R}$ provides an orientation function (its sign changes with the permutation of the vectors of a holonomic basis) and the manifold is orientable.

But on a 4 dimensional vector space one can define other operations, of special interest when the 4 dimensions have not the same properties. For any orthonormal basis $(\varepsilon_i)_{i=0}^3$:

space reversal is the change of basis :
$i = 1,2,3 : \widetilde{\varepsilon}_i = -\varepsilon_i$
$\widetilde{\varepsilon}_0 = -\varepsilon_0$

time reversal is the change of basis :
$i = 1,2,3 : \widetilde{\varepsilon}_i = \varepsilon_i$
$\widetilde{\varepsilon}_0 = -\varepsilon_0$

These two operations change the value of the determinant, so they are not represented by matrices of $SO(3,1)$:

space reversal matrix : $S = \begin{bmatrix} 1 & 0 \\ 0 & -I_3 \end{bmatrix}$

time reversal matrix : $T = \begin{bmatrix} -1 & 0 \\ 0 & I_3 \end{bmatrix}$

$ST = -I_4$

The matrices of the subgroups $SO_k(3,1), k = 1,2,3$ are generated by the product of any element of $SO_0(3,1)$ by either S or T.

Is the universe orientable ? Following our assumption, if there is a metric, it is orientable. However one can check for experimental proofs. In a universe where all observers have the same time, the simple existence of stereoisomers which do not have the same chemical properties suffices to answer positively : we can tell to a distant observer what we mean by "right" and "left" by agreeing on the property of a given product. In a space-time universe one needs a process with an outcome which discriminates an orientation. All chemical reactions starting with a balanced mix of stereoisomers produce an equally balanced mix (stereoisomers have the same level of energy). However there are experiments involving the weak interactions (CP violation symmetry in the decay of neutral kaons) which show the required property. So we can state that the 4 dimensional universe is orientable, and then we can distinguish orientation preserving gauge transformations.

A change of gauge, physically, implies some transport of the frame (one does not jump from one point to another) : we have a map : $\chi : I \to SO(3,1)$ such that at each point of the path $p_o : I \to M$ defined on a interval I of $\mathbb{R}$, $\chi(t)$ is an isometry. The path which is followed matters. In particular it is connected. The frame $(\varepsilon_i)_{i=0}^3$ is transported by : $\widetilde{\varepsilon}_i(\tau) = \chi(t)\varepsilon_i(0)$. So $\{[\chi(\tau)], t \in I\}$, image of the connected interval I by a continuous map is a connected subset of $SO(3,1)$, and because $\chi(0) = Id$ it must be the component of the identity. So the right group to consider is the **connected component of the identity** $SO_0(3,1)$

Time like and space like vectors

The causal structure is then fully defined by the metric.

At any point m one can discriminate the vectors $v \in T_mM$ according to the value of the scalar product $\langle v, v\rangle$.

Definition 42 ***Time like*** *vectors are vectors v such that* $\langle v, v\rangle < 0$ *with the signature (3,1) and* $\langle v, v\rangle > 0$ *with the signature (1,3)*

Space like *vectors are vectors v such that* $\langle v, v\rangle > 0$ *with the signature (3,1) and* $\langle v, v\rangle < 0$ *with the signature (1,3)*

Moreover the subset of time like vectors has two disconnected components (this is no longer true in universes with more than one "time component" (Maths.307)). So one can discriminate these components and, in accordance with the assumptions about the velocity of material bodies, it is logical to consider that their velocity is **future oriented**. And one can distinguish gauge transformations which preserve this time orientation.

Definition 43 *We will assume that the future orientation is given in a gauge by the vector* ε_0. *So a vector u is time like and future oriented if :*

$\langle u, u\rangle < 0, \langle u, \varepsilon_0\rangle < 0$ *with the signature (3,1)*

$\langle u, u\rangle > 0, \langle u, \varepsilon_0\rangle > 0$ *with the signature (1,3)*

A matrix $[\chi]$ of $SO_0(3,1)$ preserves the time orientation iff $[\chi]_0^0 > 0$ and this will always happen if $[\chi] = \exp[K(w)]\exp[J(r)]$ that is if $[\chi] \in SO_0(3,1)$.

A gauge transformation which preserves both the time orientation, and the global orientation must preserve also the spatial orientation.

In GR it is common to use "Killing vector fields" : they are vector fields V such that their flow, which is always a diffeomorphism, preserves the scalar product : it is an isometry. This is equivalent to say that the Lie derivative of the metric along V : $\pounds_V g = 0$. They are a crucial tool in all studies focused on the metric, but we will not use them in this book.

3.1.8 Velocities have a constant Lorentz norm

The velocity $\frac{dp_o}{d\tau}$ is a vector which is defined independently of any basis, for any observer it is transversal to $\Omega_3(t)$. It is legitimate to say that it is future oriented, and so it must be time-like. One of the basic assumptions of Relativity is that it has a constant length, as measured by the metric, identical for all observers. So it is possible to use the norm of the velocity to define a standard rate at which the clocks run.

Because the proper time of any material body can be defined as the time on the clock of an observer attached to the body this proposition is extended to any particle.

The time is not measured with the same unit as the lengths, used for the spatial components of the velocity. The ratio ξ^i/t has the dimension of a spatial speed. So we make the general assumption that for any observer or particle the velocity is such that $\left\langle \frac{dp}{d\tau}, \frac{dp}{d\tau} \right\rangle = -c^2$ where τ is the proper time. Notice that c is a constant, with no specific value. This is consistent with the procedures used to measure the time of events occurring at a distant spatial location.

And we sum up :

Proposition 44 *The velocity $\frac{dp}{d\tau}$ of any particle or observer is a time like, future oriented vector with Lorentz norm*

$$\left\langle \frac{dp}{d\tau}, \frac{dp}{d\tau} \right\rangle = -c^2 \tag{3.9}$$

(with signature (3,1) or c^2 with signature (1,3)) where c is a fundamental constant.

3.1.9 Standard chart of an observer

With the previous propositions we can define the standard chart of an observer.

Theorem 45 *For any observer there is a vector field $\mathbf{O} \in \mathfrak{X}(TM)$ which is future oriented, with length $\langle \mathbf{O}(m), \mathbf{O}(m) \rangle = -1$, normal to $\Omega_3(t)$ and such that : $\mathbf{O}(p_0(t)) = \frac{1}{c}\frac{dp_o}{dt}$ where $\frac{dp_o}{dt}$ is the velocity of the observer at each point of his world line.*

Proof. For an observer the function $f_o : \Omega \to \mathbb{R}$ has for derivative a one form $f'_o(m) \neq 0$ such that $\forall v \in T_m\Omega_3(t) : f'_o(m)v = 0$. Using the metric, it is possible to associate to $f'_o(m)$ a vector : $\mathbf{O}(m) = gradf_o : \langle \mathbf{O}(m), v \rangle = f'_o(m)v$ which is unique up to a scalar (Maths.1608). Thus $\mathbf{O}(m)$ is normal to $\Omega_3(t)$. Along the world line of the observer $\mathbf{O}(m)$ is in the direction of the velocity of the observer. And it is always possible to choose $\mathbf{O}(m)$ such that it is future oriented and with length $\langle \mathbf{O}(m), \mathbf{O}(m) \rangle = -1$ ■

As a consequence :

Theorem 46 $\Omega_3(t)$ *are space like hypersurfaces, with unitary normal* $\mathbf{O} \in \mathfrak{X}(TM)$

Using the vector field $\mathbf{O}$, and any chart φ_Ω of $\Omega(0)$ there is a standard chart associated to an observer.

Definition 47 *The **standard chart** on M of any observer is defined as :*
$\varphi_o : \mathbb{R}^4 \to \Omega :: \varphi_o\left(\xi^0, \xi^1, \xi^2, \xi^3\right) = \Phi_O(ct, x)$
$\xi^0 = ct$, $\varphi_\Omega\left(\xi^1, \xi^2, \xi^3\right) = x$ *in any chart of* $\Omega(0)$

c is required in $\Phi_O(ct, x)$ so that :

$$\xi^0 = ct \tag{3.10}$$

which makes all the coordinates homogeneous in units [Length].
The holonomic basis associated to this chart is such that :
$\partial\xi_0 = \frac{\partial\varphi_o}{\partial\xi^0} = \frac{1}{c}\frac{\partial}{\partial t}\Phi_{\varepsilon_0}(ct, x) = \mathbf{O}$

$$\mathbf{O}(m) = \partial\xi_0 \tag{3.11}$$

For any point $m = \varphi_o\left(\xi^0, \xi^1, \xi^2, \xi^3\right) = \Phi_O(ct, x)$ the point x is the point where the integral curve of $\mathbf{O}$ passing by m crosses $\Omega_3(0)$.

So the main characteristic of an observer can be summed in the vector field $\mathbf{O}$ (which is equivalently deduced from the function f_o). From this vector field it is possible to define any standard chart, by choosing a chart on $\Omega_3(0)$. In this construct the spatial location of the observer does not matter any longer : the only restriction is that he belongs to $\Omega_3(t)$ and he follows a trajectory which is an integral curve of the vector field $\mathbf{O}$: $p_o(t) = \varphi_o(t, x_0)$ for some fixed $x_0 \in \Omega_3(0)$.

According to the principle of locality any measure is done locally : the state of any system at t is represented by the measures done over $\Omega_3(t)$. The system itself can be defined as the "physical content" of $\Omega_3(t)$ and its evolution as the set $\{\Omega_3(t), t \in [0, T]\}$. The physical system itself is observer dependant. The vector field $\mathbf{O}$ defines a special chart, but also the system itself. Two observers who do not share the vector field $\mathbf{O}$ do not perceive the same system. So actually this is a limitation of the Principle of Relativity : it holds but only when the observers agree on the system they study. And of course the observers who share the same $\mathbf{O}$ have a special interest.

3.1.10 Trajectory and speed of a particle

A particle follows a world line $q(\tau)$, parametrized by its proper time. Any observer sees only one instance of the particle, located at the point where the world line crosses the hypersurface $\Omega_3(t)$ so we have a relation between τ and t. This relation identifies the respective location of the observer and the particle on their own world lines. With the standard chart of the observer it is possible to measure the velocity of the particle at any location, and of course at the location where it belongs to $\Omega_3(t)$.

The trajectory (parametrized by t) of any particle in the standard chart of an observer is :
$q(t) = \Phi_O(ct, x(t)) = \varphi_o\left(ct, \xi^1(t), \xi^2(t), \xi^3(t)\right)$
By differentiation with respect to t :
$\frac{dq}{dt} = c\mathbf{O}(q(t)) + \frac{\partial}{\partial x}\Phi_O(ct, x(t))\frac{\partial x}{\partial t}$
$\frac{\partial}{\partial x}\Phi_O(ct, x(t))\frac{\partial x}{\partial t} = \sum_{\alpha=1}^{3}\frac{d\xi_\alpha}{dt}\partial\xi_\alpha \in T_m\Omega_3(t)$ so is orthogonal to $\mathbf{O}(q(t))$

Definition 48 *The **spatial speed** of a particle on its trajectory with respect to an observer is the vector of* $T_{q(t)}\Omega_3(t)$ *:*
$\overrightarrow{v} = \frac{\partial}{\partial x}\Phi_O(ct, x(t))\frac{\partial x}{\partial t} = \sum_{\alpha=1}^{3}\frac{d\xi^\alpha}{dt}\partial\xi_\alpha$

Thus for any particle in the standard chart of an observer :

$$V(t) = \frac{dq}{dt} = c\mathbf{O}(q(t)) + \overrightarrow{v} \tag{3.12}$$

For the observer in the standard chart we have :
$\frac{dp_0}{dt} = c\mathbf{O}(p_0(t)) \Leftrightarrow \overrightarrow{v} = 0$
Notice that the velocity, and the spatial speed, are measured *in the chart of the observer at the point $q(t)$ where is the particle.* Because we have defined a standard chart it is possible to measure the speed of a particle located at a point $q(t)$ which is different from the location of the observer. And we can express the relation between τ and t.

Theorem 49 *The proper time τ of any particle and the corresponding time of any observer t are related by :*

$$\frac{d\tau}{dt} = \sqrt{1 - \frac{\|\overrightarrow{v}\|^2}{c^2}} \tag{3.13}$$

where $\overrightarrow{v}$ is the spatial speed of the particle, with respect to the observer and measured in his standard chart.
The velocity of the particle is :

$$\frac{dp}{d\tau} = \frac{1}{\sqrt{1 - \frac{\|\overrightarrow{v}\|^2}{c^2}}}\left(\overrightarrow{v} + c\mathbf{O}(m)\right) \tag{3.14}$$

Proof. i) Let be a particle A with world line :
$p : \mathbb{R} \to M :: m = p(\tau) = \Phi_u(\tau, a)$ with $a = \Phi_u(0, a) = p(0)$
In the standard chart $\Phi_O(ct, x)$ of the observer O its trajectory is :
$q : \mathbb{R} \to M :: m = q(t) = \Phi_O(ct, x(t))$
So there is a relation between t, τ :
$m = p(\tau) = \Phi_u(\tau, a) = q(t) = \Phi_O(ct, x(t))$
By differentiation with respect to t :
$\frac{d}{dt} q(t) = c\mathbf{O}(p_A(t)) + \overrightarrow{v}$
$\frac{dq}{dt} = \overrightarrow{v} + c\mathbf{O}(m)$
$\frac{dq}{dt} = \frac{dp}{d\tau}\frac{d\tau}{dt}$
$\left\langle \frac{dp}{d\tau}, \frac{dp}{d\tau} \right\rangle = -c^2$
$\left\langle \frac{dq}{dt}, \frac{dq}{dt} \right\rangle = -c^2 \left(\frac{d\tau}{dt}\right)^2$
$\left\langle \frac{dq}{dt}, \frac{dq}{dt} \right\rangle = \langle \overrightarrow{v}, \overrightarrow{v} \rangle_3 - c^2$ because $\mathbf{O}(m) \perp \Omega_3(t)$
$\|\overrightarrow{v}\|^2 - c^2 = -c^2 \left(\frac{d\tau}{dt}\right)^2$
and because $\frac{d\tau}{dt} > 0 : \frac{d\tau}{dt} = \sqrt{1 - \frac{\|\overrightarrow{v}\|^2}{c^2}}$
ii) The velocity of the particle is :
$\frac{dp}{d\tau} = \frac{dq}{dt}\frac{dt}{d\tau} = \frac{1}{\sqrt{1 - \frac{\|\overrightarrow{v}\|^2}{c^2}}}\left(\overrightarrow{v} + c\mathbf{O}(m)\right)$ ∎
As a consequence :

$$\|\overrightarrow{v}\|_3 < c \tag{3.15}$$

$V(t) = \frac{dp}{dt}$ is the measure of the motion of the particle with respect to the observer : it can be seen as the relative velocity of the particle with respect to the observer. It involves $\overrightarrow{v}$ which has the same meaning as usual, but we see that in Relativity one goes from the 4 velocity $u = \frac{dp}{d\tau}$ (which has an absolute meaning) to the relative velocity $V(t) = \frac{dp}{dt} = \frac{dp}{d\tau}\frac{d\tau}{dt} = u\sqrt{1 - \frac{\|\overrightarrow{v}\|^2}{c^2}}$ by a scalar. If we have two particles A,B, with their path $q_A(\tau_A), q_B(\tau_B)$ can we define their relative motion, for instance of B relative to A ? The simplest way to do it in relativity is to consider A as an observer, then $V_{B/A}(\tau_A) = \frac{dq_B}{d\tau_A} = u_B\sqrt{1 - \frac{\|\overrightarrow{v}_{B/A}\|^2}{c^2}}$ which is defined *in the chart* associated to the observer A. And the velocity of a particle is a vector localized at the position of the particle.

3.2 FIBER BUNDLES

The location of a particle is absolute : this is the point in the physical Universe that it occupies at some time. But the measure of this location is relative to the observer, starting with the time at which the particle is at a given place. Similarly the velocity of a particle or an observer is absolute : in its definition there is no reference to a chart or a frame. This is an essential point in Relativity. State that the velocity of a particle is absolute confers to the variable a specific status : it is a geometric vector. In the remarks following the Theorem 2 in the QM Chapter, we noticed that the status - vector or not - of a variable is not arbitrary : it is part of the assumptions of the theory. Velocity is an intrinsic property of material bodies and particles, the measure of this velocity depends on the observer : it is relative.

This remark extends to all measures. A physical measure in itself has no meaning if one does not know how it has been done. The label "done by the observer O" is necessary. So we cannot contend ourselves with maps $X : M \rightarrow E$. We need a way to attach a tag, identifying the way the measure has been done, to the value of the variable. The mathematical tool to achieve that is the fiber bundle formalism. This is more than a sophisticated mathematical theory, it embodies the relation between measure (the value) and conditions of the measure (the gauge).

3.2.1 Fiber bundles theory

(see Math.Part VI)

General fiber bundle

A fiber bundle, denoted $P(M, F, \pi_P)$, is a manifold P, which is locally the product of two manifolds, the base M and the standard fiber F, with a projection : $\pi_P : P \rightarrow M$ which is a surjective submersion. The subset of P : $\pi_P^{-1}(m)$ is the fiber over m. It is usually defined over a collection of open subsets of M, patched together, but we will assume that on the area Ω there is only one component (the fiber bundles are assumed to be trivial)[3]. A **trivialization** is a map :

$\varphi_P : M \times F \rightarrow P :: p = \varphi_P(m, v)$

and any element of P is projected on M : $\forall v \in F : \pi_P(\varphi_P(m, v)) = m$.So it is similar to a chart, but the arguments are points of the manifolds.

A **section p** on P is defined by a map : $v : M \rightarrow F$ and $\mathbf{p} = \varphi_P(m, v(m))$. The set of sections is denoted $\mathfrak{X}(P)$.

A fiber bundle can be defined by different trivializations. In a **change of trivialization** the *same* element p is defined by a different map φ_P : this is very similar to the charts for manifold.

$p = \varphi_P(m, v) = \widetilde{\varphi}_P(m, \widetilde{v})$

and there is a necessary relation between v and $\widetilde{v}$ (m stays always the same) depending on the kind of fiber bundle.

Principal bundle

If $F = G$ is a Lie group then P is a **principal bundle** : its elements are a value $g(m)$ of G localized at a point m.

p will usually define the basis used to measure vectors, so p is commonly called a gauge. There is a special gauge which can be defined at any point (it will usually be the gauge of the observer)

[3] If, in the mathematical definition of fiber bundles, the concept of collection of open subsets is essential, in all the practical consequences, notably with regard to the computation rules, the concept of change of trivialization is equivalent and has a clear physical meaning. So we can restrict ourselves to trivial bundles without loss of rigor.

: the **standard gauge,** the element of the fiber bundle such that : $\mathbf{p}(m) = \varphi_P(m, 1)$. The standard gauge is arbitrary : it reflects the free will of the observer, and as such is not submitted to any physical law. Its definition, with respect to measures, is done in protocols which document the experiments. There is no such thing as a given, natural, "field of gauges".

A principal bundle $P(M, G, \pi)$ is characterized by the existence of the right action of the group G on the fiber bundle P :

$\rho : P \times G \to P :: \rho(p, g') = \rho(\varphi_P(m, g), g') = \varphi_P(m, g \cdot g')$

which does not depend on the trivialization. So that any $p \in P$ can be written : $p = \varphi_U(m, g) = \rho(\mathbf{p}, g)$ with the standard gauge $\mathbf{p} = \varphi_P(m, 1)$.

A change of trivialization is induced by a map : $\chi : M \to G$ that is by a section $\chi \in \mathfrak{X}(P)$ and :

$p = \varphi_P(m, g) = \widetilde{\varphi}_P(m, \chi(m) \cdot g) = \widetilde{\varphi}_P(m, \widetilde{g}) \Leftrightarrow \widetilde{g} = \chi(m) \cdot g$ ($\chi(m)$ acts *on the left*)

$\chi(m)$ can be identical over M (the change is said to be global) or depends on m (the change is local).

The expression of the elements of a section change as :

$\sigma \in \mathfrak{X}(P) :: \sigma = \varphi_P(m, \sigma(m)) = \widetilde{\varphi}_P(m, \widetilde{\sigma}(m)) \Leftrightarrow \widetilde{\sigma}(m) = \chi(m) \cdot \sigma(m)$

$$\sigma(m) = \varphi_P(m, \sigma(m)) = \widetilde{\varphi}_P(m, \chi(m) \cdot \sigma(m)) \tag{3.16}$$

A change of trivialization induces a **change of standard gauge** :

$\mathbf{p}(m) = \varphi_P(m, 1) = \widetilde{\varphi}_P(m, \chi(m))$

$\to \widetilde{\mathbf{p}}(m) = \widetilde{\varphi}_P(m, 1) = \widetilde{\varphi}_P\left(m, \chi(m) \cdot \chi(m)^{-1}\right) = \mathbf{p}(m) \cdot \chi(m)^{-1}$

$$\left[\begin{array}{c} \mathbf{p}(m) = \varphi_P(m, 1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m) \cdot \chi(m)^{-1} : \\ \sigma(m) = \varphi_P(m, \sigma(m)) = \widetilde{\varphi}_P(m, \chi(m) \cdot \sigma(m)) \end{array}\right] \tag{3.17}$$

So changes of trivialization and change of gauge are the same operations, and we will consider usually a change of gauge.

Vector bundle

If $F = V$ is a vector space then P is a **vector bundle** and it has at each point the structure of a vector space :

$w_m = \varphi_P(m, w), w'_m = \varphi_P(m, w'), \alpha, \beta \in \mathbb{R}$:

$\alpha w_m + \beta w'_m = \varphi_P(m, \alpha w + \beta w')$

A holonomic basis is defined by a basis $(\varepsilon_i)_{i \in I}$ of V and : $\varepsilon_i(m) = \varphi_P(m, \varepsilon_i)$. It plays the same role as a standard gauge, and one needs a mean to go from one holonomic basis to another. It is provided usually by the action of a group, so usually vector bundles are defined as associated vector bundles. The principal bundle defines locally a standard with respect to which the measure is done.

Associated fiber bundle

Whenever there is a manifold F, a left action λ of G on F, one can built an **associated fiber bundle** denoted $P[F, \lambda]$ comprised of couples :

$(p, v) \in P \times F$ with the equivalence relation : $(p, v) \sim \left(p \cdot g, \lambda\left(g^{-1}, v\right)\right)$

The result belong to a fixed set, but its value is labeled by the standard which is used and related to a point of a manifold.

It is convenient to define these couples by using the standard gauge on P:

$$(\mathbf{p}(m), v) = (\varphi_P(m,1), v) \sim \left(\varphi_P(m,g), \lambda\left(g^{-1}, v\right)\right) \tag{3.18}$$

A standard gauge is nothing more than the use of an arbitrary standard, represented by 1, with respect to which the measure is done. *This is not a section* : the standard gauge is the embodiment of the free will of the observer, who can choose the way he proceeds to the measure, which is not fixed by any physical law. A change of standard gauge in the principal bundle impacts all associated fiber bundles (this is similar to the change of units) :

$$\left[\begin{array}{l} \mathbf{p}(m) = \varphi_P(m,1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m)\cdot\chi(m)^{-1} \\ v_p = (\mathbf{p}(m), v) = (\widetilde{\mathbf{p}}(m), \widetilde{v}) : \widetilde{v} = \lambda(\chi(m), v) \end{array}\right] \tag{3.19}$$

Similarly for the components of a section :
$\mathbf{v} \in \mathfrak{X}(P[V,\lambda]) :: \mathbf{v}(m) = (\mathbf{p}(m), v(m)) = \left(\mathbf{p}(m)\cdot\chi(m)^{-1}, \lambda(\chi(m), v)\right)$

If F is a vector space V and $[V,\rho]$ a representation of the group G then we have an **associated vector bundle** $P[V,\rho]$ which has locally the structure of a vector space. It is convenient to define a **holonomic basis** $(\varepsilon_i(m))_{i=1}^n$ from a basis $(\varepsilon_i)_{i=1}^n$ of V by : $\varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i)$ then any vector of $P[V,\rho]$ reads :

$$v_m = (\mathbf{p}(m), v) = \left(\mathbf{p}(m), \sum_{i=1}^n v^i\varepsilon_i\right) = \sum_{i=1}^n v^i\varepsilon_i(m) \tag{3.20}$$

A change of standard gauge $\mathbf{p}(m) = \varphi_P(m,1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m)\cdot\chi(m)^{-1}$ in the principal bundle impacts all associated vector bundles.

The holonomic basis of a vector bundle changes as :
$\varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i) \to$
$\widetilde{\varepsilon}_i(m) = (\widetilde{\mathbf{p}}(m), \varepsilon_i) = \left(\mathbf{p}(m)\cdot\chi(m)^{-1}, \varepsilon_i\right)$
$\sim \left(\left(\mathbf{p}(m)\cdot\chi(m)^{-1}\right)\cdot\chi(m), \rho\left(\chi(m)^{-1}\right)\varepsilon_i\right)$
$= \left(\mathbf{p}(m), \rho\left(\chi(m)^{-1}\right)(\varepsilon_i)\right) = \rho\left(\chi(m)^{-1}\right)\varepsilon_i(m)$
so that the components of a vector in the holonomic basis change as :
$v_m = \sum_{i=1}^n v^i\varepsilon_i(m) = \sum_{i=1}^n \widetilde{v}^i\widetilde{\varepsilon}_i(m) = \sum_{i=1}^n \widetilde{v}^i\rho(\chi(m))^{-1}\varepsilon_i(m)$
$\Rightarrow \widetilde{v}^i = \sum_j [\rho(\chi(m))]_j^i v^j$

$$\left[\begin{array}{l} \mathbf{p}(m) = \varphi_P(m,1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m)\cdot\chi(m)^{-1} : \\ \varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i) \to \widetilde{\varepsilon}_i(m) = \rho(\chi(m))^{-1}\varepsilon_i(m) \\ v^i \to \widetilde{v}^i = \sum_j [\rho(\chi(m))]_j^i v^j \end{array}\right] \tag{3.21}$$

The set of sections on $P[V,\rho]$, denoted $\mathfrak{X}(P[V,\rho])$, is an infinite dimensional vector space. In a change of standard gauge the components of a section change as :
$v \in \mathfrak{X}(P[V,\rho]) :: v(m) = \sum_{i=1}^n v^i(m)\varepsilon_i(m) = \sum_{i=1}^n \widetilde{v}^i(m)\widetilde{\varepsilon}_i(m)$
$\Leftrightarrow \widetilde{v}^i(m) = \sum_j \left[\rho\left(\chi(m)^{-1}\right)\right]_j^i v^j(m)$
so that $(\mathfrak{X}(P[V,\rho]), \rho)$ is an infinite dimensional representations of the group G.

I have given with great precision the rules in a change of gauge, as they will be used quite often (and are a source of constant mistakes ! For help see the Formulas in the Annex). They are necessary to ensure that a quantity is intrinsic : if it is geometric, its measure must change according to the rules. And conversely if it changes according to the rules, then one can say that it is intrinsic (this is similar to tensors). A quantity which is a vector of a fiber bundle is geometric with regard the conditions 1 of the 2nd chapter. Because this is a source of confusion, I will try to stick to these precise terms :

- a section = a point of a fiber bundle whose value is defined for each $m \in M$, this is a geometric object

- a gauge = a point of the principal bundle of P, this is a geometric object, which does not depend on a trivialization

- a standard gauge = a specific element of P, whose definition depends of the trivialization. This is not a section.

- a change of trivialization does not change the points of P, the gauge or the sections, but change the standard gauge and the way the points of P are defined with respect to the standard gauge

- it is equivalent to define a change of trivialization by the change of maps $\varphi_P \to \widetilde{\varphi}_P$ or by the change of standard gauge : $\mathbf{p}(m) = \varphi_P(m,1) \to \widetilde{\mathbf{p}}(m) = \widetilde{\varphi}_P(m,1)$

Notice that the elements of a section stay the same, but their definition changes, meanwhile the holonomic bases are defined by different elements. This is very similar to what we have in any vector space in a change of basis : the vectors of the basis change, the other vectors stay the same, but their components change.

Scalar product and norm

Whenever there is a scalar product (bilinear symmetric of Hermitian two form) $\langle\rangle$ on a vector space V, so that (V,ρ) is a *unitary* representation of the group G : $\langle \rho(g) v, \rho(g) v' \rangle = \langle v, v' \rangle$, the scalar product can be extended on the associated vector bundle $P[V,\rho]$:

$$\langle (\mathbf{p}(m), v), (\mathbf{p}(m), v') \rangle_{P[V,\rho]} = \langle v, v' \rangle_W \tag{3.22}$$

If this scalar product is definite positive, with any measure μ on the manifold M (usually the Lebesgue measure associated to a volume form as in the relativist context), one can define the spaces of integrable sections $L^q(M,\mu,P[V,\rho])$ of $P[V,\rho]$ (by taking the integral of the norm pointwise). For q = 2 they are Hilbert spaces, and unitary representation of the group G. Notice that the signature of the scalar product is that of the product defined on $P[V,\rho]$, *the metric on M is not involved.*

If there is a **norm** on V, that is a map :
$\|\| : V \to \mathbb{R}_+$
such that :
$\|X\| \geq 0$
$\|X\| = 0 \Leftrightarrow X = 0$
$\|kX\| = |k| \|X\|$
$\|X + X'\| \leq \|X\| + \|X'\|$
which does not depend on ρ :
$\forall g \in G : \|\rho(g) X\| = \|X\|$
then one can define a norm pointwise on $P[V,\rho]$:
$\|(p(m), v)\| = \|v\|$
$(p(m), v) \sim (p(m)\rho(g^{-1}), \rho(g) v)$

$\left\| \left(p(m) \rho\left(g^{-1}\right), \rho(g) v \right) \right\| = \|\rho(g) v\| = \|v\|$
and the space of integrable maps :
$L^1(\mathfrak{X}(P[V,\rho])) = \left\{ X \in \mathfrak{X}(P[V,\rho]), \int_\Omega \|X\| \mu < \infty \right\}$
is a separable Fréchet space if Ω is a compact subset.

There are several fiber bundles in the Geometry of the Universe. The simplest is the usual tangent bundle TM over M, which is a vector bundle associated to the choice of an invertible map at each point (the gauge group is $SL(\mathbb{R}, 4)$). Another one comes from the standard chart of an observer ;

Definition 50 *For any observer there is a* ***fiber bundle structure*** $\mathbf{M}_o(\mathbb{R}, \Omega(0), \pi_0)$ *on* M *with base* $\mathbb{R}$ *and :*
projection : $\pi_o(m) = f_0(m)$
trivialization : $\Phi_{\varepsilon_0} : \mathbb{R} \times \Omega(0) \to \Omega :: \Phi_{\varepsilon_0}(ct, x) = m$

3.2.2 Standard gauges associated to an observer

Frames and bases are used to measure components of vectorial quantities. Following the Principle of Locality any physical map, used to measure the *components of a vector* at a point m in M, must be done at m, that is in a local frame. Observers belong to $\Omega_3(t)$ and can do measures at any point of $\Omega_3(t)$.

They can measure components of vectors in the holonomic basis $(\partial\xi_\alpha)_{\alpha=0}^3$ given by a chart. This basis changes with the location but is fixed, in the meaning that it depends only on the choice of the spatial chart (to measure x on $\Omega_3(0)$) and the 4th vector $\partial\xi_0$ is defined, at any point of $\Omega_3(t)$, by the velocity of the observer (where ever he is located in $\Omega_3(t)$).

One property of the observers is that they have freedom of gauge : they can decide to measure the components of vectors in another basis than $(\partial\xi_\alpha)_{\alpha=0}^3$: usually, and this is what we will assume, they choose an orthonormal basis. This can be done by choosing 3 spatial vectors at a point, and we assume that they can extend the choice at any other point of $\Omega_3(t)$). However for the time vector the observer has actually no choice : this is necessarily the vector field $\mathbf{O}$ which is normal to $\Omega_3(t)$) and future oriented, and in the same direction as $\partial\xi_0$.

We will call such orthonormal bases a Standard gauge. They are abitrary, chosen by the observer, with the restriction about the choice of ε_0, and implemented all over $\Omega_3(t)$. They can be defined with respect to the holonomic basis of a chart.

This is equivalent to assume that, *for each observer*, there is a principal bundle $P_o(M, SO_0(3,1), \pi_p)$, a gauge $\mathbf{p}(m) = \varphi_P(m, 1)$ and an associated vector bundle $P_o\left[\mathbb{R}^4, \imath\right]$ where $\left(\mathbb{R}^4, \imath\right)$ is the standard representation of $SO_0(3,1)$. It defines at each point an holonomic orthonormal basis : $\varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i)$.To sum up :

Proposition 51 *For each observer there is :*
a principal ***fiber bundle structure*** $\mathbf{P}_o(M, SO_0(3,1), \pi_p)$ *on* M *with fiber the connected component of unity* $SO_0(3,1)$*, which defines at each point a* ***standard gauge*** *:* $\mathbf{p}(m) = \varphi_P(m, 1)$
an ***associated vector bundle*** *structure* $P_o\left[\mathbb{R}^4, \imath\right]$ *where* $\left(\mathbb{R}^4, \imath\right)$ *is the standard representation of* $SO_0(3,1)$*, which defines at any point* $m \in \Omega$ *the* ***standard basis*** $\varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i)$ *,i=0..3 where* $\varepsilon_0(m)$ *is orthogonal to the hypersurfaces* $\Omega_3(t)$ *to which* m *belongs.*

Notice that these structures depend on the observer. Starting with the principal bundle P_o, a change of gauge can be defined at any point by a section $\chi \in \mathfrak{X}(P_o)$ as seen above, with an impact on any associated bundle.

A standard basis is such that its time vector is $\mathbf{O}(m)$, so at the location of the observer it is in the direction of his velocity. Standard bases are not unique : their time vector is the same, but their space vectors can be rotated in $\Omega_3(t)$. Because they constitute an euclidean orthonormal basis the rotation is given by a matrix of $SO(3)$.

3.2.3 Formulas for a change of observer

Theorem 52 *For any two observers O,A the components of the vectors of the standard orthonormal basis of A, expressed in the standard basis of O are expressed by the matrix* $[\chi]$ *of* $SO_0(3,1)$, *where* $\overrightarrow{v}$ *is the instantaneous spatial speed of A with respect to O and R a matrix of* $SO(3)$ *:*

$$[\chi] = \begin{bmatrix} \frac{1}{\sqrt{1-\frac{\|v\|^2}{c^2}}} & \frac{\frac{v^t}{c}}{\sqrt{1-\frac{\|v\|^2}{c^2}}} \\ \frac{\frac{v}{c}}{\sqrt{1-\frac{\|v\|^2}{c^2}}} & I_3 + \left(\frac{1}{\sqrt{1-\frac{\|v\|^2}{c^2}}} - 1\right)\frac{vv^t}{\|v\|^2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix} \qquad (3.23)$$

Proof. Let be :

O be an observer (this will be main observer) with associated vector field $\mathbf{O}$, proper time t and world line $p_o(t)$

A be another observer with associated vector field $\mathbf{O'}$, proper time τ

Both observers use their standard chart φ_o, φ_A and their standard orthonormal basis, whose time vector is in the direction of their velocity. The location of A on his world line is the point m such that A belongs to the hypersurface $\Omega_3(t)$

The velocity of A at m :

$\frac{dp_A}{d\tau} = c\varepsilon'_0(m)$ by definition of the standard basis of A

$\frac{dp_A}{d\tau} = \frac{1}{\sqrt{1-\frac{\|\overrightarrow{v}\|^2}{c^2}}}(\overrightarrow{v} + c\varepsilon_0(m))$as measured in the standard basis of O

The matrix $[\chi]$ to go from the orthonormal basis $(\varepsilon_i(m))_{i=0}^3$ to $(\varepsilon'_i(m))_{i=0}^3$ belongs to $SO_0(3,1)$. It reads :

$[\chi(t)] = \begin{bmatrix} \cosh\sqrt{w^tw} & w^t\frac{\sinh\sqrt{w^tw}}{\sqrt{w^tw}} \\ w\frac{\sinh\sqrt{w^tw}}{\sqrt{w^tw}} & I_3 + \frac{\cosh\sqrt{w^tw}-1}{w^tw}ww^t \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix}$

for some $w \in \mathbb{R}^3, R \in SO(3)$

The elements of the first column of $[\chi(t)]$ are the components of $\varepsilon'_0(m)$, that is of $\frac{1}{c}\frac{dp_A}{d\tau}$ expressed in the basis of O :

$\cosh\sqrt{w^tw} = \frac{1}{\sqrt{1-\frac{\|v\|^2}{c^2}}}$

$w\frac{\sinh\sqrt{w^tw}}{\sqrt{w^tw}} = \frac{\overrightarrow{v}}{c}\frac{1}{\sqrt{1-\frac{\|v\|^2}{c^2}}}$

$w = k\overrightarrow{v} \Rightarrow w^tw = k^2\|\overrightarrow{v}\|^2$

which leads to the classic formula with

$w = \frac{v}{\|v\|}\arg\tanh\left\|\frac{v}{c}\right\| = \frac{1}{2}\frac{v}{\|v\|}\ln\left(\frac{c+\|\overrightarrow{v}\|}{c-\|\overrightarrow{v}\|}\right) \sim \frac{1}{2}\frac{v}{\|v\|}\ln\left(1+2\frac{\|\overrightarrow{v}\|}{c}\right) \sim \frac{v}{c}$ ■

Some key points to understand these formulas :

- They hold *for any observers* O, A, who use their standard orthonormal basis (the time vector is oriented in the direction of their velocity). There is no condition such as inertial frames.

- The points of M where O and A are located can be different, $O \in \Omega_3(\tau), A \in \Omega_3(\tau) \cap \Omega_3(t)$. The spatial speed $\overrightarrow{v}$ is a vector belonging to the space tangent at $\Omega_3(\tau)$ at the location m of A (and not at the location of O at t) and so is the relative speed of A with respect to the point m of M, which is fixed for O.

- The formulas are related to the standard orthonormal bases $(\varepsilon_i(m))_{i=0}^3$ of O and $(\varepsilon_i'(m))_{i=0}^3$ of A located at the point m of $\Omega_3(t)$ where A is located.

- These formulas apply to the components of vectors in the standard orthonormal bases. Except in SR, there is no simple way to deduce from them a relation between the coordinates in the charts of the two observers.

- The formula involves a matrix $R \in SO(3)$ which represents the possible rotation of the spatial frames of O and A, as it would be in Galilean Geometry.

These formulas have been verified with a great accuracy, and *the experiments show that c is the speed of light.* This is an example of a theory which is checked by the consequences that can be drawn from its basic assumptions.

We will see below how these formulas apply in Special Relativity.

If we take $\frac{v}{c} \to 0$ we get $[\chi] = \begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix}$, that is a rotation of the usual space. The Galilean Geometry is an approximation of SR when the speeds are small with respect to c. Then the velocities are $\frac{d\mu_A}{d\tau} = (\overrightarrow{v} + c\varepsilon_0)$ with a common vector ε_0.

3.2.4 The Tetrad

The principal fiber bundle P_G

So far we have defined a chart φ_o and a fiber bundle structure P_o for an observer : the construct is based on the trajectory of the observer, and his capability to extend his frame over the hypersurfaces $\Omega_3(t)$. With the formulas above we see how one can go from one observer to another, and thus relate the different fiber bundles P_o. The computations in a change of frame can be done with measures done by the observers, and have been checked experimentally. So it is legitimate to assume that there is a more general structure of principal bundle, denoted $\mathbf{P}_G(M, SO_0(3,1), \pi_G)$,over M. In this representation the bases used by any observer is just a choice of specific trivialization, or equivalently of standard gauge, and one goes from one trivialization to another with the matrix $[\chi]$.

Proposition 53 *There is a unique structure of principal bundle* $\mathbf{P}_G(M, SO_0(3,1), \pi_G)$ *with base* M*, standard fiber* $SO_0(3,1)$ *. A change of observer is given by a change of trivialization on* P_G*.*

The standard gauge $\mathbf{p}(m) = \varphi_G(m,1)$ *is, for any observer, associated to his standard basis* $\varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i)$*..*

If it is easy to define such a mathematical structure, it is necessary to understand its physical meaning.

Charts on a manifold are a way to locate a point. As such they are arbitrary and fixed. They are only related to the manifold structure. From a chart one can define a vector space at any point, through small translations $(\partial\xi_\alpha)_{\alpha=0}^3$. along the coordinates : they define a holonomic basis, only linked to the chart. Physically, for the observer, there is no way to measure a change in the spatial vectors $(\partial\xi_\alpha)_{\alpha=1}^3$ of a holonomic basis : for him they represent fixed directions (given for instance by far away stars). Because the translation in time is necessarily done along the time coordinate, that is along the world line of the observer, the vector $\partial\xi_0$ is a time vector, proportional to the 4 dimensional velocity of the observer. This is the standard chart of the observer, and physically one cannot conceive other charts : a physical chart is always the standard chart of some observer. If, in a theoretical model, one sets some chart with respect to which the quantities are to be measured, this is actually the choice of a specific observer which is done, and this is equivalent to the choice of a vector field $\mathbf{O}$. The vector field $\mathbf{O}$ and the time of an observer define the system itself.

With the Fiber Bundle structure we add several components.

i) A specific basis $(\varepsilon_i)_{i=0}^3$ defined at any point. As for any vectors their components can be measured in the holonomic basis, but the two bases are independent. The vectors $(\varepsilon_i)_{i=0}^3$ are arbitrarily fixed by the observer, who has freedom of gauge. If the observer chooses the 3 vectors $(\varepsilon_i)_{i=1}^3$ in his space, then they belong to the space $T_m\Omega_3(t)$ tangent at $\Omega_3(t)$, and they constitute an euclidean basis. In any orthonormal basis there is a time like vector ε_0, orthogonal to the 3 vectors $(\varepsilon_i)_{i=1}^3$. Thus if $(\varepsilon_i)_{i=1}^3$ belong to $T_m\Omega_3(t)$, because there is a unique unitary future oriented normal to $T_m\Omega_3(t)$, we have necessarily $\varepsilon_0 = \mathbf{O}$. This is the standard basis, or the standard gauge, which is defined up to the choice of $(\varepsilon_i)_{i=1}^3$ and two standard basis at the same point are related by a matrix of $SO(3)$.

ii) Even if their components are known, the comparison between two orthonormal bases (standard or not) at different location (notably at different time) requires specific tools (connections) because they belong to different vector spaces. So there is no simple definition of a "constant" orthonormal basis and this is the topic of inertial observers seen in the Chapter 5.

iii) The condition for 4 vectors to be orthogonal depends on the metric, which changes with the location. It is proven in Differential Geometry (Maths.1609) that there is no chart such that its holonomic basis is orthogonal at each point (the manifolds with this property, which is not assumed for M, are special and said to be parallelizable). This is due to the fact that a metric is an object which is added to the structure of manifold, it does not come with it. And there is no reason why it would be constant [4](for the time being we do not make any assumption about the factors which can explain this varying metric). So an orthonormal basis cannot have fixed components in any chart, even if the observer strives to keep them as still as possible.

iv) With the structure of fiber bundle independent of an observer, it is possible to compute the impact of a change of gauge with a change of ε_0. It implies a change of observer, moreover the formulas hold for observers located at the same point, and using their standard bases.

v) A change of gauge is given by a section χ (global or not) of $\mathbf{P}_G$, the vectors of the standard basis transform according to the matrix $[\chi]$. The *operation is associative : the combination of relative motions is represented by the product of the matrices,* which is convenient.

Tetrad

The vectors of a standard basis (the tetrad) can be expressed in the holonomic basis of any chart (of an observer or not).

$$\varepsilon_i(m) = \sum_{\alpha=0}^{3} P_i^\alpha(m)\,\partial\xi_\alpha \Leftrightarrow \partial\xi_\alpha = \sum_{i=0}^{3} P_\alpha^{\prime i}(m)\,\varepsilon_i(m) \tag{3.24}$$

where $[P]$ is a real invertible matrix (which has no other specific property, it does not belong to $SO(3,1)$) and we denote

[4]Even in an affine space, such as in SR, there is no reason why the metric should be constant. This is an additional assumption in SR.

Notation 54 $[P'] = [P]^{-1} = \left[P'^{i}_{\alpha}\right]$.

The components of a vector change as :
$u = \sum_{\alpha=0}^{3} u^{\alpha}\partial\xi_{\alpha} = \sum_{i=0}^{3} U^{i}\varepsilon_i(m) \Leftrightarrow U^i = \sum_{\alpha=0}^{3} P'^{i}_{\alpha}u^{\alpha} \Leftrightarrow u^{\alpha} = \sum_{i=0}^{3} P^{\alpha}_{i}U^{i}$
The dual of $(\partial\xi_{\alpha})_{\alpha=0}^{3}$ is $(d\xi^{\alpha})_{\alpha=0}^{3}$ with the defining relation :
$d\xi^{\alpha}(\partial\xi_{\beta}) = \delta^{\alpha}_{\beta}$.
The dual $\left(\varepsilon^i(m)\right)_{i=0}^{3}$ is :

$$\varepsilon^i(m) = \sum_{i=0}^{3} P'^{i}_{\alpha}(m)\, d\xi^{\alpha} \Leftrightarrow d\xi^{\alpha} = \sum_{i=0}^{3} P^{\alpha}_{i}(m)\,\varepsilon^i(m) \tag{3.25}$$

Notice that, if it is usual to associate to a vector $u \in T_mM$ a covector : $u^* = \sum_{\alpha\beta} g_{\alpha\beta}u^{\beta}d\xi^{\alpha}$, then $u^*(v) = \langle u, v\rangle$ so that $(\varepsilon_i)^* \neq \varepsilon^i : \varepsilon_i^*(v) = \eta_{ii}v^i \neq \varepsilon^i(v) = v^i$.

In the fiber bundle representation the vectors of the tetrad are variables which are vectors $\varepsilon_i \in \mathfrak{X}(TM)$ or covectors $\varepsilon^i \in \mathfrak{X}(TM^*)$. The quantities $(P^{\alpha}_{i}(m))_{i=1}^{3}$ (called vierbein) and $\left(P'^{i}_{\alpha}(m)\right)_{i=1}^{3}$ are the components of the vectors $\varepsilon_i(m)$ or the covectors $\varepsilon^i(m)$ in any chart. They can be measured, if one has a chart. They depend on the observer, change with the location m and in a change of chart as the components of a vector or a covector. The quantities $\varepsilon_i(m)$ are geometric and physical quantities. So they are one of the variables in any model in GR : as such they replace the metric g. However it is obvious that $[P]$ is defined, in any chart, up to a matrix of $SO(3,1)$, so there is some freedom in the choice of the gauge, and we will see the consequences in the specification of a lagrangian.

In a change of gauge on the principal bundle $P_G : \mathbf{p}(m) = \varphi_P(m,1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m)\cdot\chi(m)^{-1}$ the holonomic basis becomes with $[\chi(m)] \in SO_0(3,1)$
$\varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i) \to \widetilde{\varepsilon}_i(m) = [\chi(m)]^{-1}\varepsilon_i(m)$
$\sum_{\alpha=0}^{3} \widetilde{P}^{\alpha}_{i}(m)\,\partial\xi_{\alpha} = \left[\chi(m)^{-1}\right]^{j}_{i} \sum_{\alpha=0}^{3} P^{\alpha}_{j}(m)\,\partial\xi_{\alpha}$

$$\mathbf{p}(m) = \varphi_P(m,1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m)\cdot\chi(m)^{-1} : [P] \to \left[\widetilde{P}\right] = [\chi(m)]^{-1}[P] \tag{3.26}$$

With respect to the standard chart of the observer :
$\varepsilon_0(p_o(t)) = \partial\xi_0 \Rightarrow P'^{i}_{0} = \delta^{i}_{0}$
$\alpha = 1,2,3 : \frac{\partial}{\partial\xi^{\alpha}}\varphi_o\left(\xi^0,\xi^1,\xi^2,\xi^3\right) = \partial\xi_{\alpha} = \frac{\partial}{\partial x}\Phi_{\varepsilon_0}(ct,x)\frac{\partial x}{\partial\xi^{\alpha}} \in T_m\Omega_3(t) \Rightarrow P'^{0}_{\alpha} = 0$
so
$[P] = \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix}; [Q] = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix}$
$[P'] = \begin{bmatrix} 1 & 0 \\ 0 & Q' \end{bmatrix}; [Q'] = \begin{bmatrix} P'_{11} & P'_{12} & P'_{13} \\ P'_{21} & P'_{22} & P'_{23} \\ P'_{31} & P'_{32} & P'_{33} \end{bmatrix}$
$[Q][Q'] = I_3$
$[Q]$ is fixed by the observer, it expresses his "freedom of gauge". However, as it has been said before, he cannot hope to keep $[Q]$ fixed : the change of the metric with the location implies that, with respect to any chart, $[Q]$ must change. The change in the metric is measured by the deformation of the matrix $[Q]$ and, as for deformable solid, represents a "deformation tensor". Indeed in the SR context, where the metric is constant, one can assume that it is possible to keep : $[P] = I_4, [Q] = I_3$ (they are the "inertial observers").

Metric

The scalar product can be computed from the components of the tetrad. By definition :

$g_{\alpha\beta}(m) = \langle \partial\xi_\alpha, \partial\xi_\beta \rangle = \sum_{ij=0}^3 \eta_{ij} [P']^i_\alpha [P']^j_\beta$

The induced metric on the cotangent bundle (Maths.1608) is denoted with upper indexes :

$g^* = \sum_{\alpha\beta} g^{\alpha\beta} \partial\xi_\alpha \otimes \partial\xi_\beta$

and its matrix is $[g]^{-1}$:

$g^{\alpha\beta}(m) = \sum_{ij=0}^3 \eta^{ij} [P]^\alpha_i [P]^\beta_j$

$$[g]^{-1} = [P][\eta][P]^t \Leftrightarrow [g] = [P']^t [\eta][P'] \tag{3.27}$$

It does not depend on the gauge on P_G :

$[\widetilde{g}] = \left[\widetilde{P'}\right]^t [\eta] \left[\widetilde{P'}\right] = [P']^t \left[\chi(m)^{-1}\right]^t [\eta] \left[\chi(m)^{-1}\right] [P'] = [P']^t [\eta][P']$

So in the standard chart of the observer : $g^{00} = -1$.

$[g] = [P']^t [\eta][P'] = \begin{bmatrix} -1 & 0 \\ 0 & [g]_3 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & [Q']^t [Q'] \end{bmatrix}$

$[g]^{-1} = [P][\eta][P]^t = \begin{bmatrix} -1 & 0 \\ 0 & [g]_3^{-1} \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & [Q][Q]^t \end{bmatrix}$

and $[g]_3$ is definite positive.

The metric defines a **volume form** on M (Maths.1609). Its expression in any chart is, by definition :

$\varpi_4(m) = \varepsilon_0 \wedge \varepsilon_1 \wedge \varepsilon_2 \wedge \varepsilon_3 = \sqrt{|\det[g]|} d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$[g] = [P']^t [\eta][P'] \Rightarrow \det[g] = -(\det[P'])^2 \Rightarrow \sqrt{|\det[g]|} = \det[P']$

assuming that the standard basis of $P_G\left[\mathbb{R}^4, \imath\right]$ is direct.

$$\varpi_4 = \det[P'] \, d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3 \tag{3.28}$$

Induced metric

The metric on M induces a metric on any submanifold but it is not necessarily non degenerated (Maths.19.3.1).

On hypersurfaces the metric is non degenerated if the unitary normal n is such that $\langle n, n \rangle \neq 0$ (Maths.1642). The induced volume form is (Maths.1644) :

$\mu_3 = i_n \varpi_4 = \det[P'] \, d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3 (n)$

For $\Omega_3(t)$ the unitary normal n is ε_0 , the induced metric is Riemannian and the volume form ϖ_3 is :

$\varpi_3 = i_{\varepsilon_0} \varpi_4 = \det[P'] \, d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3 (\varepsilon_0)$

$= \det[P'] \, d\xi^0 (\varepsilon_0) \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$= \det[P'] \, d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$$\varpi_3 = \det[P'] \, d\xi^1 \wedge d\xi^2 \wedge d\xi^3 \tag{3.29}$$

and conversely :

$\varpi_4 = \varepsilon_0 \wedge \varpi_3 = \det[P'] \, d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

ϖ_3 is defined with respect to the coordinates ξ^1, ξ^2, ξ^3 but the measure depends on $\xi^0 = t$.

For a curve C, represented by any path : $p:\mathbb{R}\rightarrow C::m=p(\theta)$ the condition is $\left\langle\frac{dp}{d\theta},\frac{dp}{d\theta}\right\rangle\neq 0$. The volume form on any curve defined by a path : $q:\mathbb{R}\rightarrow M$ with tangent $V=\frac{dq}{d\theta}$ is$\sqrt{|\langle V,V\rangle|}d\theta$. So on the trajectory $q(t)$ of a particle it is

$$\varpi_1(t)=\sqrt{-\langle V,V\rangle}dt \tag{3.30}$$

For a particle there is the privileged parametrization by the proper time, and as $\left\langle\frac{dp}{d\tau},\frac{dp}{d\tau}\right\rangle=-c^2$ the length between two points A,B is :

$\ell_p=\int_{\tau_A}^{\tau_B}\sqrt{-\left\langle\frac{dp}{d\tau},\frac{dp}{d\tau}\right\rangle}d\tau=\int_{\tau_A}^{\tau_B}cd\tau=c(\tau_B-\tau_A)$

This is an illustration of the idea that all world lines correspond to a travel at the same speed.

3.2.5 From particles to material bodies

In Mechanics a material body is comprised of "material points" that is elements of matter whose location is a single geometric point, and change with time in a consistent way : their trajectories do not cross, and the body keeps some cohesion, which is represented by a deformation tensor for deformable solid bodies. In Relativity the material points, particles in our terminology, follow independent world lines, which do not cross, and thus can be represented by a field of vectors u, future oriented with length $\langle u,u\rangle=-c^2$, such that, at some time 0, the particles are all together in a compact subset $\varpi(0)$ of a 3 dimensional space like submanifold. Then the location of any particle of the material body is given by $\Phi_u(\tau,a)$ where τ is its proper time and $a\in\varpi(0)$ its location at $\tau=0$. The area swept by the body is $\widehat{\omega}=\{\Phi_u(\tau,a):\tau\in\mathbb{R},a\in\omega(0)\}$ and we have the function : $f:\widehat{\omega}\rightarrow\mathbb{R}:f(m)=\tau:\exists a\in\omega(0):\Phi_u(\tau,a)=m$. The function f defines a folliation in diffeomorphic compact 3 dimensional hypersurfaces $\omega(\tau)$ which can be seen as the state of the material body at τ (Maths.1503). So $\Phi_u(\tau,a)$ can be seen as a chart of $\widehat{\omega}$, and the material body has a unique proper time τ. We can then give a definition of a material body which is independent of any observer.

Definition 55 *A **material body** is defined by a field of vectors u, future oriented with length $\langle u,u\rangle=-c^2$, and a compact subset $\varpi(0)$ of a 3 dimensional space like submanifold. The body is located at its proper time τ on the set $\omega(\tau)$ diffeomorphic to $\varpi(0)$.*

The vector field $u\in\mathfrak{X}(TM)$ does not depend on a chart, but for any observer O the trajectories of the material points of the body will follow a vector field $V\in\mathfrak{X}(TM)$: the curves do not change, but they are traveled with the parameter t and not τ.

A material point can be labeled by its position $y\in\omega(0)$ then its location p is

$\Phi_u(\tau,a)$ at the proper time τ along the world line going through y

$\Phi_{\varepsilon_0}(t,x(t))=\varphi_O(t,x(t))$ at the time t of the observer in the chart of the observer given by his vector field ε_0

and $\tau,t,y,x(t)$ are related by :

$u=\frac{dp}{d\tau}$

$V(t)=\frac{dp}{dt}=u\frac{\sqrt{-\langle V,V\rangle}}{c}=c\varepsilon_0+\overrightarrow{v}$

$\overrightarrow{v}=\frac{dx}{dt}$

$x(0)=a$

At $t=0$ for the observer the set of points of the material body is $\widehat{\omega}(0)=\widehat{\omega}\cap\Omega(0)$

At $t>0$ for the observer the set of points of the material body is $\widehat{\omega}(t)=\widehat{\omega}\cap\Omega(t)$

$\widehat{\omega}(t)=\{\varphi_O(t,x(t)),x(0)\in\widehat{\omega}(0)\}$

$=\{\Phi_V(t,x(0)),x(0)\in\widehat{\omega}(0)\}$

$= \{\Phi_u(\tau(t), a), a \in \widehat{\omega}(0)\}$
and $\omega(\tau(t)) = \{\Phi_u(\tau(t), y), a \in \omega(0)\}$

So *the characterization of a material body is observer dependant* : they do not see the same body.

However we will assume :

Proposition 56 *For any material body there are observers with proper time t such that, at $t = 0$ they observe the entire material body : $\omega(0) \subset \Omega(0)$*

Then at any given time t : $\widehat{\omega}(t) = \omega(\tau(t))$. This is a legitimate assumption, which will be mainly used to compute the characteristics of material bodies.

In Mechanics a solid is a material body such that the distance between any two of its points is constant. $\omega(\tau)$ is a Riemannian manifold, with the metric g_τ induced by the metric g on M. g_τ defines the length of a curve on $\omega(\tau)$ and the distance between two of its points is then the minimum of the length of all the lines which join the points. Because $\omega(\tau)$ is compact such a minimum exists, however the metric g_τ itself depends on the point and τ, so this concept of solid cannot be extended in Relativity.

3.2.6 Special Relativity

All the results of this chapter hold in Special Relativity. This theory, which is still the geometric framework of QTF and Quantum Physics, adds two assumptions : the Universe M can be represented as an affine space, and the metric does not depend on the location (these assumptions are independent). As consequences :

- the underlying vector space $\overrightarrow{M}$ (the Minkovski space) is common to all observers : the vectors of all tangent spaces to M belong to $\overrightarrow{M}$

- one can define orthonormal bases which can be freely transported and compared from a location to another

- because the scalar product of vectors does not depend on the location, at each point one can define time-like and space-like vectors, and a future orientation (this condition relates the mathematical and the physical representations, and $\overrightarrow{M}$ is not simply $\mathbb{R}^4$)

- there are fixed charts $\left(O, (\varepsilon_i)_{i=0}^3\right)$, called frames, comprised of an origin (a location O in M : a point) and an orthonormal basis $(\varepsilon_i)_{i=0}^3$. There is necessarily one vector such that $\langle \varepsilon_i, \varepsilon_i \rangle = -1$. It is possible to define, non unique, orthonormal bases such that ε_0 is timelike and future oriented.

- the coordinates of a point m, in any frame $\left(O, (\varepsilon_i)_{i=0}^3\right)$, are the components of the vector OM. The transition maps which give the coordinates of m in another frame $\left(A, (\widetilde{\varepsilon}_i)_{i=0}^3\right)$ are then given by the formulas :

$OM = \sum_{i=0}^3 x_i \varepsilon_i$
$AM = \sum_{i=0}^3 \widetilde{x}_i \widetilde{\varepsilon}_i$
$OM = OA + AM = \sum_{i=0}^3 L_i \varepsilon_i + \sum_{i=0}^3 \widetilde{x}_i \widetilde{\varepsilon}_i$
$\widetilde{\varepsilon}_i = \sum_{j=0}^3 [\chi]_i^j \varepsilon_i, [\chi] \in SO(3,1)$

However one needs to go from this abstract representation to a physical definition of frames.

Observers can label points which are in their present with their proper time. The role of the function $f(m) = t$ is crucial, because it defines the 3 dimensional hypersurfaces $\Omega(t)$. They are not necessarily hyperplanes, but they must be space like and do not cross each other : a point m cannot belong to 2 different hypersurfaces. These hypersurfaces define the vector field $\varepsilon_0 (m)$ to which belongs the velocity of the observer (up to c). In SR one can compare vectors at different points, and usually the vectors $\varepsilon_0 (m)$ are different from one location to another. They are identical only if $\Omega (t)$ are hyperplanes normal to a vector ε_0, which implies that the world line of the observer is a straight line, and because the proper time is the parameter of the flow, if the motion of the observer is a translation at a constant spatial speed. These observers are called **inertial**. Notice that this definition is purely geometric and does not involve gravitation or inertia : the motion of an observer is absolute, and inertial observers are such that their velocity is a constant vector.

Observers can define a standard chart $\varphi_o \left(\xi^0, \xi^1, \xi^2, \xi^3\right)$ with $\xi^0 = ct, \partial\xi_0 = \varepsilon_0 (m)$ with the flow of $\varepsilon_0 (m)$ and a chart φ_Ω of $\Omega (0)$ which provides the coordinates ξ^1, ξ^2, ξ^3 of the point x where the integral curves of ε_0 passing through m crosses $\Omega (0)$. The general results hold and such a chart can always be defined. However this chart is usually not defined by a frame $\left(O, (\varepsilon_i)_{i=0}^3\right)$: the vectors of the basis must be constant, and notably ε_0 so this is possible only if the observer is inertial : *a frame can be associated to an observer only if this is an inertial observer.*

For inertial observers the integral curves are straight lines parallel to ε_0.Any spatial basis $(\varepsilon_i)_{i=1}^3$ of $\Omega (0)$ can be transported on $\Omega (t)$. The standard chart is then similar to a frame in the 4 dimensional affine space $\left(O (0), (\varepsilon_i)_{i=0}^3\right)$ with origin $O (0)$, the 3 spatial vectors $(\varepsilon_i)_{i=1}^3$ and the time vector ε_0. The coordinates of a point $m \in \Omega_3 (t)$ are :

$\overrightarrow{O (0) m} = ct\varepsilon_0 + \sum_{i=1}^3 \xi^i \varepsilon_i$ where $\overrightarrow{O (t) m} = \sum_{i=1}^3 \xi^i \varepsilon_i$

and the velocity of a particle with trajectory $q (t)$, as measured by O, is :

$V = c\varepsilon_0 + \overrightarrow{v}$ with $\overrightarrow{v} = \frac{d}{dt}\overrightarrow{O (t) q (t)} = \sum_{i=1}^3 \frac{d\xi^i}{dt}\varepsilon_i$ because $\frac{d\varepsilon_i}{dt} = 0$

If there is another inertial observer with standard chart defined by a frame $\left(A (0), (\widetilde{\varepsilon}_i)_{i=0}^3\right)$ the coordinates of $A (t)$, as seen by $O (t)$, are :

$\overrightarrow{O (0) A (t)} = ct\varepsilon_0 + \sum_{i=1}^3 \xi^i (t) \varepsilon_i$ where $\overrightarrow{O (t) A (t)} = \sum_{i=1}^3 \xi^i (t) \varepsilon_i$

The spatial speed of A is : $\overrightarrow{v} = \frac{d}{dt}\overrightarrow{O (t) A (t)} = \sum_{i=1}^3 \frac{d\xi^i}{dt}\varepsilon_i$

We can then implement the general results for the change of basis : $\varepsilon_i \to \widetilde{\varepsilon}_i$.

As for the change of coordinates we have :

$\overrightarrow{O (0) m} = ct\varepsilon_0 + \sum_{1=1}^3 \xi^i \varepsilon_i$

$\overrightarrow{A (0) m} = c\tau\widetilde{\varepsilon}_0 + \sum_{1=1}^3 \widetilde{\xi}^i \widetilde{\varepsilon}_i$

$\left[\widetilde{\xi}^i\right] = [\chi]^{-1} \left[\xi^i\right]$

$\overrightarrow{O (0) m} = \sum_{i=0}^3 \xi^i \varepsilon_i = L + \sum_{i=0}^3 \widetilde{\xi}^i \widetilde{\varepsilon}_i$

with a constant vector $\overrightarrow{O (0) A (0)} = L$.

So the transformation of the coordinates is given by the product of a fixed translation and a fixed rotation in the Minkovski space. The set of such transformations is a group, called the Poincaré's group.

This result holds only for two inertial observers, and we need a physical mean to tell what are these observers. The usual answer is that they do not feel a change in the inertial forces to which they are submitted. This is similar to the Galilean observers of Classic Mechanics. For non inertial observers the general formulas hold, but the charts cannot be defined through frames as in an affine space.

The concept of material body presented above holds. But if $\varpi (0)$ belongs to a hyperplane then the $\varpi (\tau)$ will be hyperplanes only if they are all orthonormal to a common vector, that is if the vector field which defines the material body is a constant vector : the body must be in a uniform translation (and not rotate on itself).

The formulas of the Lorentz transformations have a tremendous importance in all aspects of Relativist Physics, they are of a constant use, as well as the Poincaré's group which is the starting point in the identification of particles. However any demonstration based on frames, as it is usually done, holds only for inertial observers. A physical theory which is valid only for the study of bodies in uniform translation would be of little interest. As we have proven in this chapter, Relativist Geometry can be explained, in a rigorous and quite simple way, without the need of inertial observers. And these are required only for the use of frames. It would be a pity to loose the deep import of Relativity in order to keep a familiar, but not essential, mathematical tool. As a consequence the role assigned to the Poincaré's group must be revisited.

3.3 SOME ISSUES ABOUT RELATIVITY

It is useful to review here some issues which arise frequently about Relativity.

3.3.1 Preferred frames

Relativity is often expressed as "all inertial frames are equivalent for the Physical Laws". We have seen above that actually inertial frames are required only to define coordinates in affine space : this is a non issue in GR, and in SR it is possible to achieve the usual results with the use of standard charts which are not given by orthogonal frames. But, beyond this point, this statement is misleading.

The Theory of Relativity is more specific than the Principle of Relativity, it involves inertia and gravitation (that we will see in the next chapters), but this is at first a Theory about the Geometry of the Universe, and it shows that the geometric measures (of lengths and time) are specific to each observer. The Universe which is Scientifically accessible - meaning by the way of measures, data and figures - depends on the observer. We can represent the Universe with 4 dimensions, conceive of a 4 dimensional manifold which extends over the past and the future, but we must cope with the fact that we are stuck into our present, and it is different for each of us. The reintegration of the observer in Physics is one of the most important feature of Relativity, and the true meaning of the celebrated formulas for a change of frames. An observer is an object in Physics, and as such some properties are attached to it, among them the free will : the possibility to choose the way he proceeds to an experiment, without being himself included in the experiment. But as a consequence the measures are related to his choice.

Mathematics give powerful tools to represent manifolds, in any dimensions. And it seems easy to formulate any model using any chart as it is commonly done. This is wrong from a physical point of view. There is no banal chart or frame : it is always linked to an observer, there is a preferred chart, and so a preferred frame for an observer. It is not related to inertia : it is a matter of geometry, and a consequence of the fundamental symmetry breakdown. The observer has no choice in the selection of the time vector of his orthonormal basis, if he wants to change the vector, he has to change his velocity, and this is why the formulas in a change of frames are between two different observers moving with respect to each other. And not any change is possible : an observer cannot travel in the past, or faster than light. These features are clear when one sticks to a chart of an observer, as we will do in this book. Not only they facilitate the computations, they are a reminder of the physical meaning of the chart. This precision is specially important in the fiber bundle formalism, which is, from this point of view, a wise precaution as compared to the usual formalism using undifferentiated charts.

3.3.2 Time travel

The distinction between future and past oriented vectors come from the existence of the Lorentz metric. As it is defined everywhere, it exists everywhere, and along any path. It is not difficult to see that the border between the two kinds of vectors is for null vectors $\langle u, u \rangle = 0$. So a particle which would have a path such that its velocity is past oriented should, at some point, have a null velocity, and, with regard to another observer located at the same point, travel at the speed of light. Afterwards his velocity would be space like ($\langle u, u \rangle > 0$) before being back time like but past oriented. Clearly this would be a discontinuity on the path and "Scotty engages the drive" from Star Treck has some truth.

But the main issue with time travel lies in the fact that, if ever we would be able to come back to the location where we have been in the past (meaning a point of the universe located in our past), we would not find our old self. The idea that we exist in the past assumes that we exist at any time along our world line, as a frozen copy of ourselves. This possibility is sometimes invoked, but it raises another one : what makes us feel that each instant of time is different ? If we do not travel physically along our world line, what does move ? And of course this assumption raises many other issues in Physics, among them the potential violations of the Principle of Causality which are the bread and butter of science fiction books on time travel.

3.3.3 Twins paradox

The paradox is well known : one of the twins embarks in a rocket and travels for some time, then comes back and finds that he is younger than his twin who has stayed on Earth. This paradox is true (and has been checked with particles) and comes from two relativist features : the Universe is 4 dimensional, and the definition of the proper time of an observer.

Because the Universe is 4 dimensional, to go from a point A to a point B there are several curves. Each curve can be travelled according to different paths. We have assumed that observers move along a curve according to a specific path, their world line, and then :

$\ell_{AB} = c\left(\tau_B - \tau_A\right)$

Because the curves are different, the elapsed proper time is usually different.

The proper time is the time measured by a clock attached to the observer, it is his biological time. Assuming that all observers travel along their world lines with a velocity such that at $\left\langle \frac{dp_o}{d\tau}, \frac{dp_o}{d\tau} \right\rangle = -c^2$ is equivalent to say that, with respect to their clock, they age at the same rate. So if they travel along different curves there is no reason that the total duration of their travel would be the same.

Whom of the two twins would have aged the most ? It is not easy to do the computation in GR, but simpler in the SR context.

We can define a fixed frame $\left(O, (\varepsilon_i)_{i=0}^3\right)$ with origin O at the time $t = 0$, A is spatially immobile with respect to this frame, moves along the time axis and his coordinates are then : $OA : p_A\left(\tau_A\right) = c\tau_A\varepsilon_0$

The twin B moves in the direction of the first axis. His coordinates are then : $OB : p_B\left(\tau_B\right) = c\tau_B\varepsilon_0 + x_B\left(\tau_B\right)\varepsilon_1$

The spatial speed of B with respect to A is : $\frac{dOB}{d\tau_A} = V\left(\tau_B\right)\varepsilon_1$

The velocity of B is : $u_B = \frac{dOB}{d\tau_B} = \frac{1}{\sqrt{1-\frac{V^2}{c^2}}}\left(V\varepsilon_1 + c\varepsilon_0\right)$

To be realistic we must assume that B travels at a constant acceleration, but needs to brake before reaching first his turning point, then A. In the first phase we have for instance :

$V = \gamma c\tau_B$ with $\gamma = \frac{1}{\sqrt{1-\frac{V^2}{c^2}}}$

$p_B\left(\tau_B\right) = \int_0^{\tau_B} \frac{c}{\sqrt{1-(\gamma t)^2}}\left(\gamma t\varepsilon_1 + \varepsilon_0\right) dt = \frac{c}{\gamma}\left[\sqrt{1-y^2}\varepsilon_1 + \varepsilon_0 \arcsin y\right]_0^{\gamma\tau_B}$

A full computation gives : $\frac{\tau_A}{\tau_B} = \frac{\arcsin v_M}{v_M}$ where v_M is the maximum speed in the travel, which gives for $v_M = c : \frac{\tau_A}{\tau_B} = 1.57$ that is less than what is commonly assumed.

The Sagnac effect, used in accelerometers, is based on the same idea : two laser beams are sent in a loop in opposite direction : their 4 dimensional paths are not the same, and the difference in the 4 dimensional lengths can be measured by interferometry.

3.3.4 Can we travel faster than light ?

The relation in a change of gauge gives the transformation of the components of vectors in the gauges of two observers at the same point. The quantity $\sqrt{1 - \frac{\|v\|^2}{c^2}}$ tells us that, under the assumptions that we have made, the relative spatial speed of two observers must be smaller than c. It is also well known, and experimentally checked, that the energy required to reach c would be infinite. But the real purpose of the question is : can we shorten the time needed to reach a star ? As we have seen in the twins paradox, this time is : $\int_A^B d\tau = c(\tau_B - \tau_A)$ that is the relativist distance between two points A,B. So it depends only on the path, whatever we do, even with a "drive"... The issue is then : are there shortcuts ? The usual answer is that light always follows the shortest path. However it relies on many assumptions. We will see that light propagates at c, this does not imply that the field uses the shortest path, which is another issue. And asking the backing of photon does not bring much, as the path followed by a photon is just another assumption. The answer lies in our capability to compute the trajectory of a material body. It is possible to model the trajectories of particles in GR (this is one of the topic of Chapter 7), but their solutions rely on the knowledge of the gravitational field, which is far from satisfying, all the more so in interstellar regions. So, from my point of view, the answer is : perhaps.

3.3.5 Cosmology

General Relativity has open the way to a "scientific Cosmology", that is the study of the whole Universe and in particular of its evolution, through mathematical models. These theories will never achieve a full scientific status, because they lack one of the key criteria : the possibility to experiment with other universes. They can provide plausible explanations, but not falsifiable ones. This is reflected in the choice of the parameters which are used in the models : one can fine tune them in order to fit with observations, essentially astronomical observations, and represent in a satisfying way "what it is", but not tell "why is it so".

One of the issue of Cosmology is that of the observer, who is an essential part of Relativity. Particles (and galaxies can be considered as particles at this scale) follow world lines. Their location, which is absolute in GR, is precisely defined with respect to a proper time, but this time is specific to each particle. An observer can follow particles which are in his present, and establish a relation between his proper time and that of these particles. A Cosmological model is a model for an observer who would have access to the locations of all the particles and the universe, and indeed the existence of a universal time, which provides a folliation in hypersurfaces analogous to $\Omega_3(t)$, is one of their key component.

3.3.6 The expansion of the Universe

A manifold by itself can have some topological properties. It can be compact. It can have holes, defined through homotopy (Maths.10.4.1) : there is a hole if there are curves in M which cannot be continuously deformed to be reduced to a point. A hole does not imply some catastrophic feature : a doughnut has a hole. Thus it does not imply that the charts become singular. But there are only few purely topological features which can be defined on a manifold, and they are one of the topic of Differential Geometry. In particular a manifold has no shape to speak of.

The metric on M is an addition to the structure of the Universe. It is a mathematical feature from which more features can be defined on M, such that curvature (we will see it in another chapter). In GR the metric, and so the curvature of M at a point, depends on the distribution of matter. It is customary (see Wald) to define singularities in the Universe by singularities of geodesics, but geodesics are curves whose definition depends on the metric. A singularity for the metric, as Black holes or Bing Bang, is not necessarily a singular point for the manifold itself.

From some general reasoning and Astronomical observations, it is generally assumed that the Universe has the structure of a fiber bundle with base $\mathbb{R}$ (a warped Universe) which can be seen as the generalization of M_o, that we have defined above for an observer. Thus there is some universal time (the projection from M to $\mathbb{R}$) and a folliation of M in hypersurfaces similar to $\Omega_3(t)$, which represent the present for the observers who are located on them (see Wald and Peebles for more on this topic). This is what we have defined as a material body : the part of the universe on which stands all matter would be a single body moving together since the Big Bang (the image of an inflating balloon). So there would not be any physical content before or after this $\Omega_3(t)$ (inside the balloon), but nothing can support this interpretation, or the converse, and probably it will never be.

The Riemannian metric $\varpi_3(t)$ on each $\Omega_3(t)$ is induced by the metric on M, and therefore depends on the universal time t. In the most popular models it comes that the distance between two points on $\Omega_3(t)$, measured by the Riemannian metric, increases with t, and this is the foundation of the narrative about an expanding universe, which is supported by astronomical observations. But, assuming that these models are correct, this needs to be well understood. The change of the metric on $\Omega_3(t)$ makes that the volume form $\varpi_3(t)$ increases, but the hypersurfaces $\Omega_3(t)$ belong to the same manifold M, which does not change with time. The physical universe would be a deformable body, whose volume increases inside the unchanged container. Moreover it is generally assumed that material points, belonging to the same material body but traveling on their own world lines, stick together : they are not affected by this dilation, only the vacuum which separates material bodies.

Chapter 4

KINEMATICS

Kinematics is the next step after Geometry. We go further in the physical world, and try to understand what are the relations between the motion of material bodies and the forces which are applied to them. All material bodies manifest some resistance to a change of their motion, either in direction or in speed. This feature is the inertia, and is measured by different quantities which incorporate, in one way or another, the mass of the material body. The mass m is a characteristic of the material body : it does not change with the motion, the forces or the observer. And from motion and mass are defined key quantities : the momenta.

The Newton's law : $\overrightarrow{F} = m\overrightarrow{\gamma}$ is expressed, more appropriately by : $\overrightarrow{F} = \frac{d\overrightarrow{p}}{dt}$ where $\overrightarrow{p}$ is the momentum. The inertial forces are, by construct, the opposite of the forces which are necessary to change the momenta of a material body. So, Kinematics is, in many ways, the Physics of inertia. The issue of the origin of these inertial forces, which appear everywhere and with a great strength, will be seen in the next chapter. Let us see now how one goes from Geometry, that is motion, to Kinematics, that is inertia.

The study of rotations and rotational momenta in the 4 dimensional Universe will lead to a new representation of the momenta, based on Spinors, which are vectors in a 4 dimensional abstract space, to the introduction of antiparticles and of spin.

4.1 TRANSLATIONAL AND ROTATIONAL MOMENTA

4.1.1 Translational Momentum in the relativist context

In Newtonian Physics the bridge between Geometry and Kinetics is hold by momenta. And, because motion of material bodies involves both a translational motion and a rotational motion, there are linear momenta and rotational momenta.

In Galilean Geometry the linear momentum of a particle is simply : $\overrightarrow{p} = \mu \overrightarrow{v}$ with a constant scalar μ which is the inertial mass. It has the unit dimension $[M][L][T]^{-1}$. Its natural extension in the relativist context is the quadri-vector : $P = \mu u$ where $u = \frac{dm}{d\tau}$ is the velocity. This generalization has three consequences :

- the quadri-vector P is intrinsic : its definition does not depend on an observer

- but its measure depends on the observer. In his standard basis it reads :

$P = \mu \frac{1}{\sqrt{1-\frac{\|\overrightarrow{v}\|^2}{c^2}}} (\overrightarrow{v} + c\varepsilon_0(m))$

- its Lorentz norm does not depend on the spatial speed : $\langle P, P \rangle = -\mu^2 c^2$

In the relativist context location and motion are absolute. If the Universe has a physical meaning, then each of its points is singular, and this is clearly represented by a manifold. The proper time, and the derivative of the location with respect to the proper time, are defined without any reference to a frame, so the vector velocity u is absolute, and this property has been used to compute the rules in a change of frames. If motion is absolute, *its measure is relative*, depends on the observer and its measure changes according to geometric rules, because they are geometric quantities. The spatial speed appears when an observer has been chosen. The definition of the momentum by $P = \mu u$ is consistent with the idea that the kinematic features of a particle are intrinsic, and can be represented by a quantity which does not depend on an observer (even if its measure depends on it).

This is a big change from the Newtonian definitions : the momentum $\overrightarrow{p} = \mu \overrightarrow{v}$ as well as the kinetic energy $\frac{1}{2}\mu \|\overrightarrow{v}\|^2$ are relative and depend on the observer.

If we keep the concept of Energy as measured by [Momentum]×[Speed] then the energy of the particle $\langle P, u \rangle = -\mu c^2$ is constant and, for an observer, is split between a part related to the spatial speed $\mu \frac{\|\overrightarrow{v}\|^2}{1-\frac{\|\overrightarrow{v}\|^2}{c^2}}$, corresponding to a kinetic energy, and a part which is stored in the particle $-\mu \frac{c^2}{1-\frac{\|\overrightarrow{v}\|^2}{c^2}}$. But, if one wants to keep the principle of conservation of energy, one has to accept that mass itself can be transformed into energy, according to the famous relation $E = \mu c^2$.

However there are several interpretations of these concepts. Physicists like to keep a concept of momentum linked to the spatial velocity and, with a fixed mass, define the linear momentum as : $\overrightarrow{p_r} = \mu \frac{\overrightarrow{v}}{\sqrt{1-\frac{\|\overrightarrow{v}\|^2}{c^2}}}$, that is the spatial part of P. Then they define the Energy E of the particle by : $E^2 = c^2 \|\overrightarrow{p}_r\|^2 + \mu^2 c^4$ that is one part corresponding to a kinetic energy, and another one to an energy at rest. This sums up to define the energy by rewriting Pc with the two components :

$Pc = (c\overrightarrow{p_r}, E) \Rightarrow \langle Pc, Pc \rangle = -\mu^2 c^4 = c^2 \|\overrightarrow{p}_r\|^2 - E^2$

And $E = c^2 \mu \frac{1}{\sqrt{1-\frac{\|\overrightarrow{v}\|^2}{c^2}}}$ is the projection of $Pc = \mu c$ along the axis ε_0 : $E = \langle Pc, \varepsilon_0 \rangle = \mu c \left\langle \frac{dm}{d\tau}, \varepsilon_0 \right\rangle$.

The introduction of c in $c\overrightarrow{p_r}$ is necessary in order to have the same unit Energy in both parts. In this formulation Pc is a 4 vector, and its components change according to the Lorentz formula, so E depends on the observer. The advantages of this expression is that for small speed it gives :

$E = c^2\mu\frac{1}{\sqrt{1-\frac{\|\overrightarrow{v}\|^2}{c^2}}} \simeq c^2\mu\left(1+\frac{1}{2}\frac{\|\overrightarrow{v}\|^2}{c^2}\right) = \frac{1}{2}\mu\left\|\overrightarrow{v}\right\|^2 + \mu c^2$
and it can be adapted to massless particles such as photons.

The only physical quantity which has a clear meaning, is independent of an observer, and is characteristic of a particle, is Pc, which has the dimension of energy. The mass at rest, independent from an observer, is $\mu = \frac{1}{c}\sqrt{-\langle Pc, Pc\rangle}$. The usual energy is computed by taking the component of Pc along the direction of $\varepsilon_0(m)$. So it depends on the observer. These quantities are individualized and localized : they are linked to the particle and its position. And we remind that, even if we use freely the words "energy of a particle", it does not mean that this quantity has an absolute value, but rather it measures the energy that could be potentially exchanged by the particle with other physical objects. So if one can say, with Feynman (in *Lectures in Physics*), that the conservation of energy is a law without exception, the trouble is that in the Relativist context the definition of energy itself depends on the observer and is fairly subtle.

In writing $Pc = (c\overrightarrow{p_r}, E)$ the energy E and p_r are seen as independent variables (which they are not). In the usual interpretation of QM to E and p_r are associated operators acting on wave functions. The substitution in $E^2 = c^2\left\|\overrightarrow{p}_r\right\|^2 + \mu^2c^4$:$E \to i\hbar\frac{\partial}{\partial t}; p_{r\alpha} \to -i\hbar\partial_\alpha$ gives the Klein-Gordon equation : $\left(\Box + \mu^2\right)\psi = 0$ which, checked for the spectrum of Hydrogen, provided wrong results. Dirac proposed another equation, based on the relation $E = \sqrt{c^2\left\|\overrightarrow{p}_r\right\|^2 + \mu^2c^4}$ which would be on first order in the derivatives :
$E = A.p_r + B\mu \to i\hbar\frac{\partial\psi}{\partial t} = (Ai\hbar\nabla + B\mu)\psi$ and one can check that this is possible only if ψ is a vectorial quantity (and no longer a scalar function). Moreover to be Lorentz equivariant A, B must be 4×4 complex matrices, built from a set of matrices with the relation : $\gamma_i\gamma_j + \gamma_j\gamma_i = 2\eta_{ij}I_4$. The wave vectors ψ are then vectors, belonging to a 4 dimensional complex vector space F, which is the representation of a group which acts through the matrices γ. They are called spinors. Meanwhile in the usual interpretation ψ is a scalar function such that $|\psi(m)|^2$ gives the probability of finding a particle in m, the kinematics characteristics of the particle are represented by a spinor, and the Dirac's equation shows that, along its trajectory, ψ must meet some conservation law. The scheme has been extended to account for the action of the fields, and leads to the standard model. In QM the translational momentum is seem mainly as an operator, and used with Fourier transform. With the Dirac's equation the formalization of the concept is further from the usual vector $\overrightarrow{p}_r = \mu\overrightarrow{v}$.

Moreover material bodies have a rotation, and a rotational momentum $\overrightarrow{J}$ which depends on the mass and the shape of the body. Momenta are characteristics features of material bodies, and the bases for the definition of forces. According to the Newton's law: $\overrightarrow{F} = \frac{d\overrightarrow{p}}{dt}$ and $\overrightarrow{\gamma} = \frac{d\overrightarrow{J}}{dt}$ for a torque $\overrightarrow{\gamma}$. Rotation and rotational momentum are topics which are more complicated than it seems, in the relativist context in particular. Moreover at the atomic scale particles show properties which look like rotation, and specific features, which have lead to the concepts of spins. These are the main topics of this chapter.

4.1.2 The issues of the concept of rotation

Galilean Geometry

Rotation

The concept of rotation is well defined in Mathematics : this is the operation which transforms the *orthonormal basis* of a vector space into another. In Galilean Geometry it is represented by a matrix R of the group $SO(3)$. This is a compact, 3 dimensional Lie group, of matrices such that $R^tR = I$. Because of this relation the Lie algebra $so(3) = T_1 SO(3)$ is the vector space of 3×3 real antisymmetric matrices. If we take the following matrices as basis of $so(3)$:

$$\kappa_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} ; \kappa_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} ; \kappa_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

then any matrix of $so(3)$ reads :
$\sum_{i=1}^{3} r^i [\kappa_i] = [j(r)]$ with the operator

$$j : \mathbb{R}^3 \to L(\mathbb{R}, 3) :: [j(r)] = \begin{bmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{bmatrix} \tag{4.1}$$

The operator j is very convenient to represent quantities which are rotated [1] and has many nice algebraic properties (see formulas in the Annex) and we will use it often in this book.

For any vector u : $\sum_{ij=1}^{3} [j(r)]^i_j u^j \varepsilon_i = \overrightarrow{r} \times \overrightarrow{u}$ with the cross product $\times$.

The vector r is just the components of a vector in a Lie algebra, using a specific basis κ. However there is a natural correspondence between r and geometric characteristics of a rotation.

The axis of rotation, which is by definition the unique eigen vector of $[g]$ with eigen value 1 and norm 1 in the standard representation of $SO(3)$, has for components $\begin{bmatrix} r^1 \\ r^2 \\ r^3 \end{bmatrix} / \sqrt{r^t r}$

Similarly one can define the angle θ of the rotation resulting from a given matrix, and $\theta = \sqrt{r^t r}$

Proof. For any vector u of norm 1 : $\langle u, [g] u\rangle = \cos\theta$ where θ is an angle which depends on u and $[g] = \exp[j(r)]$. With the formula above, and using $[j(r)][j(r)] = [r][r]^t - \langle r, r\rangle I$ and $\langle u, [j(r)] u\rangle = 0$ we get :

$\langle u, [g] u\rangle = 1 + \left(\langle u, r\rangle^2 - \langle r, r\rangle\right) \frac{1-\cos\sqrt{r^t r}}{r^t r}$

which is minimum for $\langle u, r\rangle = 0$ that is for the vectors orthogonal to the axis, and :

$\cos\theta = \cos\sqrt{r^t r}$ ■

Rotational motion

We use freely the same word "rotation" for the operation to go from one orthonormal basis to another, and for the motion (the instantaneous rotation around an axis), but these are two distinct concepts.

If 2 orthonormal bases are in relative motion, at any time t we have some rotation $R(t) \in SO(3)$ and naturally one defines the instantaneous rotation through the derivative $\frac{dR}{dt}$. To have a more convenient quantity one considers $R(t)^{-1} \frac{dR}{dt} \in so(3)$. And we are fortunate because $SO(3)$ is a compact Lie group, and the exponential is surjective :

$\forall g \in SO(3), \exists \kappa = \sum_{i=1}^{3} r^i [\kappa_i] \in so(3) : g = \exp\kappa$

$\frac{d}{dt} \exp t\kappa|_{t=\theta} = (\exp\theta\kappa)\kappa$

It is easy to show that :

$[g] = \exp[j(r)] = I_3 + [j(r)] \frac{\sin\sqrt{r^t r}}{\sqrt{r^t r}} + [j(r)][j(r)] \frac{1-\cos\sqrt{r^t r}}{r^t r}$

A rotation at constant speed ϖ is such that : $\theta(t) = \varpi t = t\sqrt{r^t r}$ and has for period $T = 2\pi/\sqrt{r^t r}$.

[1] It is similar to the Levi-Civitta tensor ϵ but, in my opinion, much easier to use.

The derivative $\frac{dR}{dt}$ of R is $\frac{dR}{dt} = R(t)\kappa$ where $\kappa \in so(3)$ and $R(t)^{-1}\frac{dR}{dt} = \kappa = j(r)$. So the instantaneous rotational motion can be assimilated to a rotation which would be at a constant rotational speed $\sqrt{r^t r}$: $R(t) = \exp tj(r)$

So we have a very satisfying representation of geometric rotations : a rotation R can be defined by a single vector, which is simply related to essential characteristics of the transformation, and an instantaneous rotational movement can also be represented by a single vector r. But, as one can see, this model is less obvious than it seems. It relies on the fortuitous fact that the Lie algebra has the same dimension as the Euclidean space (the dimension of $so(n)$ is $\frac{n(n-1)}{2}$) and is compact.

Spin group

Moreover this mathematical representation is not faithful. The same rotation can be defined equally by the opposite axis, and the opposite angle. This is related to the mathematical fact that $SO(3)$ is not the only group which has $so(3)$ as Lie algebra. The more general group is the Spin group $Spin(3)$ which has also for elements the scalars + 1 and - 1, so that $R(t)$, corresponding to $\left(r, \sqrt{r^t r}\right)$ and $-R(t)$, corresponding to $\left(-r, -\sqrt{r^t r}\right)$ can represent the same physical rotational motion. Actually, the group which should be used to represent rotations in Galilean Geometry is $Spin(3)$, which makes the distinction between the two rotations, and not $SO(3)$. This is not a problem in Mathematics, but in Physics the distinction matters : in the real world one goes from one point to another along a path, by a continuous transformation which preserves the orientation of a vector, thus the orientation of $\overrightarrow{r}$ matters [2]. A single vector of $\mathbb{R}^3$ cannot by itself properly identify a physical rotation, one needs an additional parameter which is ± 1 (to tell which one of the two orientations of $\overrightarrow{r}$ is chosen, with respect to a direction, the spatial speed on the path.

Rotational momenta

But to represent rotation of material bodies by geometric rotations, as above, raises several issues in Physics.

We could expect that the total rotational momentum of a body is the sum of the rotational momentum of its components, as it happens with the translational momentum. But material points have no attached frame in Mechanics. So actually the rotational momentum is defined only at the level of the body, through a geometric rotation, and it has a meaning only for material bodies which keep some shape, represented by fixed relations between the positions of the material points. So this is doable for solids, but even for them there is a problem. If the solid has a cylindrical symmetry, by definition it is impossible to measure a geometric rotation around the axis, however the physical rotation can be measured by a rotational inertia. And some force fields, such as magnetism, can exercise a pointwise action, represented by a torque, so that the implementation of the Newton's law to the rotational momentum becomes muddled for particles without structure, like the atoms.

So, even if in Mechanics it is convenient to assimilate physical rotation with geometric rotation, they are not the same and the kinematics of rotating bodies is more complicated.

[2] In his book "The road to reality" Penrose gives a nice, simple trick with a belt and book to show this fact.

Displacements in Galilean Geometry

In the Newtonian Geometry space and time are independent. So one has to considers independently a translation and a rotation to represent a motion.

In Mathematics the operation to go from an orthogonal frame $\left(O, (e_i)_{i=1}^3\right)$ to another one $\left(G, (e_i')_{i=1}^3\right)$ combines a translation $D \in T$ and a rotation $g \in SO(3)$, which belongs to the group of displacement, which is the semi-direct product $T \ltimes SO(3)$. The "semi" implies some relations (Maths.1824) which makes the structure of the group of displacements more complicated than the direct product $T \times SO(3)$. In particular the Lie algebra of $T \ltimes SO(3)$ is not the direct product of the Lie algebras $T_1 T \times so(3)$. This can be seen in the law for the composition of speeds for rotating bodies. So it is mathematically incorrect to represent a motion, combining a translation and a rotation, as the simple couple of vectors $(\overrightarrow{v}, \overrightarrow{r})$. But this is the common practice of Physicists, for the good reason that they consider separately the motion of the center mass G, and the rotational motion, which has a meaning only for solids : they do not represent the same concept.

The group of displacements in Relativist Geometry

In Relativist Geometry the focus is usually put on the Poincaré's group, the semi product of the group $SO(3,1)$ of rotations and of the 4 dimensional translations. This is the simple generalization of the group of displacements of Galilean Geometry, with the same problem as for 3 dimensions. However this raises several additional issues.

The Poincaré's group represents the operation to go from one orthonormal frame $\left(O, (\varepsilon_i)_{i=0}^3\right)$ to another $\left(A, (e_i)_{i=0}^3\right)$. So its use is valid only in SR, and for inertial observers. According to the Principle of Locality the location (O) of the origin of the frame has no physical meaning : we should compare two frames, located at the same point (as we did to prove the formulas to go from one observer to another). The translation to go from O to A has nothing to do in the matter, and the spatial speed $\overrightarrow{v}$ in the formulas is the relative speed with respect to a "copy" of the observer who would be at the same location as the body. All the more so that, to go from O to A involves, in any physical measure, a continuous path : one does not jump from one point to another in the physical world, and the path matters because the universe is no longer isotropic. Moreover the usual concept of solid is not generalizable in Relativity, thus it would be hazardous to base the Kinematics theory on a concept which is not physical.

To represent the motion of an orthogonal frame with respect to another this is not the Poincaré's group which is relevant, but its Lie algebra, which is not the simple product of the Lie algebra.

But there is an additional issue in 4 dimension. The exponential is not surjective for $SO(3,1)$, which is not a compact group. We have $[\chi] = \exp[K(w)]\exp[J(r)]$ where $[K(w)], [J(r)] \in so(3,1)$ thus the derivative $\frac{d\chi}{dt}$ gives a more complicated expression, where $\frac{dw}{dt}, \frac{dr}{dt}$ are mixed with (w, r). In particular appears $\frac{dw}{dt}$, that is the derivative of the spatial speed.

There has been attempts to extend the concept to the group of isometries (that is maps $f: M \to M$ such that the derivative $f'(m)$ preserves the scalar product) but this just introduces additional complications and the objections above still holds.

In Galilean Geometry the points of view of the Physicists and of the Mathematicians are not the same, even if they both use two vectors and the same number of parameters, they have not the same meaning. In Relativity the discrepancy is worse : an element of the Poincaré's group is defined by 10 parameters, of which 4 (related to OA) are not physically relevant. Relativity, Special or General, is an extension of Newtonian Physics, and the common perception that the motion of a material body can be described with 6 parameters only still holds. For a physicist the motion of a material body is related to the instantaneous change of its location and disposition (as it is done in Galilean Geometry with $(\overrightarrow{v}, \overrightarrow{r})$) and not to the transformation between fixed frames (the vector $\overrightarrow{L} = \overrightarrow{OG}$ has no physical interest). So the Lorentz matrix (defined by two 3 dimensional vectors (w, r) related to $(\overrightarrow{r}, \overrightarrow{v})$) only is significant from this point of view, and w can be clearly (even if it is in a complicated way) related to $\overrightarrow{v}$. Moreover we have seen that the formulas hold for any observer.

Our purpose here is to find a way to represent kinematic characteristics of material bodies, by vectorial quantities. This is not to find the formulas in a change of coordinates (which involve the charts, not the bases in the tangent space), so the Poincaré's group is of no use. But the Lorentz group is essential because it gives the rules for the transformation of the components of vectors.

Momenta in the fiber bundle representation

Fiber bundles provide an efficient representation of the geometry of the Universe, notably in the GR picture. So it is legitimate to look at what it can provide on this issue. The aim is to represent the kinematic characteristics of material bodies, incorporating both their geometric motion (translation and rotation) and inertial features, in a single quantity, which can be implemented for particles, that is without the need to involve a fixed structure of the body. To define a motion, or a momentum, we need first to find a way to define, at the same point, two orthonormal bases, one specific to the body, and the other to the observer, which can be compared. This is just the purpose of the principal bundle : we have two gauges $p_B = \varphi_G(m, g(m)), \mathbf{p}_O(m) = \varphi_G(m, \mathbf{1})$ at the same point m, which defines two orthogonal frames (two tetrads) in the associated vector bundle $P_G[\mathbb{R}^4, SO_0(3,1)]$.

We assume that there is some quantity S, which characterizes the kinematics of the body and stays constant in the gauge p_B on its worldline, meanwhile its measure by an observer, done with $\mathbf{p}_O$ at each point $m = q(t)$ of the trajectory of the body, can change with the observer. This is just the purpose of associated bundles. So we assume the existence of an associated bundle $P_G[E, \gamma]$ with some manifold E and an action γ of $SO_0(3,1)$ on E. The gauge of the observer is $\mathbf{p}_o(t) = \varphi_G(q(t), \mathbf{1})$ and the gauge of the body is: $p_B(t) = \varphi_G(q(t), g(t(\tau)))$ for some $g(t(\tau)) \in SO_0(3,1)$.

The value measured by the observer in his gauge is then $S(t)$ but it is a constant S_0 measured in the gauge of the body and we have, by definition :

$(\mathbf{p}_o(t), S(t)) = (p_B(t), S_0) = (\mathbf{p}_o(t) \cdot g(t(\tau)), S_0) \sim (\mathbf{p}_o(t), \gamma(g(t(\tau)), S_0))$

which is equivalent to say that $S(t) = \gamma(g(\tau(t)), S_0)$ with S_0 constant.

The relation $S(t) = \gamma(g(\tau(t)), S_0)$ is just the consequence of our very general assumptions. The measure of $S(t)$ varies locally according to the observer (the measure is relative) but its intrinsic value does not change (it is absolute).

S_0 and $S(t)$ do not represent the momentum : this is done by the derivative with respect to the time t of the observer. We want a vectorial representation of the momentum, in a fixed vector space, that is on a vector bundle.

The formula $S(t) = \gamma(g(\tau(t)), S_0)$ gives by derivation a vector on the tangent bundle to the associated bundle, which is an associated bundle $TP_G[TE, TSO_0(3,1)]$ (Maths.2086). If we want to have a fixed vector space, E must be itself a vector space, $P_G[E, \gamma]$ must be a vector bundle, and (E, γ) a representation of $SO_0(3,1)$.

$\gamma(g,S) = \gamma(g)S$
Then the momenta are represented by
$\frac{dS}{dt} = \gamma'(g)\frac{dg}{dt}S_0 = \gamma'_g(1)L'_{g^{-1}g}\left(\frac{dg}{dt}\right)S_0$ and $L'_{g^{-1}g}\left(\frac{dg}{dt}\right) \in so(3)$
They are vectorial quantities, which can be linearly combined.

But to get a full profit of this representation we have to adopt an entirely new point of view. We cannot any longer view the particle as living in M and spinning in its tangent space. Actually the particle lives in E, which happens to be associated to P_G. Its trajectory is a curve in E, which projects on a curve in M. E can be seen as the physical world (at least a part of it), that we can represent through networks of frames in M. So $\frac{dg}{dt}$ cannot be seen properly as a motion, it is only a characteristic of the particle (such as mass and charge). Experience shows that it can be measured through geometric frames but this does not imply the existence of a real spinning motion of the particle. And one cannot assimilate S with a momentum.

The issue that we face is then to precise E and γ. We will make the following, reasonable, assumptions :

i) E is some vector space, so that we have an associated vector bundle $P_G[E,\gamma]$

ii) It implies that (E,γ) is a representation of $SO_0(3,1)$.

iii) We have seen that the right group which should be used is $Spin(3,1)$, so (E,γ) is a representation of $Spin(3,1)$.

iv) This representation should be finite dimensional (we consider here the value of $S(t)$ at some point).

iv) If (E,γ) is a representation of $Spin(3,1)$, then $(E,\gamma'(1))$ is a representation of its Lie algebra, both are subsets of the Clifford algebra $Cl(3,1)$ so that, if γ is a linear map, then $\gamma'(1) = \gamma$ and this is not a big leap forward to assume that (E,γ) is a representation of the Clifford algebra itself.

Then the quantity S is a vector of E, called a **spinor**.

4.2 CLIFFORD ALGEBRAS AND SPINORS

Spinors, as well as the spin, cannot be properly understood without a look at their mathematical berth, which is Clifford algebra. This is a fascinating algebraic structure on vector spaces which is seen in details in Maths.Part 9. The results which will be used in this book are summarized in this section, the proofs are given in the Annex.

4.2.1 Clifford algebra and Spin groups

Clifford Algebras

A Clifford algebra $Cl(F, \langle\rangle)$ is an algebraic structure, which can be defined on any vector space $(F, \langle\rangle)$ on a field K ($\mathbb{R}$ or $\mathbb{C}$) endowed with a bilinear *symmetric* form $\langle\rangle$. The set $Cl(F, \langle\rangle)$ is defined from K, F and a product, denoted $\cdot$, with the property that for any two *vectors* u, v :

$$\forall u, v \in F : u \cdot v + v \cdot u = 2 \langle u, v \rangle \tag{4.2}$$

A Clifford algebra is then a set which is larger than F : it includes *all vectors* of F, plus scalars, and any linear combinations of products of vectors of F. A Clifford algebra on a n dimensional vector space is a 2^n dimensional vector space on K, and an algebra with $\cdot$. Clifford algebras built on vector spaces on the same field, with same dimension and bilinear form with same signature are isomorphic. On a 4 dimensional real vector space $(F, \langle\rangle)$ endowed with a Lorentz metric there are two structures of Clifford Algebra, denoted $Cl(3,1)$ and $Cl(1,3)$, depending on the signature of the metric, and they are *not* isomorphic. In the following we will state the results for $Cl(3,1)$, and for $Cl(1,3)$ only when they are different.

The easiest way to work with a Clifford algebra is to use an orthonormal basis of F. On any 4 dimensional real vector space $(F, \langle\rangle)$ with a bilinear symmetric form of signature (3,1) or (1,3) we will denote :

Notation 57 $(\varepsilon_i)_{i=0}^3$ *is an orthonormal basis with scalar product :* $\langle \varepsilon_i, \varepsilon_i \rangle = \eta_{ii}$

So we have the relation :

$$\varepsilon_i \cdot \varepsilon_j + \varepsilon_j \cdot \varepsilon_i = 2\eta_{ij} \tag{4.3}$$

Then a basis of the Clifford algebra is a set comprised of 1 and all ordered products of $\varepsilon_i, i = 0...3$.

In any orthonormal basis there is a fourth vector which is such that $\varepsilon_i \cdot \varepsilon_i = -1$ (for the signature $(3,1)$) of $+1$ (for the signature $(1,3)$). In the following of this book we will always assume that the orthonormal basis is such that ε_0 is the 4th vector : $\langle \varepsilon_0, \varepsilon_0 \rangle = -1$ with signature (3,1) and $\langle \varepsilon_0, \varepsilon_0 \rangle = +1$ with signature (1,3). We will label this vector ε_0.

Notation 58 $\varepsilon_5 = \varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$

$\varepsilon_5 \cdot \varepsilon_5 = -1$

Spin group

The group $Pin(3,1)$ is the subset of the Clifford algebra $Cl(3,1)$:

$Pin(3,1) = \{u_1 \cdot u_2 ... \cdot u_k, \langle u_p, u_p \rangle = \pm 1, u_p \in F\}$. $Pin(3,1)$ is a Lie group,

Spin(3,1) is its subgroup where we have an even number of vectors :

$Spin(3,1) = \{u_1 \cdot u_2 ... \cdot u_{2k}, \langle u_p, u_p \rangle = \pm 1, u_p \in F\}$

and similarly for $Pin(1,3)$ and $Spin(1,3)$.
Notice that the *scalars* ± 1 belong to the groups. The identity element is the scalar 1.
$Pin(3,1)$ and $Pin(1,3)$ are not isomorphic. $Spin(3,1)$ and $Spin(1,3)$ are isomorphic.

Adjoint map

For any $s \in Pin(3,1)$, the map, called the **adjoint map** :

$$\mathbf{Ad}_s : Cl\,(3,1) \to Cl\,(3,1) :: \mathbf{Ad}_s w = s \cdot w \cdot s^{-1} \tag{4.4}$$

is such that

$$\forall w \in F : \mathbf{Ad}_s w \in F \tag{4.5}$$

and it preserves the scalar product on F :

$$\forall u, v \in F, s \in Pin(3,1) : \langle \mathbf{Ad}_s u, \mathbf{Ad}_s v \rangle_F = \langle u, v \rangle_F \tag{4.6}$$

Moreover :

$$\forall s, s' \in Pin(3,1) : \mathbf{Ad}_s \circ \mathbf{Ad}_{s'} = \mathbf{Ad}_{s \cdot s'} \tag{4.7}$$

Because the action $\mathbf{Ad}_s$ of $Spin(3,1)$ on F gives another vector of F and preserves the scalar product, it can be represented by a 4×4 orthogonal matrix. Using any orthonormal basis $(\varepsilon_i)_{i=0}^3$ of F, then $\mathbf{Ad}_s$ is represented by a matrix $\Pi\,(\mathbf{Ad}_s) = [h\,(s)] \in SO\,(3,1)$.
$v = \sum_{i=0}^3 v^i \varepsilon_i \to \widetilde{v} = \mathbf{Ad}_s v = \sum_{i=0}^3 v^i s \cdot \varepsilon_i \cdot s^{-1} = \sum_{i=0}^3 \widetilde{v}^i \varepsilon_i$
$\widetilde{v}^i = \sum_{j=0}^3 [h\,(s)]^i_j \, v^j$
To two elements $\pm s \in Spin(3,1)$ correspond a single matrix $[h\,(s)]$. $Spin(3,1)$ is the double cover (as manifold) of $SO(3,1)$. $Spin(3,1)$ has two connected components (which contains either +1 or -1) and its connected component, that we will denote for brevity also $Spin\,(3,1)$, is simply connected and is the universal cover group of $SO_0\,(3,1)$. So with the Spin group one can define two physical rotations, corresponding to opposite signs.

Lie algebra of the Spin group

Theorem 59 *The elements of the Lie algebra $T_1 Spin\,(3,1)$ belong to the Clifford algebra and can be written as the linear combination of elements $\varepsilon_i \cdot \varepsilon_j$*

As any algebra $Cl\,(F, \langle\rangle)$ is a Lie algebra with the bracket :
$\forall w, w' \in Cl\,(F, \langle\rangle) : [w, w'] = w \cdot w' - w' \cdot w$
and the Lie algebra $T_1 Spin\,(3,1)$ of $Spin(3,1)$ is a subset of $Cl\,(3,1)$ (Maths.532).
The derivative $\Pi'\,(1) : T_1 Spin\,(3,1) \to so\,(3,1)$ is an isomorphism of Lie algebras. The inverse map : $\Pi'\,(1)^{-1} : so(3,1) \to T_1 Spin\,(3,1)$ is an isomorphism of Lie algebras which reads (Maths.534) with any orthonormal basis $(\varepsilon_i)_{i=0}^3$ of F :
$\Pi'\,(1)^{-1} : so(3,1) \to T_1 Spin\,(3,1) :: \Pi'\,(1)^{-1}\,([\kappa]) = \frac{1}{4} \sum_{i,j=0}^3 ([\kappa]\,[\eta])^i_j \, \varepsilon_i \cdot \varepsilon_j$
and any element of $T_1 Spin\,(3,1)$ is such expressed in the basis of $Cl\,(F, \langle\rangle)$: it is the linear combinations of the ordered products of all the four vectors of a basis.
With any orthonormal basis and the following choices of basis $(\overrightarrow{\kappa}_a)_{a=1}^6$ of $T_1 Spin\,(3,1)$ then $\Pi'\,(1)^{-1}$ takes a simple form with an adequate ordering of the vectors (which is convenient and will be used quite often) :

$\Pi'(1)^{-1}([\kappa_1]) = \overrightarrow{\kappa}_1 = \frac{1}{2}\varepsilon_3 \cdot \varepsilon_2,$
$\Pi'(1)^{-1}([\kappa_2]) = \overrightarrow{\kappa}_2 = \frac{1}{2}\varepsilon_1 \cdot \varepsilon_3,$
$\Pi'(1)^{-1}([\kappa_3]) = \overrightarrow{\kappa}_3 = \frac{1}{2}\varepsilon_2 \cdot \varepsilon_1,$
$\Pi'(1)^{-1}([\kappa_4]) = \overrightarrow{\kappa}_4 = \frac{1}{2}\varepsilon_0 \cdot \varepsilon_1,$
$\Pi'(1)^{-1}([\kappa_5]) = \overrightarrow{\kappa}_5 = \frac{1}{2}\varepsilon_0 \cdot \varepsilon_2,$
$\Pi'(1)^{-1}([\kappa_6]) = \overrightarrow{\kappa}_6 = \frac{1}{2}\varepsilon_0 \cdot \varepsilon_3$
where $([\kappa_a])_{a=1}^6$ is the basis of $so(3,1)$ already noticed such that :
$[\kappa] = K(w) + J(r) = \sum_{a=1}^3 r^a [\kappa_a] + w^a [\kappa_{a+3}]$
$a = 1,2,3 : \overrightarrow{\kappa}_a = -\frac{1}{2}\epsilon(a,i,j)\,\varepsilon_i \cdot \varepsilon_j, a = 4,5,6 : \overrightarrow{\kappa}_a = \frac{1}{2}\varepsilon_0 \cdot \varepsilon_{a-3}$ is a basis of $T_1 Spin(3,1)$
We will use extensively the convenient (the order of the indices matters) :

Notation 60 *for both $Cl(3,1), Cl(1,3)$:*

$$\upsilon(r,w) = \frac{1}{2}\left(w^1\varepsilon_0 \cdot \varepsilon_1 + w^2\varepsilon_0 \cdot \varepsilon_2 + w^3\varepsilon_0 \cdot \varepsilon_3 + r^3\varepsilon_2 \cdot \varepsilon_1 + r^2\varepsilon_1 \cdot \varepsilon_3 + r^1\varepsilon_3 \cdot \varepsilon_2\right) \tag{4.8}$$

With this notation, whatever the orthonormal basis $(\varepsilon_i)_{i=0}^3$, any element X of the Lie algebras $T_1 Spin(3,1)$ or $T_1 Spin(1,3)$ reads :

$$X = \upsilon(r,w) = \sum_{a=1}^{3} r^a \overrightarrow{\kappa}_a + w^a \overrightarrow{\kappa}_{a+3} \tag{4.9}$$

with $(r,w) \in \mathbb{R}^3 \times \mathbb{R}^3$ then $X = \upsilon(r,w)$ is the image of :
$\Pi'(1)(\upsilon(r,w)) = K(w) + J(r) \in so(3,1)$ if $X \in T_1 Spin(3,1)$
$\Pi'(1)(\upsilon(r,w)) = -(K(w) + J(r)) \in so(1,3)$ if $X \in T_1 Spin(1,3)$
The bracket on the Lie algebra:
$[\upsilon(r,w), \upsilon(r',w')]$
$= \upsilon(r,w) \cdot \upsilon(r',w') - \upsilon(r',w') \cdot \upsilon(r,w)$
$= \upsilon(j(r)r' - j(w)w', j(w)r' + j(r)w')$
With signature (1,3) :
$[\upsilon(r,w), \upsilon(r',w')] = -\upsilon(j(r)r' - j(w)w', j(w)r' + j(r)w')$

We have the identity :
$\upsilon(r,w) \cdot \varepsilon_5 = \varepsilon_5 \cdot \upsilon(r,w) = \upsilon(-w,r)$

Expression of elements of the spin group

Theorem 61 *The elements of the Spin groups read in both signatures, with the related $a, (w^j, r^j)_{j=1}^3, b$ real scalars :*

$$\begin{bmatrix} s = a + \upsilon(r,w) + b\varepsilon_5 \\ a^2 - b^2 = 1 + \frac{1}{4}(w^t w - r^t r) \\ ab = -\frac{1}{4} r^t w \\ (a + \upsilon(r,w) + b\varepsilon_5)^{-1} = a - \upsilon(r,w) + b\varepsilon_5 \end{bmatrix} \tag{4.10}$$

The exponential is not surjective on $so(3,1)$ or $T_1Spin(3,1)$: for each $\upsilon(r,w) \in T_1Spin(3,1)$ there are two elements $\pm\exp\upsilon(r,w) \in Spin(3,1)$:
$\exp t\upsilon(r,w) = \pm\sigma_w(t) \cdot \sigma_r(t)$ with opposite sign [3]:

[3] These quite awful formulas, which will not be used, show the interest to use the Clifford algebra representation and not the group SO(3,1) itself.

$\sigma_w(t) = \sqrt{1 + \frac{1}{4}w^t w \sinh^2 \frac{1}{2}t\sqrt{w^t w}} + \sinh \frac{1}{2}t\sqrt{w^t w}\upsilon(0, w)$
$\sigma_r(t) = \sqrt{1 - \frac{1}{4}r^r r \sin^2 t\frac{1}{2}\sqrt{r^t r}} + \sin t\frac{1}{2}\sqrt{r^t r}\upsilon(r, 0)$
And we have the identity (Maths.1768) : $\forall \upsilon(r, w) \in T_1 Spin(3, 1)$:

$$\exp \mathbf{Ad}_g \upsilon(r, w) = \mathbf{Ad}_g \exp \upsilon(r, w) = g \cdot \exp \upsilon(r, w) \cdot g^{-1} \tag{4.11}$$

The product $s \cdot s$' reads with the operator j introduced previously :
$(a + \upsilon(r, w) + b\varepsilon_5) \cdot (a' + \upsilon(r', w') + b'\varepsilon_5) = a" + \upsilon(r", w") + b"\varepsilon_5$
with :
$a" = aa' - b'b + \frac{1}{4}(w^t w' - r^t r')$
$b" = ab' + ba' - \frac{1}{4}(w^t r' + r^t w')$
and in $Spin(3, 1)$:
$r" = \frac{1}{2}(j(r)r' - j(w)w') + a'r + ar' - b'w - bw'$
$w" = \frac{1}{2}(j(w)r' + j(r)w') + a'w + aw' + b'r + br'$
and in $Spin(1, 3)$:
$r" = \frac{1}{2}(j(r)r' - j(w)w') + a'r + ar' + b'w + bw'$
$w" = -\frac{1}{2}(j(w)r' + j(r)w') + a'w + aw' + b'r + br'$

Scalar product on the Clifford algebra

There is a scalar product on $Cl(F, \langle\rangle)$ defined by :
$\langle u_{i_1} \cdot u_{i_2} \cdot ... \cdot u_{i_n}, v_{j_1} \cdot v_{j_2} \cdot ... \cdot v_{j_n}\rangle = \langle u_{i_1}, v_{j_1}\rangle \langle u_{i_2}, v_{j_2}\rangle ... \langle u_{i_n}, v_{j_n}\rangle$

It does not depend on the choice of a basis, and any orthonormal basis defined as above is orthonormal :
$\langle \varepsilon_{i_1} \cdot \varepsilon_{i_2} \cdot ... \cdot \varepsilon_{i_n}, \varepsilon_{j_1} \cdot \varepsilon_{j_2} \cdot ... \cdot \varepsilon_{j_n}\rangle = \eta_{i_1 j_1} ... \eta_{i_n j_n} \epsilon(i_1 ..., i_n, j_1, ... j_n)$
the latter term is the signature of the permutation $(i_1 ..., i_n, j_1, ... j_n)$

This scalar product on $Cl(3, 1), Cl(3, 1)$ has the signature $(8, 8)$: it is non degenerate but neither definite positive or negative. It is invariant by **Ad**.

$$\forall w, w' \in Cl(F, \langle\rangle) : \langle \mathbf{Ad}_s w, \mathbf{Ad}_s w'\rangle_{Cl(E,\langle\rangle)} = \langle w, w'\rangle_{Cl(E,\langle\rangle)} \tag{4.12}$$

$(Cl(3, 1), \mathbf{Ad})$ is a unitary representation of $Spin(3, 1)$ and $(Cl(1, 3), \mathbf{Ad})$ a unitary representation of $Spin(1, 3)$.

It reads for elements of $T_1 Spin(3, 1)$:

$$\langle \upsilon(r, w), \upsilon(r', w')\rangle_{Cl} = \frac{1}{4}\left(r^t r' - w^t w'\right) \tag{4.13}$$

4.2.2 Symmetry breakdown

Clifford algebra Cl(3)

The elements of $SO(3, 1)$ are the product of spatial rotations (represented by $\exp J(r)$) and boosts, linked to the speed and represented by $\exp K(w)$. We have similarly a decomposition of the elements of $Spin(3, 1)$. But to understand this topic, from both a mathematical and a physical point of view, we need to distinguish the abstract algebraic structure and the sets on which the structures have been defined.

From a vector space $(F, \langle\rangle)$ endowed with a scalar product one can built only one Clifford algebra, which has necessarily the structure $Cl(3, 1)$: as a set $Cl(3, 1)$ must comprise all the vectors of F. But from any vector subspace of F one can built different Clifford algebras : their algebraic structure depends on the dimension of the vector space, and on the signature of the metric induced on the vector subspace. To have a Clifford algebra structure $Cl\,(3)$ on F one needs a 3 dimensional vector subspace on which the scalar product is definite positive, so it cannot include any vector such that $\langle u,u\rangle < 0$ (and conversely for the signature $(1,3)$: the scalar product must be definite negative). The subsets of F which are a 3 dimensional vector subspace and do not contain any vector such that $\langle u,u\rangle < 0$ are not unique [4]. So we have different subsets of $Cl(3,1)$ with the structure of a Clifford algebra $Cl\,(3)$, all isomorphic but which do not contain the same vectors. Because the Spin Groups are built from elements of the Clifford algebra, we have similarly isomorphic Spin groups $Spin(3)$, but with different elements.

The simplest way to deal with these issues is to fix an orthonormal basis. Any orthonormal basis of F contains one vector such that $\langle \varepsilon_i,\varepsilon_i\rangle = -1$ (or $+1$ with the signature $(1,3)$). If we exclude this vector we can generate a vector subspace $F\,(\varepsilon_0) = Span\,(\varepsilon_i)_{i=1}^3$ and then a Clifford algebra $Cl(3)$. So the identification of a specific set with the structure of $Cl(3)$ sums up to single out such a vector, that we will denote as ε_0.

Decomposition of the elements of the Spin group

Let us choose a vector $\varepsilon_0 : \langle \varepsilon_0,\varepsilon_0\rangle = -1$ (or $+1$ with the signature $(1,3)$). Then there is a unique vector subspace $F^\perp$ orthogonal to ε_0, where the scalar product is definite positive, and from $\left(F^\perp,\langle\rangle\right)$ one can build a unique set which is a Clifford algebra with structure $Cl\,(3)$. Its spin group has the structure $Spin\,(3)$ which has for Lie algebra $T_1 Spin\,(3)$. As proven in the Annex it can be identified with the subset of $Spin(3,1)$ such that : $\mathbf{Ad}_{s_r}\varepsilon_0 = s_r\cdot\varepsilon_0\cdot s_r^{-1} = \varepsilon_0$ and it reads :

$$Spin\,(3) = \left\{ s_r = \epsilon\sqrt{1-\frac{1}{4}r^t r} + \upsilon\,(r,0)\,, r\in\mathbb{R}^3, r^t r\leq 4, \epsilon=\pm 1 \right\}$$

$Spin(3)$ is a compact group, with 2 connected components. The connected component of the identity is comprised of elements with $\epsilon = 1$ and can be assimilated to $SO\,(3)$ [5].

The elements of $Spin(3)$ are generated by vectors belonging to the subspace $F\,(\varepsilon_0)$ spanned by the vectors $(\varepsilon_i)_{i=1}^3$. They have a special physical meaning : they are the spatial rotations for an observer *with a velocity in the direction of* ε_0. In the tangent space $T_m M$ of the manifold M all rotations (given by $Spin\,(3,1)$) are on the same footing. But, because of our assumptions about the motion of observers (along time like lines), any observer introduces a breakdown of symmetry : some rotations are privileged. Indeed the spatial rotations are special, in that they are the ones for which the axis belongs to the physical space.

For a given ε_0, and then set $Spin\,(3)$, one can define the **quotient space** $SW = Spin\,(3,1)\,/Spin\,(3)$. This is not a group (because $Spin(3)$ is not a normal subgroup) but a 3 dimensional manifold, called a homogeneous space (Maths.22.4.3). It is characterized by the equivalence relation :

$\forall s,s'\in Spin\,(3,1) : s\sim s' \Leftrightarrow \exists s_r\in Spin(3) : s' = s\cdot s_r$

The projection : $\pi_w : Spin\,(3,1)\to SW$ is a submersion, its derivative $\pi'_w\,(s)$ is surjective. $Spin\,(3,1)$ is a principal fiber bundle $Spin\,(3,1)\,(SW, Spin\,(3)\,,\pi_w)$ and there is a smooth transitive left action of $Spin\,(3,1)$ on SW :

[4] The set of 3 dimensional vector subspaces of F with a definite positive (or negative) metric is a 3 dimensional smooth manifold, called a Stiefel manifold, isomorphic to the set of matrices $SO(4)/SO(1)\simeq SO\,(3)$.

[5] It is formally $SO\,(3)$ plus the scalar $\{1\}$

$\lambda : Spin(3,1) \times SW \to SW : \lambda(s, s_w) = \pi_w(s \cdot s_w)$ (Maths.1813)

$\forall s_w, s'_w, \exists s \in Spin(3,1) : s'_w = \lambda(s, s_w) = \pi_w(s \cdot s_w)$

This structure is very useful, because it enables us to write any element of the spin group as a product $s_w \cdot s_r$. Physically it means that we choose first a world line (represented by a vector ε_0) which provides $s_w \in SW$, then a rotation in the space represented by a rotation $s_r \in Spin(3)$. It works as follows.

The principal bundle structure of $Spin(3,1)$ means that there are trivializations : $\varphi : SW \times Spin(3) \to Spin(3,1)$ and one can prove (see annex) that, for a given vector ε_0, any element $s \in Spin(3,1)$ can be written uniquely (up to sign) : $s = s_w \cdot s_r$ with $s_w \in SW, s_r \in Spin(3)$:

$$\forall s \in Spin(3,1) : s = \epsilon(a_w + \upsilon(0,w)) \cdot \epsilon(a_r + \upsilon(r,0))$$

In each class of SW there are only two elements of $Spin(3,1)$ which can be written as : $s_w = a_w + \upsilon(0,w)$, and they have opposite sign : $\pm s_w$ belong to the same class of SW, they are specific representatives of the projection of s on the homogeneous space SW. The elements of $SW = Spin(3,1)/Spin(3)$ are coordinated by w, and the matrix $[K(w)]$ corresponds to a gauge transformation for an observer moving with a spatial speed $\overrightarrow{v}$ parallel to w, without spatial rotation. If we choose s_r in the connected component of the identity then $\epsilon a_r > 0$ and ϵ is fixed by the sign of $a_w : a_r a_w = a$, that is by the choice for w as the same direction as $\overrightarrow{v}$ or the opposite.

The decomposition depends on the choice of ε_0.

Decomposition of the Lie algebra

To each Clifford bundle $Cl(3)$ is associated a unique Lie algebra $T_1Spin(3)$ which is a subset of $Cl(3)$ and thus of $Cl(3,1)$.

In any orthonormal basis an element of $T_1Spin(3,1)$ reads :

$X = \upsilon(r,0) + \upsilon(0,w)$ and $\upsilon(r,0) \in T_1Spin(3), \upsilon(0,w) \in T_1SW$

The vectors r, w depends on the basis (they are components), however the elements $\upsilon(r,0), \upsilon(0,w) \in T_1Spin(3,1)$ depend only on the choice of ε_0 as we will see now.

For any given vector $\varepsilon_0 : \varepsilon_0 \cdot \varepsilon_0 = -1$ let be the linear map :

$\theta(\varepsilon_0) : T_1Spin(3,1) \to T_1Spin(3,1) : \theta(\varepsilon_0)(X) = \varepsilon_0 \cdot X \cdot \varepsilon_0$

If is easy to see that for any basis built with ε_0 :

$\forall a = 1,2,3 : \varepsilon_0 \cdot \overrightarrow{\kappa}_a \cdot \varepsilon_0 = -\overrightarrow{\kappa}_a$

$\forall a = 4,5,6 : \varepsilon_0 \cdot \overrightarrow{\kappa}_a \cdot \varepsilon_0 = \overrightarrow{\kappa}_a$

Thus $\theta(\varepsilon_0)\upsilon(r,w) = \upsilon(-r,w)$

$\theta(\varepsilon_0)$ has two eigen values ± 1 with the eigen spaces :

$L_0 = \{X \in T_1Spin(3,1) : \theta(\varepsilon_0)(X) = -X\} = \{\upsilon(r,0), r \in \mathbb{R}^3\}$

$P_0 = \{X \in T_1Spin(3,1) : \theta(\varepsilon_0)(X) = X\} = \{\upsilon(0,w), w \in \mathbb{R}^3\}$

$T_1Spin(3,1) = L_0 \oplus P_0$

Thus L_0, P_0 and the decomposition depend only on the choice of ε_0 and $L_0 = T_1Spin(3), P_0 \simeq T_1SW$.

$\theta(\varepsilon_0)$ commutes with the action of the elements of $Spin(3)$:

$\forall s_r \in Spin(3), X \in T_1Spin(3,1)$:

$\mathbf{Ad}_{s_r}\theta(\varepsilon_0)(X) = s_r \cdot \varepsilon_0 \cdot X \cdot \varepsilon_0 \cdot s_r^{-1} = \varepsilon_0 \cdot s_r \cdot X \cdot s_r^{-1} \cdot \varepsilon_0 = \theta(\varepsilon_0)(\mathbf{Ad}_{s_r}(X))$

with $\mathbf{Ad}_{sr}\varepsilon_0 = s_r \cdot \varepsilon_0 \cdot s_r^{-1} = \varepsilon_0$

The vector subspaces L_0, P_0 are globally invariant by $Spin(3)$: in a change of basis with $s_r \in Spin(3)$:

$\mathbf{Ad}_{s_r} = \begin{bmatrix} \left[1 + a_r j(r) + \frac{1}{2} j(r) j(r)\right] & 0 \\ 0 & \left[1 + a_r j(r) + \frac{1}{2} j(r) j(r)\right] \end{bmatrix}$

$X = \upsilon(x, 0) \to \upsilon\left(\left[1 + a_r j(r) + \frac{1}{2} j(r) j(r)\right] x, 0\right)$

$X = \upsilon(0, y) \to \upsilon\left(0, \left[1 + a_r j(r) + \frac{1}{2} j(r) j(r)\right] y\right)$

L_0 is a Lie subalgebra, $[L_0, L_0] \subset L_0, [L_0, P_0] \subset P_0, [P_0, P_0] \subset L_0$

This is a Cartan decomposition of $T_1 Spin(3,1)$ (Maths.1742). It depends on the choice of ε_0 but not of the choice of $(\varepsilon_i)_{i=1}^3$.

Let us define the projections :

$\pi_L(\varepsilon_0) : T_1 Spin(3,1) \to L_0 ::$

$\pi_L(\varepsilon_0)(X) = \frac{1}{2}(X - \theta(\varepsilon_0)(X)) = \frac{1}{2}(X - \varepsilon_0 \cdot X \cdot \varepsilon_0) = \upsilon(r, 0)$

$\pi_P(\varepsilon_0) : T_1 Spin(3,1) \to P_0 ::$

$\pi_L(\varepsilon_0)(X) = \frac{1}{2}(X + \theta(\varepsilon_0)(X)) = \frac{1}{2}(X + \varepsilon_0 \cdot X \cdot \varepsilon_0) = \upsilon(0, w)$

$X = \pi_L(\varepsilon_0)(X) + \pi_P(\varepsilon_0)(X)$

and the projections commute with the action of the elements of $Spin(3)$:

$\forall s_r \in Spin(3), X \in T_1 Spin(3,1)$:

$\pi_L(\varepsilon_0)(\mathbf{Ad}_{s_r}(X)) = \mathbf{Ad}_{s_r}(\pi_L(\varepsilon_0)(X))$

$\pi_P(\varepsilon_0)(\mathbf{Ad}_{s_r}(X)) = \mathbf{Ad}_{s_r}(\pi_P(\varepsilon_0)(X))$

The scalar product on the Clifford algebra reads in $T_1 Spin(3,1)$

$\langle \upsilon(r, w), \upsilon(r', w') \rangle_{Cl} = \frac{1}{4}(r^t r' - w^t w')$

and then it is definite positive on $T_1 Spin(3) = L_0$ and definite negative on P_0.

$\theta(\varepsilon_0)$ preserves the scalar product and L_0, P_0 are orthogonal, thus :

$\langle X, X \rangle_{Cl} = \langle \pi_L(X), \pi_L(X) \rangle_{Cl} + \langle \pi_P(X), \pi_P(X) \rangle_{Cl}$

Let us define the map :

$\|X\| : T_1 Spin(3,1) \to \mathbb{R}_+ : \|X\| = \sqrt{\langle \pi_L(X), \pi_L(X) \rangle_{Cl} - \langle \pi_P(X), \pi_P(X) \rangle_{Cl}}$

This is a norm on $T_1 Spin(3,1)$:

$\|X\| = 0 \Leftrightarrow \pi_L(X) = \pi_P(X) = X = 0$

$\|\lambda X\| = |\lambda| \|X\|$

$\|X + X'\|^2 = \langle \pi_L(X + X'), \pi_L(X + X') \rangle_{Cl} - \langle \pi_P(X + X'), \pi_P(X + X') \rangle_{Cl}$

$\langle \pi_L(X + X'), \pi_L(X + X') \rangle_{Cl} \leq \langle \pi_L(X), \pi_L(X) \rangle_{Cl} + \langle \pi_L(X'), \pi_L(X') \rangle_{Cl}$

$- \langle \pi_P(X + X'), \pi_P(X + X') \rangle_{Cl} \leq - \langle \pi_P(X), \pi_P(X) \rangle_{Cl} - \langle \pi_P(X'), \pi_P(X) \rangle_{Cl}$

$\Rightarrow$

$\|X + X'\|^2 \leq \|X\| + \|X'\|$

It reads :

$$\|\upsilon(r, w)\| = \frac{1}{2}\sqrt{r^t r + w^t w} = \frac{1}{2}\sqrt{\langle \pi_L(X), \pi_L(X) \rangle_{Cl} - \langle \pi_P(X), \pi_P(X) \rangle_{Cl}} \tag{4.14}$$

It depends only on the choice of ε_0.

A change of basis changes the decomposition only if it changes ε_0, that is if it is done by some $s_w = a_w + \upsilon(0, w) \in SW$. Then the elements of F or $T_1 Spin(3,1)$ do not change, but their components change. The value of the norm depends on the choice of ε_0 but, as there is always a vector such as ε_0 in any orthonormal basis, its existence is assured.

Remarks : In any Lie algebra there is a bilinear symmetric form B called the Killing form B, which does not depend on a basis and is invariant by Ad. In any orthonormal basis, defined as above, it has on $T_1 Spin(3,1)$ the same expression as in $so(3,1)$ (Maths.1669) :

$B(\upsilon(r, w), \upsilon(r', w')) = 4(r^t r' - w^t w') = 16 \langle \upsilon(r, w), \upsilon(r', w') \rangle_{Cl}$

4.2.3 Change of basis in F

Expression of the action $\mathbf{Ad}_s$ on vectors

The action of $Spin(3,1)$ on vectors of F is :

$v=\sum_{i=0}^{3} v^i\varepsilon_i \to \widetilde{v}=\mathbf{Ad}_s v=\sum_{i=0}^{3} v^i s\cdot\varepsilon_i\cdot s^{-1}=\sum_{i=0}^{3}\widetilde{v}^i\varepsilon_i$

$\widetilde{v}^i=\sum_{j=0}^{3}[h(s)]^i_j v^j$

With the expression of the elements of the Spin group $s=a+\upsilon(r,w)+b\varepsilon_5$ it is easy to express the matrix $[h(s)]$

$$[h(s)]=\begin{bmatrix} a^2+b^2+\frac{1}{4}(r^tr+w^tw) & aw^t-br^t+\frac{1}{2}w^tj(r) \\ aw-br+\frac{1}{2}j(r)w & a^2+b^2+\frac{1}{4}(r^tr+w^tw)+aj(r)+bj(w)+\frac{1}{2}(j(r)j(r)+j(w)j(w)) \end{bmatrix}$$

For a product : $\mathbf{Ad}_s\circ\mathbf{Ad}_{s'}=\mathbf{Ad}_{s\cdot s'}\to[h(s.s')]=[h(s)][h(s')]$

Then If $s=s_w\cdot s_r:[h(s)]=[h(s_w)][h(s_r)]$

If $s=a_w+\upsilon(0,w)$

$$[h(s)]=\begin{bmatrix} 2a_w^2-1 & a_ww^t \\ a_ww & 2a_w^2-1+\frac{1}{2}j(w)j(w)\end{bmatrix}$$

If $s=a_r+\upsilon(r,0)$

$$[h(s)]=\begin{bmatrix} 1 & 0 \\ 0 & 1+a_rj(r)+\frac{1}{2}j(r)j(r)\end{bmatrix}$$

$[C(r)]=1+a_rj(r)+\frac{1}{2}j(r)j(r)\in SO(3)$

Both matrices belong to $SO(3,1):[h(s)]^t[\eta][h(s)]=[\eta]$. Notice that $[h(s)]\neq\begin{bmatrix}1&0\\0&j(r)\end{bmatrix}$: this is not a usual rotation in the 3 dimensional space (this is a matrix of the group and not the Lie algebra). However its eigen vector for the eigen value 1 is $\begin{bmatrix}0\\r\end{bmatrix}$: the axis of rotation is the space vector r.

Expression of the Action $\mathbf{Ad}_s$ on the Lie algebra

The action of $Spin(3,1)$ is :

$Z=\sum_{a=1}^{6}Z_a\overrightarrow{\kappa}_a\to\widetilde{Z}=\sum_{a=1}^{6}Z_a\mathbf{Ad}_s(\overrightarrow{\kappa}_a)=\sum_{a=1}^{6}Z_as\cdot(\overrightarrow{\kappa}_a)\cdot s^{-1}=\sum_{a=1}^{6}Z_a\widetilde{\overrightarrow{\kappa}_a}=\sum_{a=1}^{6}\widetilde{Z}_a\overrightarrow{\kappa}_a$

With :

$Z=\upsilon(X,Y)\to\widetilde{Z}=\upsilon\left(\widetilde{X},\widetilde{Y}\right)$

$$\begin{bmatrix}\widetilde{X}\\ \widetilde{Y}\end{bmatrix}=[\mathbf{Ad}_s]\begin{bmatrix}X\\Y\end{bmatrix}$$

where $[\mathbf{Ad}_s]$ is a 6×6 matrix (see Annex) with $s=a+\upsilon(r,w)+b\varepsilon_5$:

$[\mathbf{Ad}_{s\cdot s'}]=[\mathbf{Ad}_s][\mathbf{Ad}_{s'}]$

With $s_w=a_w+\upsilon(0,w)$

$$[\mathbf{Ad}_s]=\begin{bmatrix}\left[1-\frac{1}{2}j(w)j(w)\right] & -[a_wj(w)] \\ [a_wj(w)] & \left[1-\frac{1}{2}j(w)j(w)\right]\end{bmatrix}=\begin{bmatrix}A&-B\\B&A\end{bmatrix}$$

and the identities :

$A=A^t,B^t=-B,AB=BA$

$A^2+B^2=I$

$$[\mathbf{Ad}_{s_w}]^{-1}=\left[\mathbf{Ad}_{s_w^{-1}}\right]=\begin{bmatrix}A&B\\-B&A\end{bmatrix}$$

With $s_r=a_r+\upsilon(r,0)$

$[\mathbf{Ad}_s] = \begin{bmatrix} \left[1+a_r j(r)+\frac{1}{2}j(r)j(r)\right] & 0 \\ 0 & \left[1+a_r j(r)+\frac{1}{2}j(r)j(r)\right] \end{bmatrix} = \begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix}$

and the identities :

$CC^t = C^tC = I_3$

$[\mathbf{Ad}_{s_r}]^{-1} = \left[\mathbf{Ad}_{s_r^{-1}}\right] = \begin{bmatrix} C^t & 0 \\ 0 & C^t \end{bmatrix}$

Change of basis in F

A change of orthonormal basis of F can be expressed by an action of the Spin group :

$s = a + \upsilon(r,w) + b\varepsilon_5 \in Spin(3,1)$: $\varepsilon_i \to \widetilde{\varepsilon}_i = \mathbf{Ad}_{s^{-1}}\varepsilon_i$

$\widetilde{\varepsilon}_i = \sum_{j=0}^3 \left[h\left(s^{-1}\right)\right]_i^j \varepsilon_j$

Then the vectors of F and $T_1Spin(3,1)$ stay the same, but their components change according to the inverse of the operations see above (as it is usual in any vector space).

$v = \sum_{i=0}^3 v^i\varepsilon_i = \sum_{i=0}^3 \widetilde{v}^i\widetilde{\varepsilon}_i = \sum_{i=0}^3 \widetilde{v}^i\mathbf{Ad}_s\varepsilon_i = \sum_{i,j=0}^3 \widetilde{v}^i \left[h(s^{-1})\right]_i^j \varepsilon_j$

$\widetilde{v}^i = \sum_{j=0}^3 [h(s)]_j^i v^j$

$Z = \sum_{a=1}^6 Z_a \overrightarrow{\kappa}_a = \sum_{a=1}^6 \widetilde{Z}_a \widetilde{\overrightarrow{\kappa}_a} = \sum_{a=1}^6 \widetilde{Z}_a \mathbf{Ad}_{s^{-1}}\left(\overrightarrow{\kappa}_a\right)$

$\upsilon(X,Y) = \widetilde{\upsilon}\left(\widetilde{X},\widetilde{Y}\right)$

$\begin{bmatrix} \widetilde{X} \\ \widetilde{Y} \end{bmatrix} = [\mathbf{Ad}_s] \begin{bmatrix} X \\ Y \end{bmatrix}$

In the following we will use fiber bundles where the change of gauge results in

$\varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i) \to \widetilde{\varepsilon}_i(m) = \mathbf{Ad}_{\chi(m)^{-1}}\varepsilon_i(m)$

Then the rules must be implemented with $\chi(m)^{-1}$. See the formulas in Annex.

4.2.4 Derivatives

Derivatives of translations

The translations on $Spin(3,1)$ are :

$L_g h = g \cdot h, R_g h = h \cdot g$

and their derivatives :

$L'_g h : T_h Spin(3,1) \to T_{g\cdot h}Spin(3,1) :: L'_g h(X_h) = g \cdot X_h$

$R'_g h : T_h Spin(3,1) \to T_{h\cdot g}Spin(3,1) :: R'_g h(X_h) = X_h \cdot g$

Their inverse are, as in any Lie groups :

$\left(L'_g h\right)^{-1} = L'_{g^{-1}}(g\cdot h)\,;\left(R'_g h\right)^{-1} = R'_{g^{-1}}(h\cdot g)$

$T_g Spin(3,1) \subset Cl(3,1)$ and there are two linear maps :

$L'_{g^{-1}}g : T_g Spin(3,1) \to T_1 Spin(3,1) :: L'_{g^{-1}}g(Z_g) = g^{-1}\cdot Z_g$

$R'_{g^{-1}}g : T_g Spin(3,1) \to T_1 Spin(3,1) :: R'_{g^{-1}}g(Z_g) = Z_g \cdot g^{-1}$

And the usual adjoint map:of Lie groups :

$Ad_g : T_1 Spin(3,1) \to T_1 Spin(3,1) :: Ad_g = L'_g g^{-1} \circ R'_{g^{-1}}1 = R'_{g^{-1}}g \circ L'_g 1$

$Ad_g Z = L'_g g^{-1} \circ R'_{g^{-1}}1(Z) = L'_g g^{-1}\left(Z\cdot g^{-1}\right) = g\cdot Z\cdot g^{-1}$

So the usual adjoint map of Lie groups is the adjoint map of the Clifford algebra :

$Ad_g Z = \left(\mathbf{Ad}_g Z\right)'_{x=1} = \mathbf{Ad}_g Z$

Derivatives on the Spin group

Let $g : M \to Spin(3,1)) :: g(m) = a(m) + \upsilon(r(m), w(m)) + b(m)\varepsilon_5$

$g'(m) : T_m M \to T_g Spin(3,1) ::$

$g'(m) u_m = a'(m) u_m + \upsilon(r'(m) u_m, w'(m) u_m) + b'(m) u_m \varepsilon_5$

where $u_m \in T_m M, a'(m) u_m, b'(m) u_m \in \mathbb{R}, r'(m) u_m, w'(m) u_m \in \mathbb{R}^3$

Thus :

$g^{-1} \cdot (a'(m) u_m + \upsilon(r'(m) u_m, w'(m) u_m) + b'(m) u_m \varepsilon_5) \in T_1 Spin(3,1)$

$(a'(m) u_m + \upsilon(r'(m) u_m, w'(m) u_m) + b'(m) u_m \varepsilon_5) \cdot g^{-1} \in T_1 Spin(3,1)$

$L'_{g^{-1}} g(g'(m)) = g^{-1} \cdot g'$

$= (a(m) - \upsilon(r(m), w(m)) + b(m)\varepsilon_5) \cdot (a'(m) + \upsilon(r'(m), w'(m)) + b'(m)\varepsilon_5)$

$= a" + \upsilon(r", w") + b"\varepsilon_5$

with :

$a" = aa' - bb' - \frac{1}{4}(w^t w' - r^t r')$

$= \frac{1}{2}(a^2 - b^2)' - \frac{1}{4}\frac{1}{2}((w^t w - r^t r))'$

$= \frac{1}{2}(a^2 - b^2 - \frac{1}{4}(w^t w - r^t r))' = 0$

$b" = (a'b + ab') + \frac{1}{4}(w^t r' + r^t w')$

$= (ab)' + \frac{1}{4}(w^t r)' = 0$

$r" = -a'r + ar' - bw' + b'w - \frac{1}{2}(j(r) r' - j(w) w')$

$w" = -a'w + aw' + br' - b'r - \frac{1}{2}(j(w) r' + j(r) w')$

$g^{-1} \cdot g' = \upsilon(-\frac{1}{2}(j(r) r' - j(w) w') - a'r + ar' + b'w - bw',$

$-\frac{1}{2}(j(w) r' + j(r) w') - a'w + aw' - b'r + br')$

Similarly :

$R'_{g^{-1}} g(g'(m) u_m) = g' \cdot g^{-1}$

$= (a'(m) + \upsilon(r'(m), w'(m)) + b'(m)\varepsilon_5) \cdot (a(m) - \upsilon(r(m), w(m)) + b(m)\varepsilon_5)$

$= a" + \upsilon(r", w") + b"\varepsilon_5$

with :

$a" = aa' - bb' - \frac{1}{4}(w^t w' - r^t r') = 0$

$b" = (a'b + ab') + \frac{1}{4}(w^t r' + r^t w') = 0$

$r" = \frac{1}{2}(j(r) r' - j(w) w') - a'r + ar' + b'w - bw'$

$w" = \frac{1}{2}(j(r) w' + j(w) r') - a'w + aw' - b'r + br'$

$g' \cdot g^{-1} = \upsilon\{\frac{1}{2}(j(r) r' - j(w) w') - a'r + ar' + b'w - bw',$

$\frac{1}{2}(j(r) w' + j(w) r') - a'w + aw' - b'r + br'\}$

With $\sigma_w = a_w + \upsilon(0, w)$

$\sigma_w^{-1} \cdot \partial_\alpha \sigma_w = \upsilon(\frac{1}{2} j(w) \partial_\alpha w, \frac{1}{4a_w}(-j(w) j(w) + 4) \partial_\alpha w)$

$\partial_\alpha \sigma_w \cdot \sigma_w^{-1} = \upsilon(-\frac{1}{2} j(w) \partial_\alpha w, \frac{1}{4a_w}(-j(w) j(w) + 4) \partial_\alpha w)$

With $\sigma_r = a_r + \upsilon(r, 0)$

$\sigma_r^{-1} \cdot \partial_\alpha \sigma_r = \upsilon(\left(\frac{1}{a_r} - \frac{1}{2} j(r) + \frac{1}{4a_r} j(r) j(r)\right) \partial_\alpha r, 0)$

$\partial_\alpha \sigma_r \cdot \sigma_r^{-1} = \upsilon(\left(\frac{1}{a_r} + \frac{1}{2} j(r) + \frac{1}{4a_r} j(r) j(r)\right) \partial_\alpha r, 0)$

Notice that $\partial_\alpha \sigma_r \cdot \sigma_r^{-1}, \partial_\alpha \sigma_r \cdot \sigma_r^{-1} \in T_1 Spin(3)$ but $\sigma_w^{-1} \cdot \partial_\alpha \sigma_w, \partial_\alpha \sigma_w \cdot \sigma_w^{-1} \notin T_1 SW$.

4.2.5 Representation of Clifford algebras

Complexification of real Clifford algebras

Any real vector space can be complexified, by extending the operations from real scalars to complex scalars (Maths.6.5.2) : as a set the vector space is enlarged by all vectors of the form $iu : F_{\mathbb{C}} = F \oplus iF$. The real scalar product is extended to a complex bilinear form $\langle\rangle_{\mathbb{C}}$, with the signature $(++++)$[6], any orthonormal basis $(\varepsilon_j)_{j=0}^3$ of F is an orthonormal basis of $F_{\mathbb{C}}$ with complex components. There is a complex Clifford algebra $Cl\left(F_{\mathbb{C}},\langle\rangle\right)$ which is the complexified of $Cl\left(F,\langle\rangle\right)$ and has the algebraic structure $Cl\left(\mathbb{C},4\right)$, the Clifford algebra on $\mathbb{C}^4$ with the bilinear *symmetric* form of signature $(++++)$. So for both signatures $Cl\left(3,1\right)$ and $Cl\left(1,3\right)$ have the same complexified structure $Cl\left(\mathbb{C},4\right)$. In $Cl\left(F_{\mathbb{C}},\langle\rangle\right)$ the product of vectors is :

$\forall u,v \in F_{\mathbb{C}} : u \odot v + v \odot u = 2\left\langle u,v\right\rangle_{\mathbb{C}}$

Any orthonormal basis of $Cl\left(3,1\right)$ or $Cl\left(1,3\right)$ is an orthonormal basis of $Cl\left(\mathbb{C},4\right)$ and : $\varepsilon_i \odot \varepsilon_j + \varepsilon_j \odot \varepsilon_i = 2\delta_{ij}$ and $\varepsilon_0 \odot \varepsilon_0 = +1$

$Cl\left(3,1\right)$ and $Cl\left(1,3\right)$ are real vector subspaces of $Cl\left(\mathbb{C},4\right)$.

There are real algebras morphisms (injective but not surjective) from the real Clifford algebras to $Cl\left(\mathbb{C},4\right)$.

With the signature (3 1) let us choose as above a vector $\varepsilon_0 \in F$ such that $\varepsilon_0 \cdot \varepsilon_0 = -1$.

Let us define the map :

$\widetilde{C} : \left(F,\langle\rangle\right) \to Cl\left(\mathbb{C},4\right) :: \widetilde{C}\left(u\right) = \left(u + \left\langle \varepsilon_0,u\right\rangle_F \varepsilon_0\right) - i\left\langle \varepsilon_0,u\right\rangle_F \varepsilon_0 = u + \left\langle \varepsilon_0,u\right\rangle_F \left(\varepsilon_0 - i\varepsilon_0\right)$

(this is just the map : $\widetilde{C}\left(\varepsilon_j\right) = \varepsilon_j, j = 1,2,3; \widetilde{C}\left(\varepsilon_0\right) = i\varepsilon_0$)

$\widetilde{C}\left(u\right) \odot \widetilde{C}\left(v\right) + \widetilde{C}\left(v\right) \odot \widetilde{C}\left(u\right)$

$= \left(u + \left\langle \varepsilon_0,u\right\rangle_F \left(\varepsilon_0 - i\varepsilon_0\right)\right) \odot \left(v + \left\langle \varepsilon_0,v\right\rangle_F \left(\varepsilon_0 - i\varepsilon_0\right)\right)$

$+ \left(v + \left\langle \varepsilon_0,v\right\rangle_F \left(\varepsilon_0 - i\varepsilon_0\right)\right) \odot \left(u + \left\langle \varepsilon_0,u\right\rangle_F \left(\varepsilon_0 - i\varepsilon_0\right)\right)$

$= 2\left\langle u,v\right\rangle_{\mathbb{C}} + 2\left\langle \varepsilon_0,v\right\rangle_F \left\langle u,\varepsilon_0 - i\varepsilon_0\right\rangle_{\mathbb{C}} + 2\left\langle \varepsilon_0,u\right\rangle_F \left\langle \varepsilon_0 - i\varepsilon_0,v\right\rangle_{\mathbb{C}}$

$+2\left\langle \varepsilon_0,u\right\rangle_F \left\langle \varepsilon_0,v\right\rangle_F \left\langle \varepsilon_0 - i\varepsilon_0,\varepsilon_0 - i\varepsilon_0\right\rangle_{\mathbb{C}}$

$= 2\left\langle u + \left\langle \varepsilon_0,u\right\rangle_F \left(\varepsilon_0 - i\varepsilon_0\right), v + \left\langle \varepsilon_0,v\right\rangle_F \left(\varepsilon_0 - i\varepsilon_0\right)\right\rangle_{\mathbb{C}}$

$= 2\left\langle \widetilde{C}\left(u\right),\widetilde{C}\left(v\right)\right\rangle_{\mathbb{C}}$

Thus, by the universal property of Clifford algebras, there is a unique real algebra morphism $C : Cl\left(3,1\right) \to Cl\left(\mathbb{C},4\right)$ such that $\widetilde{C} = C \circ \jmath$ where $\jmath$ is the canonical injection $\left(F,\langle\rangle\right) \to Cl\left(3,1\right)$ (Maths.494). We will denote for simplicity $\widetilde{C} = C$. The image $C\left(Cl\left(3,1\right)\right)$ is a real subalgebra of $Cl\left(\mathbb{C},4\right)$, which can be identified with $Cl\left(3,1\right)$ so it does not depend on the choice of ε_0 (but the map C depends on ε_0).

Similarly with $\widetilde{C}'\left(\varepsilon_j\right) = i\varepsilon_j, j = 1,2,3; \widetilde{C}'\left(\varepsilon_0\right) = \varepsilon_0$ we have a real algebra morphism $C' : Cl\left(1,3\right) \to Cl\left(\mathbb{C},4\right)$ and $C'\left(Cl\left(1,3\right)\right)$ is a real subalgebra of $Cl\left(\mathbb{C},4\right)$. Moreover $C'\left(\varepsilon_j\right) = -i\eta_{jj}C\left(\varepsilon_j\right)$ (η always correspond to the signature - + + +).

Algebraic and geometric representations

An *algebraic* representation of a Clifford algebra is a map γ which associates to each element w of the Clifford algebra a matrix $\left[\gamma\left(w\right)\right]$ and such that γ is a isomorphism of algebra : all the operations in the Clifford algebra (multiplication by a scalar, sum, Clifford product) are reproduced on the matrices. A representation is fully defined by the family of matrices $(\gamma_i)_{i=0}^3$ representing each vector $(\varepsilon_i)_{i=0}^3$ of an orthonormal basis. The choice of these matrices is not

[6] Actually the signature of a bilinear symmetric form is defined for real vector space, but the meaning will be clear for the reader. We will always work here with bilinear form and not hermitian form.

unique : the only condition is that $[\gamma_i][\gamma_j] + [\gamma_j][\gamma_i] = 2\eta_{ij}[I]$ and any family of matrices deduced by conjugation with a fixed matrix gives an equivalent algebraic representation. An element of the Clifford algebra is then represented by a linear combination of generators :

$\gamma(w) = \gamma\left(\sum_{\{i_1...i_r\}} a^{i_1...i_r}\varepsilon_{i_1}\cdot...\cdot\varepsilon_{i_r}\right) = \sum_{\{i_1...i_r\}} a^{i_1...i_r}\gamma_{i_1}...\gamma_{i_r}$

A *geometric* representation (E,γ) of a Clifford algebra is an isomorphism $\gamma : Cl \to L(E;E)$ in which $[\gamma(w)]$ is the matrix of an endomorphism of E, represented in some basis. From an algebraic representation one can deduce a geometric representation, and they are equivalent up to the choice of a basis.

We look for a geometric representation : the quantity S that we are looking for is represented, not by γ matrices, but by vectors S of the space E, which are called spinors. Higher orders spinors are tensorial products of vectors of E.

A Clifford algebra has, up to isomorphism, a unique faithful algebraic irreducible representation in an algebra of matrices (γ is a bijection). As can be expected the representations depend on the signature :

For $Cl(3,1)$ this is $\mathbb{R}(4)$ the 4×4 real matrices (the corresponding spinors are the Majorana spinors)

For $Cl(1,3)$ this is $H(2)$ the 2×2 matrices with quaternionic elements

In both cases an element of the Clifford algebra is characterized by $2^4 = 16$ real parameters.

The geometry of the universe is based upon real structures. Thus we should consider representations of $Cl(3,1)$ or $Cl(1,3)$, which raises the issue of the signature. However it happens, from experience, that the vector space E must be complex[7].

The irreducible representation of $Cl(\mathbb{C},4)$ is by 4×4 matrices on complex numbers which must meet the condition : $\gamma_j\gamma_k + \gamma_k\gamma_j = 2\delta_{jk}I_4$.

If (E,γ) is a complex representation of $Cl(\mathbb{C},4)$ then $(E,\gamma\circ C)$ is a real geometric representation of $Cl(3,1)$ on the complex vector space E : the map $\gamma\circ C : Cl(3,1) \to L(E;E)$ is a real morphism of algebras, and the maps $\gamma\circ C(w)$ are complex linear. The matrices of the real representation are $i\gamma_0$, $\gamma_j, j = 1,2,3, i\gamma_0$. Similarly $(E,\gamma\circ C')$ is a real geometric representation of $Cl(1,3)$ with matrices $\gamma_0, i\gamma_j, j = 1,2,3$.

Using this trick we see that we are fortunate, in that we have the same representation (E,γ) for both signatures, and a complex vector space E. Moreover it is easy to specify the representation through additional features of E (such as chirality as we will see). A spinor has 8 real components (vs 16 real components for elements of the real Clifford algebras) thus a spinor carries more information than a simple vector of $\mathbb{R}^4$ and this solves part of the issue of the number of parameters needed to represent the motion (both translation and rotation).

Chirality

Any Clifford algebra Cl is the direct sum of one Clifford subalgebra Cl_0 comprised of elements which are the sum of products of an even number of vectors, and a vector subspace Cl_1 comprised of elements which are the sum of products of an odd number of vectors. Moreover some Clifford algebras present a specific feature : they are the direct sum of two subalgebras which can be seen as algebras of left handed and right handed elements. This property depends on the existence of an element ϖ such that $\varpi\cdot\varpi = 1$. This element exists in any complex algebra, but not in $Cl(1,3), Cl(3,1)$. As chirality is a defining feature of particles, this is an additional argument for using $Cl(\mathbb{C},4)$.

[7]This is necessary to represent the electromagnetic fiefd.

In $Cl(\mathbb{C},4)$ the special element is : $\varpi = \pm\varepsilon_0 \odot \varepsilon_1 \odot \varepsilon_2 \odot \varepsilon_3 \in Spin(\mathbb{C},4)$. Thus there is a choice and we will use : $\varpi = \varepsilon_5 = \varepsilon_0 \odot \varepsilon_1 \odot \varepsilon_2 \odot \varepsilon_3$.

The Clifford algebra splits in two subalgebras :
$Cl(\mathbb{C},4) = Cl^R(\mathbb{C},4) \oplus Cl^L(\mathbb{C},4)$:
$Cl^R(\mathbb{C},4) = \{w \in Cl(\mathbb{C},4) : \varepsilon_5 \odot w = w\}$,
$Cl^L(\mathbb{C},4) = \{w \in Cl(\mathbb{C},4) : \varepsilon_5 \odot w = -w\}$
and any element of $Cl(\mathbb{C},4)$ can be uniquely written as : $w = w_R + w_L$
The projections from $Cl(\mathbb{C},4)$ on each subalgebra are the maps
$p_R = \frac{1}{2}(1+\varepsilon_5), p_L = \frac{1}{2}(1-\varepsilon_5)$:
$w_R = p_R \odot w, w_L = p_L \odot w$
$p_R \odot p_L = p_L \odot p_R = 0, p_R^2 = p_R, p_L^2 = p_L, p_R + p_L = 1$
We have similarly : $E = E^R \oplus E^L$ with
$E^R = \gamma_R(E), E^L = \gamma_L(E), \gamma_R = \gamma(p_R), \gamma_L = \gamma(p_L) \Rightarrow \gamma(\varepsilon_5) = \gamma_R - \gamma_L$
$u \in E : u = u_R + u_L$:
$u_R = \gamma_R(u) = \frac{1}{2}(u + \gamma(\varepsilon_5)u)$;
$u_L = \gamma_L(u) = \frac{1}{2}(u - \gamma(\varepsilon_5)u)$
For any homogeneous element $w = v_1 \odot v_2 ... \odot v_k, v_j \in \mathbb{C}^4$ we have $\varepsilon_5 \odot w = (-1)^k w \odot \varepsilon_5$
$\forall w \in Cl(\mathbb{C},4), u \in E$:
$\gamma_R(\gamma(w)u_R) = \frac{1}{2}\left(1+(-1)^k\right)\gamma(w)u_R$
k even : $\gamma_R(\gamma(w)u_R) = \gamma(w)u_R$
k odd : $\gamma_R(\gamma(w)u_R) = 0$
For k even : $\gamma(w)$ preserves both E^R, E^L (as vector subspaces)
For k odd : $\gamma(w)$ exchanges E^R, E^L
In particular the elements of the images $C(Spin(3,1))$ and $C'(Spin(1,3))$ by γ preserve both E^R, E^L. So we have reducible representations of these groups.

The choice of the representation γ

An algebraic representation is defined by the choice of its generators γ_i, and any set of generators conjugate by a fixed matrix gives an equivalent representation. We can specify the generators by the choice of a basis $(e_i)_{i=1}^4$ of E . The previous result leads to a natural choice : take $(e_i)_{i=1}^2$ as basis of E^R and $(e_i)_{i=3}^4$ as basis of E^L, then :

$$\gamma_R = \begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix}, \gamma_L = \begin{bmatrix} 0 & 0 \\ 0 & I_2 \end{bmatrix}, \gamma_5 = \gamma(\varepsilon_5) = \gamma_R - \gamma_L = \begin{bmatrix} I_2 & 0 \\ 0 & -I_2 \end{bmatrix}$$

Denote : $\gamma_j = \begin{bmatrix} A_j & B_j \\ C_j & D_j \end{bmatrix}$ with four 2×2 complex matrices.

ε_5 belongs to the Spin group $Spin(\mathbb{C},4)$, commutes with any element of $Cl_0(\mathbb{C},4)$ and anti-commutes with any vector, thus $\gamma_5\gamma_j = -\gamma_j\gamma_5$ which imposes the condition :

$$\begin{bmatrix} A_j & -B_j \\ C_j & -D_j \end{bmatrix} = -\begin{bmatrix} A_j & B_j \\ -C_j & -D_j \end{bmatrix} \Rightarrow \gamma_j = \begin{bmatrix} 0 & B_j \\ C_j & 0 \end{bmatrix}$$

The defining relations : $\gamma_j\gamma_k + \gamma_k\gamma_j = 2\delta_{jk}I_4$ lead to :

$$\begin{bmatrix} B_jC_k + B_kC_j & 0 \\ 0 & C_jB_k + C_kB_j \end{bmatrix} = 2\delta_{jk}I_4$$

$j \neq k : B_jC_k + B_kC_j = C_jB_k + C_kB_j = 0$
$j = k : B_jC_j = C_jB_j = I_2 \Leftrightarrow C_j = B_j^{-1}$

thus $(\gamma_i)_{i=0}^3$ is fully defined by a set $(B_i)_{i=0}^3$ of 2×2 complex matrices
$\gamma_j = \begin{bmatrix} 0 & B_j \\ B_j^{-1} & 0 \end{bmatrix}$
meeting : $j \neq k : B_j B_k^{-1} + B_k B_j^{-1} = B_j^{-1} B_k + B_k^{-1} B_j = 0$
which reads :
$B_j B_k^{-1} = -\left(B_j B_k^{-1}\right)^{-1} \Leftrightarrow \left(B_j B_k^{-1}\right)^2 = -I_2$
$B_j^{-1} B_k = -\left(B_j^{-1} B_k\right)^{-1} \Leftrightarrow \left(B_j^{-1} B_k\right)^2 = -I_2$

Let us define : $k = 1,2,3 : M_k = -iB_k B_0^{-1}$
The matrices $(M_k)_{k=1}^3$ are such that :
$M_k^2 = -\left(B_j B_0^{-1}\right)^2 = -I_2$
$M_j M_k + M_k M_j = -B_j B_0^{-1} B_k B_0^{-1} - B_k B_0^{-1} B_j B_0^{-1}$
$= -\left(-B_j B_k^{-1} B_0 - B_k B_j^{-1} B_0\right) B_0^{-1}$
$= B_j B_k^{-1} + B_k B_j^{-1} = 0$
that is $k = 1,2,3 : M_j M_k + M_k M_j = 2\delta_{jk} I_2$
Moreover : $\gamma_5 = \gamma_0\gamma_1\gamma_2\gamma_3 \Rightarrow$
$B_0 B_1^{-1} B_2 B_3^{-1} = I_2$
$B_0^{-1} B_1 B_2^{-1} B_3 = -I_2$
with $B_k = iM_k B_0, B_k^{-1} = -iB_0^{-1} M_k^{-1}$
$B_0 \left(-iB_0^{-1} M_1^{-1}\right) \left(iM_2 B_0\right) \left(-iB_0^{-1} M_3^{-1}\right) = I_2 = -iM_1^{-1} M_2 M_3^{-1}$
$B_0^{-1} \left(iM_1 B_0\right) \left(-iB_0^{-1} M_2^{-1}\right) \left(iM_3 B_0\right) = -I_2 = iB_0^{-1} M_1 M_2^{-1} M_3 B_0$
which reads :
$iM_2 = -M_1 M_3 = M_3 M_1$
$-M_1^{-1} M_3^{-1} = iM_2^{-1} \Leftrightarrow iM_2 = M_3 M_1$
$M_2 M_3 + M_3 M_2 = 0 = iM_1 M_3 M_3 + M_3 M_2 \Leftrightarrow iM_1 = -M_3 M_2 = M_2 M_3$
$M_1 M_2 + M_2 M_1 = 0 = iM_3 M_2 M_2 + M_2 M_1 \Rightarrow iM_3 = -M_2 M_1 = M_1 M_2$
The set of 3 matrices $(M_k)_{k=1}^3$ has the multiplication table :
$\begin{bmatrix} 1\backslash 2 & M_1 & M_2 & M_3 \\ M_1 & I & iM_3 & -iM_2 \\ M_2 & -iM_3 & I & iM_1 \\ M_3 & iM_2 & -iM_1 & I \end{bmatrix}$
which is the same as the set of Pauli's matrices :

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}; \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}; \sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{4.15}$$

$$\sigma_i^2 = \sigma_0; \text{ For } j \neq k : \sigma_j \sigma_k = \epsilon\left(j,k,l\right) i\sigma_l \tag{4.16}$$

Notation 62 $\epsilon\left(j,k,l\right)$ = *the signature of the permutation of the three different integers i,j,k, 0 if two integers are equal*

There is still some freedom in the choice of the γ_i matrices by the choice of B_0 and the simplest is : $B_0 = -iI_2 \Rightarrow B_k = \sigma_k$

Moreover, because scalars belong to Clifford algebras, one must have the identity matrix I_4 and $\gamma\left(z\right) = zI_4$

Thus :

$$\gamma_0 = \begin{bmatrix} 0 & -i\sigma_0 \\ i\sigma_0 & 0 \end{bmatrix}; \gamma_1 = \begin{bmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{bmatrix}; \gamma_2 = \begin{bmatrix} 0 & \sigma_2 \\ \sigma_2 & 0 \end{bmatrix}; \gamma_3 = \begin{bmatrix} 0 & \sigma_3 \\ \sigma_3 & 0 \end{bmatrix}; \tag{4.17}$$

The matrices γ_j are then unitary and Hermitian :

$$\gamma_j = \gamma_j^* = \gamma_j^{-1} \tag{4.18}$$

which is extremely convenient.

We will use the following (see the annex for more formulas) :

Notation 63 *$j = 1,2,3 : \widetilde{\gamma}_j = \begin{bmatrix} \sigma_j & 0 \\ 0 & \sigma_j \end{bmatrix}$*

$j \neq k, l = 1, 2, 3 : \gamma_j \gamma_k = -\gamma_k \gamma_j = i\epsilon(j, k, l)\, \widetilde{\gamma}_l$

$j = 1, 2, 3 : \gamma_j \gamma_0 = -\gamma_0 \gamma_j = i \begin{bmatrix} \sigma_j & 0 \\ 0 & -\sigma_j \end{bmatrix} = i\gamma_5 \widetilde{\gamma}_j$

Notice that the choice of the matrices is done in $Cl(\mathbb{C}, 4)$, so it is independent of the choice of signature. However we have the representations of the real algebras by the matrices $\gamma C(\varepsilon_j)$ and $\gamma C'(\varepsilon_j)$

$$\begin{bmatrix} Cl(3,1) : \gamma C(\varepsilon_j) = \gamma_j, j = 1,2,3; \gamma C(\varepsilon_0) = i\gamma_0; \gamma C(\varepsilon_5) = i\gamma_5 \\ Cl(1,3) : \gamma C'(\varepsilon_j) = i\gamma_j, j = 1,2,3; \gamma C'(\varepsilon_0) = \gamma_0; \gamma C'(\varepsilon_5) = \gamma_5 \end{bmatrix} \tag{4.19}$$

The representation that we have chosen here is not unique and others, equivalent, would hold. However from my point of view this is the most convenient because of the nice properties of the γ matrices. The choice of $\varpi = -\varepsilon_5 = -\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$ would have lead to take $\widetilde{\gamma}_j = -\gamma_j$. In the Standard Model we have a representation of $Cl(1,3)$ by the matrices : $\widetilde{\gamma}_0 = i\gamma_0, \widetilde{\gamma}_j = \gamma_j, j = 1, 2, 3$ and $\widetilde{\gamma}_5 = -i\widetilde{\gamma}_0\widetilde{\gamma}_1\widetilde{\gamma}_2\widetilde{\gamma}_3$

Expression of the matrices for the Lie algebra and the Spin groups

The

Notation 64 $\sum_{a=1}^{3} z_a \sigma_a = \sigma(z)$ *with* $z \in \mathbb{C}^3$

is very convenient as we have :

$(\sigma(z))^* = \sigma(\overline{z})$

$\sigma(z)\, \sigma(z') = \sigma(j(z)\, z') + z^t z' \sigma_0$

The matrices $\gamma C(\upsilon(r, w)), \gamma C'(\upsilon(r, w))$ are of constant use.

In Cl(3,1) :

$$\gamma C(\upsilon(r, w)) = -i\frac{1}{2}\sum_{a=1}^{3} (w^a \gamma_a \gamma_0 + r^a \widetilde{\gamma}_a) = \frac{1}{2}\begin{bmatrix} \sigma(w - ir) & 0 \\ 0 & -\sigma(w + ir) \end{bmatrix} \tag{4.20}$$

In Cl(1,3) :

$$\gamma C'(v(r,w)) = -i\frac{1}{2}\sum_{a=1}^{3}\left(w^a\gamma_a\gamma_0 - r^a\widetilde{\gamma}_a\right) = \frac{1}{2}\begin{bmatrix}\sigma(w+ir) & 0\\ 0 & -\sigma(w-ir)\end{bmatrix} \quad (4.21)$$

so one goes from one signature to the other by changing the sign of r.
The 2×2 matrices $\frac{1}{2}\sum_{a=1}^{3}(w_a - ir_a)\sigma_a$ and $\frac{1}{2}\sum_{a=1}^{3}(w_a + ir_a)\sigma_a$ belong to $SU(2)$
The elements of the spin groups are represented by the matrices :

In Cl(3,1) :

$$\gamma C(a + v(r,w) + b\varepsilon_5) = aI - i\tfrac{1}{2}\textstyle\sum_{a=1}^{3}\left(w_a\gamma_a\gamma_0 + r_a\widetilde{\gamma}_a\right) + b\gamma_5 \quad (4.22)$$

In Cl(1,3) :

$$\gamma C'(a + v(r,w) + b\varepsilon_5) = aI - i\tfrac{1}{2}\textstyle\sum_{a=1}^{3}\left(w^a\gamma_a\gamma_0 - r^a\widetilde{\gamma}_a\right) + b\gamma_5 \quad (4.23)$$

4.2.6 Scalar product of Spinors

We need a scalar product on E, preserved by a gauge transformation, that is by both $Spin(3,1)$ and $Spin(1,3)$.

Theorem 65 *The only scalar products on E, preserved by* $\{\gamma C(\sigma), \sigma \in Spin(3,1)\}$ *are* $G = \begin{bmatrix} 0 & k\sigma_0\\ \overline{k}\sigma_0 & 0\end{bmatrix}$ *with* $k \in \mathbb{C}$

Proof. It is represented in the basis of E by a 4×4 Hermitian matrix G such that : $G = G^*$
$\forall s \in Spin(3,1) : [\gamma \circ C(s)]^* G [\gamma \circ C(s)] = G$
or $\forall s \in Spin(1,3) : [\gamma \circ C'(s)]^* G [\gamma \circ C'(s)] =$
$[\gamma \circ C(s)]^* G = G[\gamma \circ C(s)]^{-1} = G\left[\gamma \circ C\left(s^{-1}\right)\right]$
$\gamma C(s) = \begin{bmatrix}(a+ib)\sigma_0 + \frac{1}{2}\sigma(w-ir) & 0\\ 0 & (a-ib)\sigma_0 - \frac{1}{2}\sigma(w+ir)\end{bmatrix}$
$\gamma C(s)^* = \begin{bmatrix}(a-ib)\sigma_0 + \frac{1}{2}\sigma(w+ir) & 0\\ 0 & (a+ib)\sigma_0 - \frac{1}{2}\sigma(w-ir)\end{bmatrix}$
$G = \begin{bmatrix} A & B\\ B^* & C\end{bmatrix}$, with $A = A^*, C = C^*$
$[\gamma \circ C(s)]^* G$
$= \begin{bmatrix}(a-ib)A + \frac{1}{2}\sigma(w+ir)A & (a-ib)B + \frac{1}{2}\sigma(w+ir)B\\ (a+ib)B^* - \frac{1}{2}\sigma(w-ir)B^* & (a+ib)C - \frac{1}{2}\sigma(w-ir)C\end{bmatrix}$
$G\left[\gamma \circ C\left(s^{-1}\right)\right]$
$= \begin{bmatrix}(a+ib)A - \frac{1}{2}A\sigma(w-ir) & (a-ib)B + \frac{1}{2}B\sigma(w+ir)\\ (a+ib)B^* - \frac{1}{2}B^*\sigma(w-ir) & (a-ib)C + \frac{1}{2}C\sigma(w+ir)\end{bmatrix}$
$(a-ib)A + \frac{1}{2}\sigma(w+ir)A = (a+ib)A - \frac{1}{2}A\sigma(w-ir)$
$(a-ib)B + \frac{1}{2}\sigma(w+ir)B = (a-ib)B + \frac{1}{2}B\sigma(w+ir)$
$(a+ib)B^* - \frac{1}{2}\sigma(w-ir)B^* = (a+ib)B^* - \frac{1}{2}B^*\sigma(w-ir)$
$(a+ib)C - \frac{1}{2}\sigma(w-ir)C = (a-ib)C + \frac{1}{2}C\sigma(w+ir)$
$\sigma(w+ir)B = B\sigma(w+ir)$
$\sigma(w-ir)B^* = B^*\sigma(w-ir)\sigma$

$2ibA$
$= \frac{1}{2}\sigma\left(w+ir\right)A + A\sigma\left(w-ir\right)$
$= \frac{1}{2}\sum_{a=1}^{3} w_a\left(\sigma_a A + A\sigma_a\right) + ir_a\left(\left(\sigma_a A - A\sigma_a\right)\right)$
$2ibC$
$= \frac{1}{2}C\sigma\left(w+ir\right)C + \sigma\left(w-ir\right)C$
$= \frac{1}{2}\sum_{a=1}^{3} w_a\left(\sigma_a C + C\sigma_a\right) + ir_a\left(\left(C\sigma_a - \sigma_a C\right)\right)$
By taking the adjoint on the two last equations :
$-2ibA$
$= \frac{1}{2}\sum_{a=1}^{3} w_a\left(A\sigma_a + \sigma_a A\right) - ir_a\left(\left(A\sigma_a - \sigma_a A\right)\right)$
$= -\frac{1}{2}\sum_{a=1}^{3} w_a\left(\sigma_a A + A\sigma_a\right) + ir_a\left(\left(\sigma_a A - A\sigma_a\right)\right) \Rightarrow A = 0$
$-2ibC$
$= \frac{1}{2}\sum_{a=1}^{3} w_a\left(\sigma_a C + C\sigma_a\right) - ir_a\left(\left(C\sigma_a - \sigma_a C\right)\right)$
$= -\frac{1}{2}\sum_{a=1}^{3} w_a\left(\sigma_a C + C\sigma_a\right) + ir_a\left(\left(C\sigma_a - \sigma_a C\right)\right) \Rightarrow C = 0$
We are left with :
$\forall w, r : \sum_{a=1}^{3}\left(w_a + ir_a\right)\sigma_a B = \sum_{a=1}^{3}\left(w_a + ir_a\right)B\sigma_a$
which implies that B commutes with all the Dirac matrices, which happens only for the scalar matrices : $B = k\sigma_0$.
$G = \begin{bmatrix} 0 & k\sigma_0 \\ \overline{k}\sigma_0 & 0 \end{bmatrix}$ ■

The scalar product will never be definite positive, so we can take $k = -i$ that is $G = \gamma_0$. And it is easy to check that it works also for the signature (1,3).

Any vector of E reads :
$u = \sum_{i=1}^{4} u^i e_i = u_R + u_L$ with $u_R = \sum_{i=1}^{2} u^i e_i, u_L = \sum_{i=3}^{4} u^i e_i$
The scalar product of two vectors u, v of E is then:

$$\left\langle \sum_{i=1}^{4} u^i e_i, \sum_{i=1}^{4} v^i e_i \right\rangle_E = [u]^*[\gamma_0][v] = i\left(u_L^* v_R - u_R^* v_L\right) \quad (4.24)$$

It is not definite positive. It is preserved both by $Spin(3,1)$ and $Spin(1,3)$.
It is definite positive on E_R and definite negative on E_L.

The basis $(e_i)_{i=1}^4$ of E is not orthonormal : $\langle e_j, e_k\rangle = i\begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$

The scalar product is *not invariant* by $T_1 Spin\left(3,1\right)$ (which acts by γ). We have the following formula :
$\left\langle \gamma C\left(\upsilon\left(r,w\right)\right)S, \gamma C\left(\upsilon\left(r,w\right)\right)S\right\rangle = -\frac{1}{2}\left(\left(w^t w - r^t r\right)\operatorname{Im} S_R^* S_L + 2\left(r^t w\right)\operatorname{Re} S_R^* S_L\right)$

4.2.7 Norm on the space E of spinors

The two chiral operators p_R, p_L on $Cl\left(\mathbb{C},4\right)$ do not belong to the images $\gamma C\left(Cl\left(3,1\right)\right), \gamma C\left(Cl\left(1,3\right)\right)$ however they define two subspaces $E = E_R \oplus E_L$ and the elements of the images $C\left(Spin\left(3,1\right)\right)$ and $C'\left(Spin\left(1,3\right)\right)$ by γ preserve both E_R, E_L. So we have operators $\gamma_R : E \to E_R, \gamma_L : E \to E_L$ such that :
$\gamma_R = \gamma_R \cdot \gamma_R; \gamma_L = \gamma_L \cdot \gamma_L$
$\gamma_R + \gamma_L = Id$

$\forall \sigma \in Spin(3,1), Spin(1,3) : \gamma C(\sigma) \circ \gamma_R = \gamma_R \circ \gamma C(\sigma); \gamma C(\sigma) \circ \gamma_L = \gamma_L \circ \gamma C(\sigma)$

$\gamma_R = \gamma(p_R), \gamma_L = \gamma(p_L)$ are complex linear maps, as images of the complex linear maps p_R, p_L by the complex linear map γ. So they preserve any real structure on E : $\gamma_R(\operatorname{Re} u + i \operatorname{Im} u) = \gamma(p_R)(\operatorname{Re} u + i \operatorname{Im} u) = \gamma(p_R)\operatorname{Re} u + i\gamma(p_R)(\operatorname{Im} u)$ and $\gamma_R = \overline{\gamma}_R, \gamma_L = \overline{\gamma}_L$ (Maths.357).

In the basis $(e_i)_{i=1}^4$:

$\gamma_R = \begin{bmatrix} \sigma_0 & 0 \\ 0 & 0 \end{bmatrix} = \gamma_R^*$

$\gamma_L = \begin{bmatrix} 0 & 0 \\ 0 & \sigma_0 \end{bmatrix} = \gamma_L^*$

There is no definite scalar product on E, but there is a norm.

Theorem 66 *The vector space E is a normed vector space with the norm, invariant by $Spin(3,1), Spin(1,3)$:*
$\|S\|_E = \sqrt{\langle \gamma_R S, \gamma_R S \rangle_E - \langle \gamma_L S, \gamma_L S \rangle_E}$

Proof. E_R, E_L are two 2 dimensional complex vector spaces, they can be endowed with a norm which is invariant by $Spin(3,1), Spin(1,3)$:

$S \in E_R : \|S\|_{E_R}^2 = \langle S, S \rangle_E$

$S \in E_L : \|S\|_{E_L}^2 = -\langle S, S \rangle_E$

The norm are invariant by γ_R, γ_L :

$S \in E_R \Leftrightarrow \exists S' \in E : S = \gamma_R(S')$

$\|\gamma_R(S)\|_{E_R} = \|\gamma_R^2(S')\|_{E_R} = \|\gamma_R(S')\|_{E_R} = \|S\|_{E_R}$

Define : $\|\|_E : E \times E \to \mathbb{R} :: \|S\|_E = \sqrt{\|\gamma_R S\|_{E_R}^2 + \|\gamma_L S\|_{E_L}^2}$

$\|S\|_E = 0 \Rightarrow \|\gamma_R S\|_{E_R} = 0; \|\gamma_L S\|_{E_L} = 0 \Rightarrow \gamma_R S = 0; \gamma_L S = 0 \Rightarrow (\gamma_R + \gamma_L)[S] = 0 = S$

$\|S + S'\|_E^2 = \|\gamma_R(S + S')\|_{E_R}^2 + \|\gamma_L(S + S')\|_{E_L}^2 = \|S + S'\|_{E_R}^2 + \|S + S'\|_{E_L}^2$

$\leq \|S\|_{E_R}^2 + \|S'\|_{E_R}^2 + \|S\|_{E_L}^2 + \|S'\|_{E_L}^2 = \|S\|_E^2 + \|S'\|_E^2$

This norm is invariant by $Spin(3,1), Spin(1,3)$:

$\|\gamma C(\sigma) S\|_E = \sqrt{\|\gamma C(\sigma) \gamma_R S\|_{E_R}^2 + \|\gamma C(\sigma) \gamma_L S\|_{E_L}^2}$

$= \sqrt{\|\gamma_R \gamma C(\sigma) S\|_{E_R}^2 + \|\gamma_L \gamma C(\sigma) S\|_{E_L}^2}$

$= \sqrt{\|\gamma C(\sigma) S\|_{E_R}^2 + \|\gamma C(\sigma) S\|_{E_L}^2} = \sqrt{\|S\|_{E_R}^2 + \|S\|_{E_L}^2} = \|S\|_E$ ■

4.3 THE SPINOR MODEL OF KINEMATICS

We have now the mathematical tools to enter the representation of kinematics of material bodies in General Relativity. First we will make some adjustments to the fiber bundles used so far to represent the geometry, to account for the introduction of the Spin group.

4.3.1 Description of the fiber bundles

The geometric fiber bundles

The geometric model is similar to the previous one, with the replacement of $SO(3,1)$ by the Spin group.

Definition 67 *The* ***principal bundle*** $P_G(M, Spin_0(3,1), \pi_G)$ *has for fiber the connected component of the unity of the Spin group, for trivialization the map :*
$\varphi_G : M \times Spin_0(3,1) \to P_G :: p = \varphi_G(m,s)$.
The ***standard gauge*** *used by observers is* $\mathbf{p}(m) = \varphi_G(m, \mathbf{1})$

A section $\sigma \in \mathfrak{X}(P_G)$ is defined by a map: $\sigma : M \to Spin(3,1)$ such that : $\sigma(m) = \varphi_G(m, \sigma(m))$ and in a change of gauge :

$$\left[\begin{array}{c} \mathbf{p}(m) = \varphi_G(m,1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m) \cdot \chi(m)^{-1} : \\ \sigma(m) = \varphi_G(m,\sigma) = \widetilde{\varphi}_G(m, \chi(m) \cdot \widetilde{\sigma}) : \widetilde{\sigma} = \chi(m) \cdot \sigma \end{array}\right] \tag{4.25}$$

Definition 68 *The* ***vectors on the tangent bundle TM*** *are represented in the associated vector bundle* $P_G[\mathbb{R}^4, \mathbf{Ad}]$ *defined through the holonomic orthonormal basis :*
$\varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i)$

So $\varepsilon_0(m) = (\mathbf{p}(m), \varepsilon_0)$ is the 4th vector both in the Clifford algebra and in the tangent space T_mM. It corresponds to the velocity of the observer. In $P_G[\mathbb{R}^4, \mathbf{Ad}]$ the components of vectors are measured in orthonormal bases.

With the equivalence relation :$(\mathbf{p}(m), v) \sim \left(\varphi_G(m,g), \mathbf{Ad}_{g^{-1}} v\right)$

In a change of gauge on P_G the holonomic basis becomes :

$$\left[\begin{array}{c} \mathbf{p}(m) = \varphi_G(m,1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m) \cdot \chi(m)^{-1} : \\ \varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i) \to \widetilde{\varepsilon}_i(m) = \mathbf{Ad}_{\chi(m)^{-1}} \varepsilon_i(m) \end{array}\right]$$

For a given observer $\varepsilon_0(p_o(t)) = \frac{1}{c}\frac{dp_o}{dt}$ is fixed along his world line.

The Lorentz scalar product on $\mathbb{R}^4$ is preserved by $\mathbf{Ad}$ thus it can be extended to $P_G[\mathbb{R}^4, \mathbf{Ad}]$.

At each point of M there are orthonormal bases provided by P_G, and the tangent vector space if endowed with a metric. So there is a structure of Clifford algebra $Cl((T_mM, g(m)))$ at each point $m \in M$, whose elements are defined through products of vectors $\varepsilon_i(m)$, and they are isomorphic to $Cl(3,1)$. This common structure is a **Clifford bundle** noted $Cl(TM)$. Its elements change in a change of gauge as products of vectors of $P_G[\mathbb{R}^4, \mathbf{Ad}]$, that is an associated vector bundle $P_G[Cl(3,1), \mathbf{Ad}]$.

Definition 69 *The* ***Clifford bundle*** $Cl(TM)$ *is the associated vector bundle* $P_G[Cl(3,1), \mathbf{Ad}]$ *defined through the basis* $(\varepsilon_i(m))_{i=0}^3$.

In a change of gauge on P_G the elements of Cl(m) transforms as :

$$\left[\begin{array}{l}\mathbf{p}(m)=\varphi_G(m,1)\to\widetilde{\mathbf{p}}(m)=\mathbf{p}(m)\cdot\chi(m)^{-1}:\\ \mathbf{X}(m)=(\mathbf{p}(m),X)=\left(\widetilde{\mathbf{p}}(m),\widetilde{X}\right):\widetilde{X}=\mathbf{Ad}_{\chi(m)^{-1}}X\end{array}\right] \tag{4.26}$$

The kinematic bundle

In addition to the previous bundles we define the associated bundle in which the spinors live :

Proposition 70 *The kinematics characteristics of particles are represented by* ***Spinors,*** *which are, at each point of the world line of the particle, vectors of the associated vector bundle* $P_G[E,\gamma C]$ *. They are measured by observers in the standard gauge defined through the holonomic basis :* $\mathbf{e}_i(m)=(\mathbf{p}(m),e_i)$

With the equivalence relation : $(\mathbf{p}(m),S)\sim\left(\varphi_G(m,g),\gamma C\left(g^{-1}\right)S\right)$ so that in a change of gauge the holonomic basis becomes :

$$\left[\begin{array}{l}\mathbf{p}(m)=\varphi_G(m,1)\to\widetilde{\mathbf{p}}(m)=\mathbf{p}(m)\cdot\chi(m)^{-1}:\\ \mathbf{e}_i(m)=(\mathbf{p}(m),e_i)\to\widetilde{\mathbf{e}}_i(m)=\gamma C\left(\chi(m)^{-1}\right)\mathbf{e}_i(m)\\ \mathbf{S}(m)=(\mathbf{p}(m),S)=\left(\widetilde{\mathbf{p}}(m),\widetilde{S}\right):\widetilde{S}=\gamma C(\chi(m))S\end{array}\right]$$

With the representation $(E,\gamma C)$ one can then define (Maths.2111), pointwise, an **action** of $Cl(TM)$ on $P_G[E,\gamma C]$:

$\Gamma: Cl(TM)\times P_G[E,\gamma C]\to P_G[E,\gamma C]::$

$\Gamma(X(m),(\varphi_G(m,\sigma),S(m)))=\left(\varphi_G(m,\sigma),\gamma C\left(\sigma^{-1}\cdot X(m)\cdot\sigma\right)S(m)\right)$

So that, if $S(m)=\gamma C(\sigma(m))S_0:\Gamma(X(m),S(m))=\gamma C\left(\mathbf{Ad}_{\sigma^{-1}}X(m)\right)S_0$

Notice that the action *is not* $\gamma C(X\cdot\sigma)S$, even if $X\cdot\sigma\in Cl(3,1)$. $P_G[E,\gamma C]$ is a vector bundle associated to P_G and not to $Cl(TM)$. $P_G[E,\gamma C]$ is a representation of the Lie algebra $T_1Spin(3,1)$ but the action is different :

$X,Y\in T_1Spin(3,1)::\gamma C([X,Y])S_0=(\gamma C(X)\circ\gamma C(Y)-\gamma C(Y)\circ\gamma C(X))S_0$

From a mathematical point of view the holonomic bases $(\varepsilon_i(m))_{i=0}^3,(\mathbf{e}_i(m))_{i=1}^4$, are defined through the same standard gauge $\mathbf{p}(m)$ chosen by the observer. This gauge is arbitrary. For the tetrad the vectors $\varepsilon_i(m)$ can be measured in the holonomic basis of any chart through P. We have nothing similar for $\mathbf{e}_i(m)$, and actually the vectors e_i of E themselves are abstract. However we will see in the following how the basis $\mathbf{e}_i(m)$ used by an observer can be related to physical phenomena (inertial observers).

The scalar product on E is preserved by γC thus it can be extended to $P_G[E,\gamma C]$ and to the space of sections $\mathfrak{X}(P_G[E,\gamma C])$ by :

$\left\langle\mathbf{S},\mathbf{S}'\right\rangle=\int_\Omega\left\langle\mathbf{S}(m),\mathbf{S}'(m)\right\rangle_E\varpi_4(m)$

Moreover we have the following :

Theorem 71 *The set of integrable sections :*

$L^1(\mathfrak{X}(P_G[E,\gamma C]),\varpi_4)=\left\{\int_\Omega\|S\|\varpi_4<\infty\right\}$

with the norm on E is a separable, infinite dimensional Fréchet space.

Proof. Consider the vector space : $\mathfrak{X}(P_G[E,\gamma C])$ endowed with the norm :
$\|\mathbf{S}\| = \int_\Omega \|S(m)\|_E \varpi_4(m)$ and the norm
$\|S(m)\|_E = \sqrt{\langle \gamma_R S, \gamma_R S\rangle_E - \langle \gamma_L S, \gamma_L S\rangle_E}$
Restrict this space to $L^1(M, P_G[E,\gamma C], \varpi_4)$
$= \left\{\mathbf{S} \in \mathfrak{X}(P_G[E,\gamma C]) \int_\Omega \|S(m)\|_E \varpi_4(m) < \infty\right\}$

This is a Fréchet space (Maths.2276). Moreover it is separable, because Ω is relatively compact and the smooth compactly supported maps are a countable basis in L^1 (see Lieb).

Because the norm is invariant by the Spin group this space does not depend on the choice of trivialization. ■

The result still holds if we impose that the sections are differentiable.

Fundamental symmetry breakdown

The observer uses the frame $\left(O, (\varepsilon_i)_{i=0}^3\right)$ to measure the components of vectors of TM, and the holonomic maps $(\mathbf{e}_i(m))_{i=0}^3$ to measure the spinors. The breakdown, specific to each observer, comes from the distinction of his present, and is materialized in his standard basis by the vector $\varepsilon_0(m)$.This choice leads to a split of the Spin group between the spatial rotations, represented by $Spin(3)$, and the homogeneous space $SW = Spin(3,1)/Spin(3)$.

We have an associated fiber bundle :
$P_W = P_G[SW,\lambda]$:
$(\mathbf{p}(m), s_w) = (\varphi_G(m,1), s_w) \sim \left(\varphi_G(m,s), \lambda\left(s^{-1}, s_w\right)\right)$
with the left action :
$\lambda : Spin(3,1) \times SW \to SW : \lambda(s, s_w) = \pi_w(s \cdot s_w)$
On the manifold P_G there is a structure of principal fiber bundle
$P_R(P_W, Spin(3), \pi_R)$ with trivialization :
$\varphi_R : P_W \times Spin(3) \to P_G ::$
$\varphi_R((\mathbf{p}(m), s_w), s_r) = \varphi_G(m, s_w \cdot s_r) = \varphi_R\left(\left(\varphi_G(m,s), \lambda\left(s^{-1}, s_w\right)\right), s_r\right)$

As the latest trivialization shows, for a given s, s_r depends on s_w in that it is a part of $s \in Spin(3,1)$.

It sums up to define the local basis in two steps : first by choosing s_w second by choosing s_r

Any section $\sigma \in \mathfrak{X}(P_G)$ can be decomposed, for a given vector field ε_0 and a fixed $\epsilon = \pm 1$, in two sections :
$\epsilon\sigma_w \in \mathfrak{X}(P_W), \epsilon\sigma_r \in \mathfrak{X}(P_R)$ with $\sigma(m) = \epsilon\sigma_w(m) \cdot \epsilon\sigma_r(m)$

The set of vectors of T_mM used to build $Spin(3)$ is defined by $\varepsilon_0(m)$.

4.3.2 Representation by Spinors

The idea is that, for any particle (that is material body which can be seen as located at a point) there is an intrinsic property which characterizes its kinematics behavior and can be represented as a spinor : a map from its world line into the fiber bundle $P_G[E,\gamma C]$. Its existence and value does not depend on any chart or gauge. In a gauge which would be fixed with respect to the particle the measure of S is S_0 which is an intrinsic characteristic of the particle. In the gauge of an observer the measure is $S = \gamma C(\sigma)(S_0)$. This sums up to define $\sigma \in Spin(3,1)$ as a change of gauge in the principal bundle $P_G(M, Spin_0(3,1), \pi_G)$ and, because the same principal bundle P_G is used to represent spinors and orthogonal frames, σ is the rotation which should be necessary so that an orthogonal basis attached to the particle coincides with the tetrad of the observer. Because any element of $Spin(3,1)$ can be written $\sigma = \sigma_w \cdot \sigma_r$ one can decompose this rotation in a rotation with respect to a time like vector ε_0, which corresponds to the velocity of the observer, and a spatial rotation in the 3 dimensional space of the observer. If this is well settled, one issue is to relate σ to measurable quantities. Usually we do not have an orthonormal frame attached to the particle. However the time axis of any orthonormal basis attached to the particle has the same direction as its velocity u, and indeed this is the only significant obvious feature of its motion. The velocity has a constant, known Lorentz norm, so we can compute the Spin element which brings ε_0 of the observer, to u.

Trajectories and the Spin Group

Theorem 72 *Any section $\sigma \in \mathfrak{X}(P_G)$ defines, for any positive function $f \in C_\infty(\Omega;\mathbb{R}_+)$ and observer, two vector fields $V \in \mathfrak{X}\left(P_G\left[\mathbb{R}^4,\mathbf{Ad}\right]\right)$ by :*

$$V(m) = f(m)\,\mathbf{Ad}_{\sigma(m)}\varepsilon_0(m) = f(m)\left(\left(2a_w^2-1\right)\varepsilon_0 + \epsilon a_w\left(w^1\varepsilon_1+w^2\varepsilon_2+w^3\varepsilon_3\right)\right) \tag{4.27}$$

where $\sigma_w(m) = \epsilon\left(a_w + \upsilon(0,w)\right)$ is the projection of σ on P_W along ε_0

Then V is time like, future oriented and $\langle V(m),V(m)\rangle = -f^2(m)$ and is invariant in a change of gauge on P_G

Conversely, for any time like, future oriented vector field $V \in \mathfrak{X}\left(P_G\left[\mathbb{R}^4,\mathbf{Ad}\right]\right)$ there are two sections $\sigma_w \in \mathfrak{X}(P_G)$ such that :

$\frac{V}{\sqrt{-\langle V,V\rangle}} = u = \mathbf{Ad}_{\sigma(m)}\varepsilon_0(m) : \sigma_w = \epsilon\left(\sqrt{\frac{1}{2}(u_0+1)} + \frac{1}{\sqrt{\frac{1}{2}(u_0+1)}}\upsilon(0,u_i)\right)$

Proof. i) $\sigma(m) = \epsilon\sigma_w(m)\cdot\epsilon\sigma_r(m)$

$\sigma_w(m) = \epsilon\left(a_w+\upsilon(0,w)\right)$ so let be $a_w>0$ which defines ϵ

ii) Define

$u \in \mathfrak{X}\left(P_G\left[\mathbb{R}^4,\mathbf{Ad}\right]\right) : u(m) = \mathbf{Ad}_{\sigma_w(m)}\varepsilon_0(m) = \left(\mathbf{p}(m),\mathbf{Ad}_{\sigma_w}\varepsilon_0\right) = \left(\mathbf{p}(m),u\right)$

$u = \sigma_w\cdot\varepsilon_0\cdot\sigma_w^{-1} = \left(\epsilon a_w+\upsilon(0,w)\right)\cdot\varepsilon_0\cdot\left(\epsilon a_w-\upsilon(0,w)\right)$

$= \left(\epsilon a_w+\upsilon(0,w)\right)\cdot\left(\epsilon a_w\varepsilon_0-\varepsilon_0\cdot\upsilon(0,w)\right)$

$= a_w^2\varepsilon_0+\epsilon a_w\left(-\varepsilon_0\cdot\upsilon(0,w)+\upsilon(0,w)\cdot\varepsilon_0\right)-\upsilon(0,w)\cdot\varepsilon_0\cdot\upsilon(0,w)$

$= a_w^2\varepsilon_0+\frac{1}{2}\epsilon a_w\left(-\varepsilon_0\cdot\varepsilon_0\cdot\left(w^1\varepsilon_1+w^2\varepsilon_2+w^3\varepsilon_3\right)-\left(w^1\varepsilon_1+w^2\varepsilon_2+w^3\varepsilon_3\right)\cdot\varepsilon_0\cdot\varepsilon_0\right)$

$-\frac{1}{4}\varepsilon_0\cdot\left(w^1\varepsilon_1+w^2\varepsilon_2+w^3\varepsilon_3\right)\cdot\varepsilon_0\cdot\varepsilon_0\cdot\left(w^1\varepsilon_1+w^2\varepsilon_2+w^3\varepsilon_3\right)$

$= a_w^2\varepsilon_0+\frac{1}{2}\epsilon a_w(w+w)+\frac{1}{4}\varepsilon_0\cdot w\cdot w$

$= a_w^2\varepsilon_0+a_w\epsilon w+\frac{1}{4}\varepsilon_0\cdot\langle w,w\rangle$

$= \left(a_w^2+\frac{1}{4}w^tw\right)\varepsilon_0+\epsilon a_w w$

$u = \left(2a_w^2-1\right)\varepsilon_0+\epsilon a_w\left(w^1\varepsilon_1+w^2\varepsilon_2+w^3\varepsilon_3\right)$

iii) $\langle u,\varepsilon_0\rangle = -\left(2a_w^2-1\right) = 1-2\left(1+\frac{1}{4}w^tw\right) = -1-\frac{1}{2}w^tw<0$

$\langle u,u\rangle = a_w^2w^tw-\left(2a_w^2-1\right)^2 = a_w^2\left(4\left(a_w^2-1\right)\right)-\left(2a_w^2-1\right)^2 = -1$

iv) $V = f(m)\,u \Rightarrow$

$\langle V,\varepsilon_0\rangle = f(m)\langle u,\varepsilon_0\rangle<0$

$\langle V,V\rangle = -f^2<0$

v) In a change of gauge on P_G :

$\mathbf{p}(m) = \varphi_G(m,1)\rightarrow\widetilde{\mathbf{p}}(m) = \mathbf{p}(m)\cdot\chi(m)^{-1}$:

$\sigma(m) = \varphi_G(m,\sigma(m)) = \widetilde{\varphi}_G(m,\chi(m)\cdot\sigma(m)) = \widetilde{\varphi}_G(m,\widetilde{\sigma}(m))$

$u(m) = \left(\mathbf{p}(m),\mathbf{Ad}_{\sigma_w}\varepsilon_0\right)\sim\left(\mathbf{p}(m)\cdot\chi(m)^{-1},\mathbf{Ad}_\chi\mathbf{Ad}_{\sigma_w}\varepsilon_0\right) = \left(\widetilde{\mathbf{p}}(m),\mathbf{Ad}_{\widetilde{\sigma_w}}\varepsilon_0\right) = \widetilde{u}(m)$

vi) Let $u = \frac{V}{\sqrt{-\langle V,V\rangle}}$then u is time like, $\langle u,u\rangle = -1, \langle u,\varepsilon_0\rangle<0$

$u = \left(\left(2a_w^2-1\right)\varepsilon_0+\epsilon a_w\left(w^1\varepsilon_1+w^2\varepsilon_2+w^3\varepsilon_3\right)\right) = \sum_{i=0}^3 u_i\varepsilon_i$

i=1,2,3 : $w_i = \epsilon u_i / a_w$
$w^t w = \frac{\sum_{i=1}^3 u_i^2}{a_w^2}$
$\langle u, u \rangle = -1 = \sum_{i=1}^3 u_i^2 - u_0^2$
$u_0^2 = 1 + \sum_{i=1}^3 u_i^2$
$\langle u, \varepsilon_0 \rangle = u_0 < 0$
$u_0 = -\sqrt{1 + \sum_{i=1}^3 u_i^2} < -1$
$1 + \frac{1}{4} w^t w = 1 + \frac{1}{4} \frac{\sum_{i=1}^3 u_i^2}{a_w^2} = \frac{1 + 4 \sum_{i=1}^3 u_i^2}{4 a_w^2} = \frac{1 + 4\left(u_0^2 - 1\right)}{4 a_w^2} = \frac{-3 + 4 u_0^2}{4 a_w^2} > \frac{-3+4}{4 a_w^2} = \frac{1}{a_w^2}$
So we can define :
$u_0 = 2 a_w^2 - 1$
$a_w = \epsilon \sqrt{\frac{1}{2} (u_0 + 1)}$
$w_i = u_i / a_w = \epsilon u_i / \sqrt{\frac{1}{2} (u_0 + 1)}$
$\sigma_w = \epsilon \left(\sqrt{\frac{1}{2} (u_0 + 1)} + \frac{1}{\sqrt{\frac{1}{2}(u_0 + 1)}} \upsilon (0, u_i) \right) = a_w + \upsilon (0, w)$
with $w_i = \frac{u_i}{\sqrt{\frac{1}{2}(u_0+1)}} = \frac{V_i}{\sqrt{\frac{1}{2}\left(\frac{V_0}{\sqrt{-\langle V,V\rangle}} + 1\right)}} \frac{1}{\sqrt{-\langle V,V \rangle}} = \frac{V_i}{\sqrt{-\frac{1}{2}\left(\langle V,V\rangle - V_0 \sqrt{-\langle V,V\rangle}\right)}}$ ∎

If we take $f(m) = c$ any section $\sigma \in \mathfrak{X}(P_G)$ defines two fields of world lines, with opposite spatial speed :

$$u = \frac{dp}{d\tau} = \frac{1}{\sqrt{1 - \frac{\|\overrightarrow{v}\|^2}{c^2}}} \left(\overrightarrow{v} + c \varepsilon_0 (m)\right) = c \left(\left(2 a_w^2 - 1\right) \varepsilon_0 + \epsilon a_w \sum_{i=1}^3 w_i \varepsilon_i \right) \tag{4.28}$$

If we take $f(m) = \frac{c}{2a_w^2 - 1}$ any section $\sigma \in \mathfrak{X}(P_G)$ defines two fields of trajectories, with opposite spatial speed :

$$V = \frac{dp}{dt} = \overrightarrow{v} + c \varepsilon_0 (m) = c \left(\varepsilon_0 + \epsilon \frac{a_w}{2 a_w^2 - 1} \sum_{i=1}^3 w_i \varepsilon_i \right) \tag{4.29}$$

$$\left[\begin{array}{l} a_w = \sqrt{\frac{1}{2}\left(1 + \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}\right)} \simeq 1 + \frac{1}{8} \frac{v^2}{c^2} \\ w = \sqrt{\frac{2}{1 + \sqrt{1 - \frac{\|\overrightarrow{v}\|^2}{c^2}}}} \frac{\overrightarrow{v}}{c} \simeq \left(1 + \frac{1}{8} \frac{v^2}{c^2}\right) \frac{\overrightarrow{v}}{c} \end{array} \right] \tag{4.30}$$

Remarks :
i) All this is defined with respect to an observer, who fixes $\varepsilon_0 (m)$
ii) $V \in \mathfrak{X}\left(P_G \left[\mathbb{R}^4, \mathbf{Ad}\right]\right)$ so can be equivalently defined as a section of TM :
$V = \sum_{\alpha=0}^3 V^\alpha \partial \xi_\alpha = \sum_{\alpha=0}^3 \sum_{i=0}^3 V^i P_i^\alpha \partial \xi_\alpha$
iii) If V is past oriented ($u_0 < 0$) or null ($\langle V, V \rangle = 0$) there is no solution :
$2 a_w^2 - 1 = \frac{1}{2}(u_0 - 1) < -\frac{1}{2} \Rightarrow a_w^2 < 1$ and $a_w^2 \neq 1 + \frac{1}{4} w^t w$
This gives a strong physical meaning to the representation of world lines by section of P_W.
iv) Any map $\sigma : \mathbb{R} \to P_G$ is projected on M as a curve, which is not necessarily time like or defines a world line.
v) From the formula above V has the dimension of a spatial speed, and w is unitless, by the use of the universal constant c, which provides a natural standard.

Spatial Spinor

A section $\sigma \in \mathfrak{X}(P_G)$ defines at each point an element of $Spin(3,1)$, which can uniquely (up to sign) be decomposed in $\sigma = \epsilon\sigma_w \cdot \epsilon\sigma_r$ with respect to a given observer.

The first component $\sigma_w = \pm(a_w + \upsilon(0,w)) \in Spin(3,1)$ defines, in the standard basis of the observer, a vector field of world lines, and a trajectory. w is aligned in the direction of the spatial speed or the opposite.

The second part $\sigma_r = a_r + \upsilon(0,r)$ belonging to one of the two connected components of $Spin(3)$ (according to the sign of a_r, it is + for the component of the identity) leaves invariant $\varepsilon_0(m)$ and defines a spatial rotation, in the hyperplane orthogonal to $\varepsilon_0(m)$.

So with a single quantity σ we can define the disposition of a basis which would be attached to the particle, with respect to the basis of the observer.

The decomposition $\sigma = \epsilon\sigma_w \cdot \epsilon\sigma_r$ and the identification of the vectors of $Spin(3)$ request a choice of $\varepsilon_0(m)$: it is observer-dependant.

Notice that the action $\mathbf{Ad}_{\sigma_r}$ on any vector of the tangent space T_mM rotates the vector, but leaves invariant $\varepsilon_0(m)$,so this is an action on the physical space $\Omega_3(t)$. In the Spinor $S = \gamma C(\sigma(m)) S_0$ this action is done on vectors of E, and not on vector of T_mM.

Definition 73 *We can then define the* ***spatial spinor*** *as :*

$$\mathbf{S}_r(t) = \gamma C\left(\sigma_w^{-1}\right)\mathbf{S}(t) = \gamma C(\sigma_r(t)) S_0 \tag{4.31}$$

For a given trajectory and observer there are two possible, opposite, values of the spatial spinor : $S_r(t) = \pm\gamma C(\sigma_r(t)) S_0$. In all cases $S(t) = \gamma C(\sigma(t)) S_0$: the total spinor stays the same, the distinction between σ_w, σ_r and the opposite values is the consequence of the breakdown of symmetry induced by the observer. The sign $\pm$ is related to a trajectory (the orientation of w with respect to the spatial speed) so one can speak of **spin up or down** with respect to the trajectory. This feature is entirely linked to the Relativist picture, and has nothing to do with QM. The name spin is used freely in Quantum Physics, and this is sometimes confusing. And to be clear I will call the present feature (spin up or down) **Relativist Spin** which takes the values of $\epsilon = \pm 1$.

If we assume that the spatial spinor is, by itself, an intrinsic feature of the particle, then one must assume that the map : $\sigma_r : \mathbb{R} \to P_R$ is continuous, thus σ_r must belong and stay in one of the two connected components of P_R. Normally the decomposition $\sigma = \epsilon\sigma_w \cdot \epsilon\sigma_r$ is continuous, and the passage to the opposite sign is, for the spatial spinor, a discontinuity, and also for the relativist spin.

The issue now is to precise what can be S_0,that we will call **inertial spinor**. The 4 dimensional relativist momentum $P = \mu u$, which is a geometric quantity, has a constant scalar product : $\langle P, P\rangle = -\mu^2 c^2$ where μ is, by definition, the mass at rest. The scalar product $\langle S, S\rangle = \langle S_0, S_0\rangle$ is preserved on the world line, so we will look at vectors S such that : $S \neq 0 \Rightarrow \langle S_0, S_0\rangle \neq 0$.

Inertial spinor

Definition

The unique scalar product on E is non degenerate, but not definite positive. So it is logical to require that S_0 belongs to some vector subspace E_0 of E, over which the scalar product is definite, either positive of negative. Moreover a change of spatial frame should change only S_r, thus E_0 should be invariant under the action of $Spin(3)$.

So there should be some vector subspace E_0 of E such that :
- it is invariant by $\gamma C(\sigma_r)$ for $\sigma_r \in Spin(3) : \forall S_0 \in E_0, s_r \in Spin(3) : \gamma C(s_r) S_0 \in E_0$
- on which the scalar product is either definite positive or definite negative :
$\forall S_0 \in E_0 : \langle S_0, S_0 \rangle_E = 0 \Rightarrow S_0 = 0$

Theorem 74 *The only vector subspace of E invariant by γC on $Spin(3)$ and over which the scalar product is definite*
- *positive is* $E_0 = \left\{ S = \begin{bmatrix} S_R \\ S_L \end{bmatrix} = \begin{bmatrix} v \\ iv \end{bmatrix}, v \in \mathbb{C}^2 \right\}$
- *negative is* $E_0' = \left\{ S = \begin{bmatrix} S_R \\ S_L \end{bmatrix} = \begin{bmatrix} v \\ -iv \end{bmatrix}, v \in \mathbb{C}^2 \right\}$

Proof. i) The scalar product on E (which does not depend on the signature) reads :
$u = \begin{bmatrix} u_R \\ u_L \end{bmatrix} \in E : \begin{bmatrix} u_R^* & u_L^* \end{bmatrix} \begin{bmatrix} 0 & -i\sigma_0 \\ i\sigma_0 & 0 \end{bmatrix} \begin{bmatrix} u_R \\ u_L \end{bmatrix} = i(u_L^* u_R - u_R^* u_L) = i\left(\overline{u}_L^t u_R - \left(\overline{u}_R^t u_L\right)\right)$
$u_L = v_L + iw_L$ with $v_L, w_L \in \mathbb{R}^2$
$u_R = v_R + iw_R$ with $v_R, w_R \in \mathbb{R}^2$
$\langle S, S \rangle = i\left((v_L^t - iw_L^t)(v_R + iw_R) - (v_R^t - iw_R^t)(v_L + iw_L)\right)$
$= i\left(v_L^t v_R + iv_L^t w_R - iw_L^t v_R + w_L^t w_R - v_R^t v_L - iv_R^t w_L + iw_R^t v_L - w_R^t w_L\right)$
$= -2\left(v_L^t w_R - w_L^t v_R\right)$
$\langle S, S \rangle = 0 \Leftrightarrow v_L^t w_R = w_L^t v_R$
So it is definite for any u such that :
$v_L = -\epsilon w_R, w_L = \epsilon v_R \Rightarrow$
$u_L = -\epsilon w_R + \epsilon i v_R = \epsilon i(v_R + iw_R) = \epsilon i u_R$
$u = \begin{bmatrix} u_R \\ \epsilon i u_R \end{bmatrix}$
$\langle S, S \rangle = 2\epsilon\left(w_R^t w_R + v_R^t v_R\right) = 2\epsilon u_R^* u_R$
It is definite positive for $\epsilon = +1$ and definite negative for $\epsilon = -1$
The sets E_0, E_0' are vector subspaces of E, and the scalar product is definite positive of negative, on them.
ii) The vector subspace must be invariant by $\gamma C(s_r)$. Which is equivalent to $S_L = \epsilon i S_R$
For any $S_0 \in E_0, E_0', s \in Spin(3,1)$
$\gamma C(a + \upsilon(r, w) + b\varepsilon_5) S_0 =$
$\begin{bmatrix} (a+ib)\sigma_0 + \frac{1}{2}\sum_a (w_a - ir_a)\sigma_a & 0 \\ 0 & (a-ib)\sigma_0 - \frac{1}{2}\sum_a (w_a + ir_a)\sigma_a \end{bmatrix} \begin{bmatrix} v \\ \epsilon iv \end{bmatrix}$
$= \begin{bmatrix} S_R \\ S_L \end{bmatrix}$
$S_R = \left((a+ib)\sigma_0 + \frac{1}{2}\sum_{a=1}^3 (w_a - ir_a)\sigma_a\right) v$
$S_L = \epsilon\left((a-ib)\sigma_0 - \frac{1}{2}\sum_{a=1}^3 (w_a + ir_a)\sigma_a\right) iv$
and $S \in E_0 \Leftrightarrow S_L = \epsilon i S_R$
$\Leftrightarrow \epsilon i\left((a-ib)\sigma_0 - \frac{1}{2}\sum_{a=1}^3 (w_a + ir_a)\sigma_a\right) v$
$= \epsilon i\left((a+ib)\sigma_0 + \frac{1}{2}\sum_{a=1}^3 (w_a - ir_a)\sigma_a\right) v$
$\Leftrightarrow \left(-ib\sigma_0 - \frac{1}{2}\sum_{a=1}^3 w_a\sigma_a\right) v = \left(ib\sigma_0 + \frac{1}{2}\sum_{a=1}^3 w_a\sigma_a\right) v$
$\Leftrightarrow \left(ib\sigma_0 + \frac{1}{2}\sum_{a=1}^3 w_a\sigma_a\right) v = 0$
This condition is met for $w = 0$ that is $s \in Spin(3)$.
iii) It is easy to see that the result does not depend on the signature:

$\gamma C'(s)\begin{bmatrix} v \\ \epsilon iv \end{bmatrix} = \begin{bmatrix} S_R \\ S_L \end{bmatrix} \Rightarrow$
$S_R = \left((a+b)\sigma_0 + \frac{1}{2}\sum_{a=1}^{3}(w_a + ir_a)\sigma_a\right)v$
$S_L = \epsilon i\left((a-b)\sigma_0 + \frac{1}{2}\sum_{a=1}^{3}(-w_a + ir_a)\sigma_a\right)v$
$\left(2b\sigma_0 + \sum_{a=1}^{3} w_a\sigma_a\right)v = 0$ ■

So particles have both a left S_L and a right S_R part, which are linked but not equal. We have one of the known features of particles : chirality.

Because $E_0 \cap E_0' = \{0\}$, $(E_0, \gamma C), (E_0', \gamma C)$ are two, non equivalent, irreducible representations of $Spin(3)$. So they can be seen as corresponding to two kinds of particles according to ϵ.

The inertial spinor is defined, from the components of the two complex vectors of S_R , by 4 real scalars.

Particles and antiparticles

The quantity :

$$\langle S_0, S_0\rangle_E = \epsilon 2S_R^* S_R \tag{4.32}$$

(with the same meaning of ϵ as above) is a scalar, which is conserved along the trajectory and we can assume that it is linked to the mass at rest M_p of the particle : $\langle S_0, S_0\rangle = M_p^2 c^2$, but because $\langle S_0, S_0\rangle$ can be negative we have to consider :

$$\langle S_0, S_0\rangle = \epsilon M_p^2 c^2 \tag{4.33}$$

where ϵ is a characteristic of the particle. We retrieve a celebrated Dirac's result from his equation. So we define :

Definition 75 ***particles*** *are such that* $S_L = iS_R$.
Their mass is $M_p = \frac{1}{c}\sqrt{\langle S_0, S_0\rangle_E} = \frac{1}{c}\sqrt{2S_R^* S_R}$
antiparticles *are such that* $S_L = -iS_R$
Their mass is $M_p = \frac{1}{c}\sqrt{-\langle S_0, S_0\rangle_E} = \frac{1}{c}\sqrt{-2S_R^* S_R}$

Do antiparticles have negative mass ? The idea of a negative mass is still controversial. Dirac considered that antiparticles move backwards in time and indeed a negative mass combined with the first Newton's law seems to have this effect. But here the world line of the particle is defined by σ_w, and there is no doubt about the behavior of an antiparticle : it moves towards the future. The mass at rest M_p is somewhat conventional, the defining relation is $\langle S_0, S_0\rangle = \epsilon M_p^2 c^2$ so we can choose any sign for M_p, and it seems more appropriate to take $M_p > 0$ both for particles and antiparticles.

Then S_R reads :
$S_R = \frac{M_p}{c\sqrt{2}}\begin{bmatrix} a \\ b \end{bmatrix}$ and $1 = \left(|a|^2 + |b|^2\right)$

It is customary to represent the polarization of the plane wave of an electric field by two complex quantities (the Jones vector) :
$E_x = E_{0x}e^{i\alpha_x}$
$E_y = E_{0y}e^{i\alpha_y}$
where (E_{0x}, E_{0y}) are the components of a vector E_0 along the axes x, y.

So we can write similarly :

$$S_R = \frac{M_p c}{\sqrt{2}} \begin{bmatrix} e^{i\alpha_1} \cos \alpha_0 \\ e^{i\alpha_2} \sin \alpha_0 \end{bmatrix} \quad (4.34)$$

Particles :

$S_0 = \frac{M_p c}{\sqrt{2}} \begin{bmatrix} e^{i\alpha_1} \cos \alpha_0 \\ e^{i\alpha_2} \sin \alpha_0 \\ ie^{i\alpha_1} \cos \alpha_0 \\ ie^{i\alpha_2} \sin \alpha_0 \end{bmatrix}$

Antiparticles :

$S_0 = \frac{M_p c^2}{\sqrt{2}} \begin{bmatrix} e^{i\alpha_1} \cos \alpha_0 \\ e^{i\alpha_2} \sin \alpha_0 \\ -ie^{i\alpha_1} \cos \alpha_0 \\ -ie^{i\alpha_2} \sin \alpha_0 \end{bmatrix}$

To each particle corresponds an antiparticle with the same mass. And particles show polarization characteristics similar to waves. This is how the inertial spinor is seen in Quantum Physics : the 3 real variables $\alpha_0, \alpha_1, \alpha_2$ define a polarization of the particle when it behaves as a wave. Each kind of elementary particles is characterized by a vector of S_0, and we will see that it is associated to its charge with respect to the EM field, and the 3 variables $\alpha_0, \alpha_1, \alpha_2$ define the magnetic moment.

The space E_0, E_0' are orthogonal, so :
$\forall S_0 \in E_0, S_0' \in E_0' : \langle \gamma C(\sigma) S_0, \gamma C(\sigma) S_0' \rangle = 0$
The definition of E_0, E_0' does not depend on the observer.

Space and time reversal

A change of orthonormal basis in $\mathbb{R}^4$ is represented by an orthogonal matrix, and in the Clifford algebra by the action $\mathbf{Ad}_s$ for some element s of the Pin group (it is not necessarily represented by an element of the connected component of the Spin group) :
$w \to \widetilde{w} = \mathbf{Ad}_s w = s \cdot w \cdot s^{-1}$
The impact on a representation is :
$\gamma(w) \to \gamma(\widetilde{w}) = [\gamma(s)] [\gamma(w)] \left[\gamma(s)^{-1}\right]$
In a change of basis in E represented by a matrix Q the components of a vector $u \in E$ change according to : $[u] \to [\widetilde{u}] = Q^{-1}[u]$ and the matrices γ representing endomorphisms change as : $\gamma \to \widetilde{\gamma} = Q^{-1}\gamma Q$. So the change of basis in $\mathbb{R}^4$ corresponds to a change of basis represented by the matrix $Q = [\gamma(s)]^{-1}$ in E, and the components of a vector S of E change as : $[u] \to [\widetilde{u}] = [\gamma(s)][u]$.

Time reversal

Time reversal is the operation :
$u = u^0\varepsilon_0 + u^1\varepsilon_1 + u^2\varepsilon_2 + u^3\varepsilon_3 \to -u^0\varepsilon_0 + u^1\varepsilon_1 + u^2\varepsilon_2 + u^3\varepsilon_3$
corresponding to $s = \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$, with $s^{-1} = \varepsilon_3 \cdot \varepsilon_2 \cdot \varepsilon_1$ in $Cl(3,1)$, $s^{-1} = \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$ in $Cl(1,3)$
$Cl(3,1)$:
$[\gamma C(\varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3)] = \gamma_3\gamma_2\gamma_1 = i \begin{bmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{bmatrix}$
$i \begin{bmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{bmatrix} \begin{bmatrix} v \\ \epsilon i v \end{bmatrix} = \begin{bmatrix} -\epsilon v \\ iv \end{bmatrix} = \begin{bmatrix} v' \\ -\epsilon i v' \end{bmatrix}$

$Cl(1,3)$:

$[\gamma C(\varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3)] = -i\gamma_1\gamma_2\gamma_3 = \begin{bmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{bmatrix}$

$\begin{bmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{bmatrix} \begin{bmatrix} v \\ \epsilon i v \end{bmatrix} = \begin{bmatrix} \epsilon i v \\ v \end{bmatrix} = \begin{bmatrix} v' \\ -\epsilon i v' \end{bmatrix}$

So with both signatures particles and antiparticles are exchanged.

Space reversal

Space reversal is the operation :

$u = u^0\varepsilon_0 + u^1\varepsilon_1 + u^2\varepsilon_2 + u^3\varepsilon_3 \to u^0\varepsilon_0 - u^1\varepsilon_1 - u^2\varepsilon_2 - u^3\varepsilon_3$

corresponding to $s = \varepsilon_0, s^{-1} = -\varepsilon_0$ in $Cl(3,1)$, $s^{-1} = \varepsilon_0$ in $Cl(1,3)$

$Cl(3,1)$:

$[\gamma C(\varepsilon_0)] = i\gamma_0 = \begin{bmatrix} 0 & \sigma_0 \\ -\sigma_0 & 0 \end{bmatrix}$

$\begin{bmatrix} 0 & \sigma_0 \\ -\sigma_0 & 0 \end{bmatrix} \begin{bmatrix} v \\ \epsilon i v \end{bmatrix} = \begin{bmatrix} \epsilon i v \\ -v \end{bmatrix} = \begin{bmatrix} v' \\ \epsilon i v' \end{bmatrix}$

$Cl(1,3)$:

$[\gamma C(\varepsilon_0)] = \gamma_0 = \begin{bmatrix} 0 & -i\sigma_0 \\ i\sigma_0 & 0 \end{bmatrix}$

$\begin{bmatrix} 0 & -i\sigma_0 \\ i\sigma_0 & 0 \end{bmatrix} \begin{bmatrix} v \\ \epsilon i v \end{bmatrix} = \begin{bmatrix} \epsilon v \\ i v \end{bmatrix} = \begin{bmatrix} v' \\ \epsilon i v' \end{bmatrix}$

So with both signatures particle and antiparticles stay in the same category.

These results are consistent with what is checked in Particles Physics, and the Standard Model. However the latter does not consider both signatures. This feature does not allow to distinguish one signature as more physical than the other.

The demonstration above is actually the equivalent - expressed in the formalism of fiber bundles and spinors - of the classic Wigner's classification of particles (see for instance Weinberg), done through the analysis of equivariance of the relativist momentum by the Poincaré's group.

The sets of Spinors

The spatial spinor S_r belongs to E_0, E_0' by construct. The action of $Spin(3)$ on E_0, E_0' is proper, continuous and free, thus (Maths.1793) the orbits have a unique structure of manifold of dimension : $\dim E_0 - \dim Spin(3) = 1$. For any value $S_0 \in E_0, E_0'$ when $\sigma_r \in Spin(3)$ then $S_r = \gamma C(\sigma_r) S_0$ stays on a curve on E_0, E_0', and conversely each vector S of E_0, E_0' belongs to a unique such curve. $Spin(3)$ is compact, so this curve is compact, $Spin(3)$ has two connected components, so the curve is formed of two, compact, connected components.

For a given particle its spinor can take any value $S = \gamma C(\sigma) S_0 = \gamma C(\sigma_w) S_r$. The relation $S_L = \epsilon i S_R$ does not hold any more at the level of the total spinor, however we have still $\langle S, S\rangle = \langle S_0, S_0\rangle_E$ which is positive for particles, and negative for anti-particles, so the distinction holds. The total spinor S belongs to a subset $\hat{E}_0$ of E larger than E_0.

Definition 76 $\hat{E}_0 = \{\gamma C(\sigma_w) S_0, \sigma_w \in Spin(3,1)/Spin(3), S_0 \in E_0\}$
$= \left\{ \left(a_w - i\frac{1}{2}\sum_{a=1}^{3} w_a\gamma_a\gamma_0 \right) S_0, S_0 \in E_0 \right\}$

with a similar set $\hat{E}'_0$ for antiparticles.

The set $\hat{E}_0, \hat{E}'_0$ are not vector spaces (because $a_w = \sqrt{1 + \frac{1}{4}w^t w)}$ but 7 dimensional real manifolds embedded in E.

However the set $\widetilde{E}_0 = Span_{\mathbb{C}}(E_0)$, the complex vector space of E comprised of all linear combinations of vectors of $\widehat{E}_0$, is a Hilbert space.

Theorem 77 *On $\widetilde{E}_0$ the scalar product is definite positive (and definite negative on $\hat{E}'_0$)*

Proof. i) $\forall \lambda \in \mathbb{C} : S \neq 0 \in \hat{E}_0 \Rightarrow \langle \lambda S, \lambda S \rangle > 0$

ii) Let us take $S = \gamma C(\sigma_w) S_0, S' = \gamma C(\sigma_w) S'_0$, and consider $S + S'$

$\langle S + S', S + S' \rangle$

$= \langle \gamma C(\sigma_w) S_0, \gamma C(\sigma_w) S_0 \rangle + \langle \gamma C(\sigma'_w) S'_0, \gamma C(\sigma'_w) S'_0 \rangle + \langle \gamma C(\sigma_w) S_0, \gamma C(\sigma'_w) S'_0 \rangle + \langle \gamma C(\sigma'_w) S'_0, \gamma C(\sigma_w) S_0 \rangle$

$= \langle S_0, S_0 \rangle + \langle S'_0, S'_0 \rangle + 2 \operatorname{Re} \langle \gamma C(\sigma_w) S_0, \gamma C(\sigma'_w) S'_0 \rangle$

$S_0 = \begin{bmatrix} S_R \\ \epsilon i S_R \end{bmatrix} \Rightarrow \gamma C(\sigma_w) S_0 = \begin{bmatrix} \left(\sqrt{1 + \frac{1}{4}w^t w} + \frac{1}{2}\sigma(w) \right) S_R \\ \epsilon i \left(\sqrt{1 + \frac{1}{4}w^t w} - \frac{1}{2}\sigma(w) \right) S_R \end{bmatrix}$

$\langle \gamma C(\sigma_w) S_0, \gamma C(\sigma'_w) S'_0 \rangle$

$= i \left(\left(-\epsilon i S_R^* \left(a_w + \frac{1}{2}\sigma(w) \right) \right) \left(\left(a'_w + \frac{1}{2}\sigma(w') \right) S_R \right) - \left(S_R^* \left(a_w - \frac{1}{2}\sigma(w) \right) \epsilon i \left(a'_w - \frac{1}{2}\sigma(w') \right) S'_R \right) \right)$

$= \epsilon S_R^* \left(2a_w a'_w + \frac{1}{4} \left(\sigma(j(w) w') + \sigma(j(w) w') + 2w^t w' \sigma_{0.} \right) \right) S'_R$

$= \epsilon \left(2a_w a'_w + \frac{1}{2}w^t w' \right) S_R^* S'_R$

$= 2\epsilon \left(\sqrt{1 + \frac{1}{4}w^t w} \sqrt{1 + \frac{1}{4}w'^t w'} + \frac{1}{4}w^t w' \right) S_R^* S'_R$

$\langle S + S', S + S' \rangle$

$= \epsilon 2 \left(S_R^* S_R + S'^*_R S'_R + \left(\sqrt{1 + \frac{1}{4}w^t w} \sqrt{1 + \frac{1}{4}w'^t w'} + \frac{1}{4}w^t w' \right) \left(S_R^* S'_R + S'^*_R S_R \right) \right)$

w, w', S_R, S'_R are independent but the expression is strongly symmetric. Let us consider first w, w'.The minimum is for : $w = w'$:

$= \epsilon 2 \left(S_R^* S_R + S'^*_R S'_R + \left(1 + \frac{1}{2}w^t w \right) \left(S_R^* S'_R + S'^*_R S_R \right) \right)$

$= \epsilon 2 \left(\left(S_R + S'_R \right)^* \left(S_R + S'_R \right) \left(1 + \frac{1}{4}(w^t w) \right) - \frac{1}{4}(w^t w) \left(S_R^* S_R + S'^*_R S'_R \right) \right)$

and the extremum is met for $S'_R = S_R$:

$= \epsilon 2 \left(4 S_R^* S_R \left(1 + \frac{1}{4}(w^t w) \right) - \frac{1}{2}(w^t w) \left(S_R^* S_R \right) \right) = \epsilon 2 \left(4 + \frac{1}{2}w^t w \right) S_R^* S_R \geq 0$ ■

Moreover we have the following :

Theorem 78 *For a given value of the inertial spinor S_0 , and a measured value $S \in \hat{E}_0$ of the spinor S, there is a unique element $\sigma \in Spin(3,1)$ such that $\gamma C(\sigma) S_0 = S$*

Proof. i) The action of $Spin(3,1)$ on E_0, E'_0 is free :

$\forall S_0 \in E_0, E'_0 : \gamma C(s) S_0 = S_0 \Leftrightarrow \sigma = 1$

$S_0 = \begin{bmatrix} v \\ \epsilon i v \end{bmatrix}$

$\gamma C(s) S_0 = S_0 \Leftrightarrow$

$S_R = \left((a + ib) \sigma_0 + \frac{1}{2} \sum_{a=1}^3 (w_a - i r_a) \sigma_a \right) v = v$

$S_L = \left((a - ib) \sigma_0 - \frac{1}{2} \sum_{a=1}^3 (w_a + i r_a) \sigma_a \right) \epsilon i v = \epsilon i v$

$\Rightarrow$

$\left(2a\sigma_0 - i \sum_{a=1}^3 r_a \sigma_a \right) v = 2v$

$\left(2ib\sigma_0 + \sum_{a=1}^3 w_a \sigma_a \right) v = 0$

$\left(\sum_{a=1}^{3} r_a \sigma_a\right) v = 2i(1-a)v$
$\left(\sum_{a=1}^{3} w_a \sigma_a\right) v = -2ibv$
$\Rightarrow$
$\sum_{a=1}^{3} r_a v^* \sigma_a v = 2i(1-a)v^*v$
$\sum_{a=1}^{3} w_a v^* \sigma_a v = -2ibv^*v$
The scalars $v^*\sigma_a v$ are real because the Dirac matrices are Hermitian, as is v^*v ,so
$\Rightarrow b = 0, a = 1$
$\Rightarrow r = w = 0$
and the only solution is $\sigma = 1$.
ii) $\gamma C(\sigma) S_0 = \gamma C(\sigma') S_0 \Rightarrow S_0 = \gamma C(\sigma^{-1}) \gamma C(\sigma') S_0 \Rightarrow \sigma^{-1} \cdot \sigma' = 1$ ■

We can assume that S_0 depends only on the type of particle, then with the knowledge of S_0 the measure of the spinor S defines uniquely the state of the particle with respect to the observer. As $S = \gamma C(\sigma) S_0$ and σ can itself be uniquely, up to spin, decomposed in $\sigma = \sigma_w \cdot \sigma_r$, we have the correspondence with the formulas in the transition between observers where $[\chi] = \exp[K(w)] \exp[J(r)]$: σ_w corresponds to the boost and σ_r corresponds to $\exp[J(r)]$, the vector r is in both cases a Lie algebra representative of an instantaneous spatial rotation.

4.3.3 Momentum and spinors

Momentum

Even If the spinor is deduced from S_0 by an element $\sigma \in Spin(3,1)$, and σ_w is obviously linked to the velocity, *the spinor S does not represent a momentum.* It belongs to a complex vector space E and combines both rotation and translation. Its geometric part σ can be seen as the disposition of a frame which would be attached to the body with respect to the frame of an observer. This is the same difference as between the rotation of a frame of a certain angle, and the motion of rotation at some rotational speed. Notice that the motion is with respect to the standard basis of an observer who would be located at the same point as the particle. It does not involve any fixed origin, and this is not a rotation around a point (like an orbit), but a *rotation at a point.*

Momenta are defined as vectorial quantities (one can add the momenta of different objects) and through derivatives.

Even if the addition of two quantities such as $\gamma C(\sigma_1) S_0 + \gamma C(\sigma_2) S_0 = \gamma C(\sigma_1 + \sigma_2) S_0$ is mathematically valid, because $\sigma_1 + \sigma_2 \in Cl(3,1)$, it has no physical meaning, because $\sigma_1 + \sigma_2 \notin Spin(3,1)$. But the derivative of the spinor reads $\frac{dS}{dt} = \gamma C\left(\frac{d\sigma}{dt}\right) S_0 = \gamma C(\sigma) \gamma C\left(\sigma^{-1} \cdot \frac{d\sigma}{dt}\right) S_0$ and $\sigma^{-1} \cdot \frac{d\sigma}{dt} \in T_1 Spin(3,1)$. And the set :

$M_0 = \{\gamma C(X) S_0, X \in T_1 Spin(3,1), S_0 \in E_0\} = \left\{ \begin{bmatrix} \sigma(\overline{z}) S_R \\ -i\sigma(z) S_R \end{bmatrix}, z \in \mathbb{C}, S_R \in \mathbb{C}^2 \right\}$

is a *real* 6 dimensional vector space. And similarly for M_0' and E_0'.

$(E, \gamma C \circ \mathbf{Ad})$ is a representation of $Spin(3,1)$, so $P_G[M_0, \gamma C \circ \mathbf{Ad}]$ is a real vector bundle.

So $\mathcal{M} = \gamma C\left(\sigma^{-1} \cdot \frac{d\sigma}{dt}\right) S_0$ can be a good candidate to represent a momentum. The instantaneous change in S is measured with respect to the gauge of the particle ($\sigma^{-1} \cdot \frac{d\sigma}{dt}$ is the right logarithmic derivative see Maths.1774).

For any $\mathcal{M} \in M_0, S_0 \in E_0$ there is a unique $v(X,Y) \in T_1 Spin(3,1)$: the Dirac matrices are a basis of the real Lie algebra $su(2)$ so that $\sigma(z)$ identifies uniquely $v(r,w)$.

We have the relation :

$$\mathcal{M} = \gamma C \left(\sigma^{-1} \cdot \frac{d\sigma}{dt} \right) S_0 \Leftrightarrow \frac{dS}{dt} = \gamma C (\sigma) \mathcal{M} \tag{4.35}$$

The exponential is not surjective on $T_1 Spin(3,1)$: to one value of $\upsilon(X,Y) \in T_1 Spin(3,1)$ can be associated 2 opposite values of :

$\sigma = \pm \exp X \exp Y$ (see annex) : they correspond to opposite values of the relativist spin. The momentum is null if $\sigma = Ct$.

To a momentum can be associated a trajectory, as we did above for σ_w.

Theorem 79 *For a given function $f : M \to \mathbb{R}, f(m) > 0$, and a given observer, one can associate a unique, up to sign, element $X = \upsilon(0,w) \in T_1 Spin(3,1)$ to a vector $V \in T_m M$ such that $V = f(m)(c\varepsilon_0(m)) \pm \sum_{a=1}^3 w_a \varepsilon_a(m)$.*

Conversely, given an element $X = \upsilon(r,w) \in T_1 Spin(3,1)$ there are two vectors such that
$V = f(m)(c\varepsilon_0(m)) \pm \sum_{a=1}^3 w_a \varepsilon_a(m)$

Proof. i) Let $V \in T_m M$ and $X = \upsilon(0,w)$ such that $V = f(m)(c\varepsilon_0(m)) + \epsilon \sum_{a=1}^3 w_a \varepsilon_a(m)$

In a change of gauge on P_G :

$\mathbf{p}(m) = \varphi_G(m,1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m) \cdot \chi(m)^{-1}$

$i = 0...3 : \varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i) \to \widetilde{\varepsilon}_i(m) = \mathbf{Ad}_{\chi(m)^{-1}} \varepsilon_i(m)$

$\overrightarrow{\kappa}_a \to \widetilde{\overrightarrow{\kappa}_a} = \mathbf{Ad}_{\chi^{-1}}(\overrightarrow{\kappa}_a)$

$V, f(m)$ do not change but the decomposition changes :

$V = f(m)(c\varepsilon_0(m) + w) = f(m)(c\widetilde{\varepsilon}_0(m) + \widetilde{w})$

so $\widetilde{w} = \frac{1}{f(m)} V - \mathbf{Ad}_{\chi(m)^{-1}} \varepsilon_0(m)$

$X = \upsilon(0,w) \to \widetilde{X} = \sum_{a=4}^6 w_a \widetilde{\overrightarrow{\kappa}_a} = \mathbf{Ad}_{\chi(m)^{-1}}$ so $X \in P_G[T_1 Spin(3,1), \mathbf{Ad}]$

ii) We have the identity :

$\forall X = \upsilon(r,w) \in T_1 Spin(3,1) : \upsilon(r,w) \cdot \varepsilon_0 - \varepsilon_0 \cdot \upsilon(r,w) = [\upsilon(r,w), \varepsilon_0] = \upsilon(0,w)$

thus for a given $X = \upsilon(r,w) \in T_1 Spin(3,1)$ in the orthonormal basis of an observer there are two vectors such that :

$X \cdot \varepsilon_0 - \varepsilon_0 \cdot X = w$

$V = f(m) \left(c\varepsilon_0(m) + \epsilon \sum_{a=1}^3 w_a \varepsilon_a(m) \right), \epsilon = \pm 1$ ■

Notice that the association holds for null vectors.

The definition above :

$\mathcal{M} = \gamma C \left(\sigma^{-1} \cdot \frac{d\sigma}{dt} \right) S_0 \Leftrightarrow \frac{dS}{dt} = \gamma C(\sigma) \mathcal{M}$

assumes that an observer has been chosen (t is the time of an observer). In order to get an observer free definition it is necessary to define the momentum not as a vector but as a one form valued in M_0 :

$$\mathcal{M} = \sum_{\alpha=0}^{3} \gamma C \left(\sigma^{-1} \cdot \partial_\alpha \sigma \right) S_0 \otimes d\xi^\alpha \in \Lambda_1(M; M_0) \tag{4.36}$$

Then along the world line :

$\mathcal{M}(u) = \gamma C \left(\sigma^{-1} \cdot \frac{d\sigma}{d\tau} \right) S_0$

and along the trajectory :

$\mathcal{M}(V) = \gamma C \left(\sigma^{-1} \cdot \frac{d\sigma}{dt} \right) S_0$

In the previous subsection we have defined σ by the property :
$\frac{V}{\sqrt{-\langle V,V\rangle}} = u = \mathbf{Ad}_{\sigma(m)}\varepsilon_0(m) = \sigma\cdot\varepsilon_0\cdot\sigma^{-1}$
By derivating along ε_0 :
$\frac{du}{dt} = \frac{d\sigma}{dt}\cdot\varepsilon_0\cdot\sigma^{-1} + \sigma\cdot\frac{d}{dt}\varepsilon_0\cdot\sigma^{-1} - \sigma\cdot\varepsilon_0\cdot\sigma^{-1}\cdot\frac{d\sigma}{dt}\cdot\sigma^{-1}$
$\frac{d}{dt}\varepsilon_0 = 0$
$\frac{du}{dt} = \sigma\cdot\left(\sigma^{-1}\cdot\frac{d\sigma}{dt}\cdot\varepsilon_0 - \varepsilon_0\cdot\sigma^{-1}\cdot\frac{d\sigma}{dt}\right)\cdot\sigma^{-1} = \mathbf{Ad}_\sigma\left[\sigma^{-1}\cdot\frac{d\sigma}{dt},\varepsilon_0\right] = \mathbf{Ad}_\sigma\left[\upsilon(X,Y),\varepsilon_0\right] = \mathbf{Ad}_\sigma\upsilon(0,Y)$
$\upsilon(0,Y) = \mathbf{Ad}_{\sigma^{-1}}\frac{du}{dt}$
As can be checked σ_r is not involved : the relations hold whatever the value of σ_r.
The usual relativist momentum is $Pc = M_pcu$, so $\upsilon(0,Y)M_pc = \mathbf{Ad}_{\sigma^{-1}}\frac{dP}{dt}$:
$\mathcal{M}(u) = \gamma C\left(\mathbf{Ad}_{\sigma^{-1}}\upsilon\left(X,\frac{dP}{dt}\right)\right)\frac{S_0}{\sqrt{\epsilon\langle S_0,S_0\rangle}}$

We can use the momentum $\mathcal{M}$ in kinematics, represented as a vector $\upsilon(X,Y)$, with variables (X,Y), then Y corresponds to the *derivative* of the usual momentum. Or, with the Newton'law : $\overrightarrow{F} = \frac{d\overrightarrow{p}}{dt}$, $\mathcal{M}$ can be seen as the inertial force.

If we define σ by the components r,w of σ_r,σ_w we have the following formula :

Theorem 80 *For the derivative with respect to any coordinate* α :

$$\partial_\alpha S = \gamma C(\sigma)\gamma C\left(\upsilon\left(\left([C(r)]^t\left(\frac{1}{2}j(w)\partial_\alpha w\right) + [D(r)]^t\partial_\alpha r\right),[C(r)]^t\left(\frac{1}{4a_w}[4-j(w)j(w)]\partial_\alpha w\right)\right)\right)S_0 \tag{4.37}$$

with :
$[C(r)] = \left[1 + a_rj(r) + \frac{1}{2}j(r)j(r)\right]$
$[D(r)] = \left[\frac{1}{a_r} + \frac{1}{2}j(r) + \frac{1}{4a_r}j(r)j(r)\right]$ *and* $[D(r)] = [C(r)][D(r)]^t$

(the same notations $[C(r)],[D(r)]$ are used all along the book)
Proof. $\partial_\alpha S = \partial_\alpha\gamma C(\sigma(m))S_0 = \gamma C(\partial_\alpha\sigma(m))S_0 = \gamma C(\sigma)\gamma C\left(\sigma^{-1}\partial_\alpha\sigma\right)S_0$
$\sigma = \sigma_w\cdot\sigma_r$
$\sigma^{-1}\partial_\alpha\sigma$
$= \sigma_r^{-1}\cdot\sigma_w^{-1}\cdot\partial_\alpha\sigma_w\cdot\sigma_r + \sigma_r^{-1}\cdot\sigma_w^{-1}\cdot\sigma_w\cdot\partial_\alpha\sigma_r$
$= \sigma_r^{-1}\cdot\left(\sigma_w^{-1}\cdot\partial_\alpha\sigma_w + \partial_\alpha\sigma_r\cdot\sigma_r^{-1}\right)\cdot\sigma_r$
$= \mathbf{Ad}_{\sigma_r^{-1}}\left(\sigma_w^{-1}\cdot\partial_\alpha\sigma_w + \partial_\alpha\sigma_r\cdot\sigma_r^{-1}\right)$
$\sigma_w^{-1}\cdot\partial_\alpha\sigma_w = \upsilon(\frac{1}{2}j(w)\partial_\alpha w,\frac{1}{4a_w}(4-j(w)j(w))\partial_\alpha w)$
$\partial_\alpha\sigma_r\cdot\sigma_r^{-1} = \upsilon\left(\left(\frac{1}{a_r} + \frac{1}{2}j(r) + \frac{1}{4a_r}j(r)j(r)\right)\partial_\alpha r,0\right) = \upsilon([D(r)]\partial_\alpha r,0)$
$\sigma_w^{-1}\cdot\partial_\alpha\sigma_w + \partial_\alpha\sigma_r\cdot\sigma_r^{-1}$
$= \upsilon\left([D(r)]\partial_\alpha r + \frac{1}{2}j(w)\partial_\alpha w,\frac{1}{4a_w}(4-j(w)j(w))\partial_\alpha w\right)$
$\left[\mathbf{Ad}_{\sigma_r^{-1}}\right] = [\mathbf{Ad}_{\sigma_r}]^t = \begin{bmatrix}[C(r)]^t & 0\\ 0 & [C(r)]^t\end{bmatrix}$
with $[C(r)]^t = \left[1 - a_rj(r) + \frac{1}{2}j(r)j(r)\right]$
$\sigma^{-1}\partial_\alpha\sigma$
$= \left[\mathbf{Ad}_{\sigma_r^{-1}}\right]\left(\sigma_w^{-1}\cdot\partial_\alpha\sigma_w + \partial_\alpha\sigma_r\cdot\sigma_r^{-1}\right)$
$= \upsilon(X_\alpha,Y_\alpha)$
with :

$X_\alpha = [C(r)]^t \left(\frac{1}{2} j(w)\, \partial_\alpha w + [D(r)]\, \partial_\alpha r\right)$
$Y_\alpha = [C(r)]^t \left(\frac{1}{4a_w} [4 - j(w)\, j(w)]\, \partial_\alpha w\right)$
$[C(r)]^t [D(r)] = \left[1 - a_r j(r) + \frac{1}{2} j(r)\, j(r)\right] \left[\frac{1}{a_r} + \frac{1}{2} j(r) + \frac{1}{4a_r} j(r)\, j(r)\right]$
$= \left[\frac{1}{a_r} - \frac{1}{2} j(r) + \frac{1}{4a_r} j(r)\, j(r)\right] = [D(r)]^t \Leftrightarrow [D(r)] = [C(r)]\, [D(r)]^t$
$X_\alpha = [C(r)]^t \left(\frac{1}{2} j(w)\, \partial_\alpha w\right) + [D(r)]^t \partial_\alpha r$ ■

We have the decomposition :
$\gamma C\left(\sigma^{-1} \cdot \partial_\alpha \sigma\right) = \gamma C\left(\upsilon\left(X_\alpha, 0\right)\right) + \gamma C\left(\upsilon\left(0, Y_\alpha\right)\right)$
thus :
$\mathcal{M} = \mathcal{M}_T + \mathcal{M}_R$
with :
$\mathcal{M}_T = \sum_{\alpha=0}^{3} \gamma C\left(\upsilon\left(X_\alpha, 0\right)\right) S_0 \otimes d\xi^\alpha$
$\mathcal{M}_R = \sum_{\alpha=0}^{3} \gamma C\left(\upsilon\left(0, Y_\alpha\right)\right) S_0 \otimes d\xi^\alpha$
where :
$X_\alpha = [C(r)]^t \left(\frac{1}{2} j(w)\, \partial_\alpha w\right) + [D(r)]^t \partial_\alpha r \simeq [D(r)]^t \partial_\alpha r$
$Y_\alpha = [C(r)]^t \frac{1}{4a_w} [4 - j(w)\, j(w)]\, \partial_\alpha w \simeq [C(r)]^t \partial_\alpha w$
$\partial_\alpha w = \left[a_w + \frac{1}{a_w} j(w)\, j(w)\right] [C(r)]\, Y_\alpha$
$\partial_\alpha r = [D(r)]\, X_\alpha - \frac{1}{2} \left[\frac{4}{a_w} - 3a_w\right] [D(r)]^t j(w)\, [C(r)]\, Y_\alpha$
$\mathcal{M}_T$ is the equivalent of the derivative of a translational momentum or an inertial torque

$\mathcal{M}_R$ is the equivalent of the derivative of a rotational momentum or a translational inertial force.

Kinetic Energy

In the usual relativist context the energy of a particle is defined as the time-component of the vector Pc, where P is the relativist translational momentum. Then the kinetic energy can be defined as the difference between this energy and the energy corresponding to the mass at rest. We have nothing equivalent here : the spinor S incorporates the rotation, is not defined as a vector belonging to $\mathbb{R}^4$ and anyway does not represent a motion. It seems legitimate to come back to Newtonian mechanics : a force applied to a particle is expressed as $\overrightarrow{F} = \frac{d\overrightarrow{p}}{dt}$ where $\overrightarrow{p}$ is the momentum, and the energy spent to change the motion is $K = \int_A^B \left\langle \overrightarrow{F}, \overrightarrow{v} \right\rangle dt \Leftrightarrow \frac{dK}{dt} = \left\langle \frac{d\overrightarrow{p}}{dt}, \overrightarrow{v} \right\rangle = \frac{1}{M_p} \left\langle \frac{d\overrightarrow{p}}{dt}, \overrightarrow{p} \right\rangle$. So, in a system, one can see the variation of the energy which is stored in the motion of the particles as : $\frac{dK}{dt} = \frac{1}{M_p} \left\langle \frac{d\overrightarrow{p}}{dt}, \overrightarrow{p} \right\rangle$. And we have a similar relation for a rotational motion. $\mathcal{M}$ corresponds to the derivative of the momentum, so the variation of the kinetic energy can be represented as:

$\frac{1}{i} \left\langle S, \frac{dS}{dt} \right\rangle = \frac{1}{i} \left\langle \gamma C(\sigma)\, S_0, \gamma C\left(\frac{d\sigma}{dt}\right) S_0 \right\rangle = \frac{1}{i} \left\langle S_0, \gamma C\left(\sigma^{-1} \frac{d\sigma}{dt}\right) S_0 \right\rangle = \frac{1}{i} \left\langle S_0, \mathcal{M}(V) \right\rangle$

Its measure depends on the observer, through the choice of t.

In the following, to keep it simple, we will call variation of the kinetic energy the quantity $\frac{1}{i} \left\langle S, \frac{dS}{dt} \right\rangle$. It is similar to the expression $\psi^* \gamma_0 \frac{\partial \psi}{\partial t}$ of QTF, and a bit far from the usual vision of Newtonian Mechanics.

Theorem 81 *The kinetic energy of a particle can be expressed as :*

$$\frac{1}{i}\left\langle S,\frac{dS}{d\tau}\right\rangle = k^t X \tag{4.38}$$

where $k \in \mathbb{R}^3$ is a fixed vector depending on S_0 and $X = [C(r)]^t \left(\frac{1}{2}j(w)\frac{dw}{dt}\right) + [D(r)]^t \frac{dr}{dt}$

Proof. For the derivative with respect to any coordinate α :

$\frac{1}{i}\langle S,\partial_\alpha S\rangle = \frac{1}{i}\langle \gamma C(\sigma) S_0, \gamma C(\sigma)\gamma C(\upsilon(X_\alpha,Y_\alpha)) S_0\rangle = \frac{1}{i}\langle S_0, \gamma C(\upsilon(X_\alpha,Y_\alpha)) S_0\rangle$

$\upsilon\left(\left([C(r)]^t\left(\frac{1}{2}j(w)\partial_\alpha w\right) + [D(r)]^t\partial_\alpha r\right), [C(r)]^t\left(\frac{1}{4a_w}[4-j(w)j(w)]\partial_\alpha w\right)\right)$

$X_\alpha = [C(r)]^t\left(\frac{1}{2}j(w)\partial_\alpha w\right) + [D(r)]^t\partial_\alpha r$

$Y_\alpha = [C(r)]^t\left(\frac{1}{4a_w}[4-j(w)j(w)]\partial_\alpha w\right)$

$\gamma C(\upsilon(X_\alpha,Y_\alpha)) = \frac{1}{2}\begin{bmatrix}\sigma(Y_\alpha - iX_\alpha) & 0\\ 0 & -\sigma(Y_\alpha + iX_\alpha)\end{bmatrix}$

Let us denote $[S_0] = \begin{bmatrix} S_R \\ \epsilon i S_R\end{bmatrix}$ with $\epsilon = +1$ for particles, and $\epsilon = -1$ for antiparticles.

$\gamma C(\upsilon(X_\alpha,Y_\alpha))[S_0] = \frac{1}{2}\begin{bmatrix}\sigma(Y_\alpha - iX_\alpha)S_R \\ -\epsilon i\sigma(Y_\alpha + iX_\alpha)S_R\end{bmatrix}$

$\gamma_0\gamma C(\upsilon(X_\alpha,Y_\alpha))[S_0] = \frac{1}{2}\begin{bmatrix}-\epsilon\sigma(Y_\alpha + iX_\alpha)S_R \\ i\sigma(Y_\alpha - iX_\alpha)S_R\end{bmatrix}$

$\left\langle S_0, \gamma C\left(\sigma^{-1}\partial_\alpha\sigma_\alpha\right) S_0\right\rangle$

$= \frac{1}{2}\left(S_R^*\left(-\epsilon\sigma(Y_\alpha + iX_\alpha)S_R\right) - \epsilon i S_R^*\left(i\sigma(Y_\alpha - iX_\alpha)S_R\right)\right)$

$= \frac{1}{2}\epsilon\sum_{a=1}^3\left(-(Y_\alpha^a + iX_\alpha^a) + (Y_\alpha^a - iX_\alpha^a)\right)(S_R^*\sigma_a S_R)$

$= -\epsilon i\sum_{a=1}^3 X_\alpha^a(S_R^*\sigma_a S_R)$

$= i\sum_{a=1}^3 X_\alpha^a k^a = ik^t X_\alpha$

where $k^a = -\epsilon(S_R^*\sigma_a S_R)$ are fixed scalars. ■

With : $\langle S_0, S_0\rangle = \epsilon M_p^2 c^2$

$S_R = \frac{M_p c^2}{\sqrt{2}}\begin{bmatrix} e^{i\alpha_1}\cos\alpha_0 \\ e^{i\alpha_2}\sin\alpha_0\end{bmatrix}$

$k = -\epsilon\frac{1}{2}M_p^2c^2 k_0 = -\epsilon\frac{1}{2}M_p^2c^2\begin{bmatrix}(\sin 2\alpha_0)\cos(\alpha_2-\alpha_1)\\ (\sin 2\alpha_0)\sin(\alpha_2-\alpha_1)\\ \cos 2\alpha_0\end{bmatrix}; k_0^t k_0 = 1$

The vector k is similar to the inertia tensor, but here this is a vector, and it encompasses both the translational and the rotational motions. And of course it holds for a particle without internal structure. We will see that k acts also as a magnetic moment.

Then the variation of kinetic energy is :

$$\frac{dK}{dt} = \frac{1}{2}M_p c^2 k_0^t\left([C(r)]^t\left([D(r)]\frac{dr}{dt} + \frac{1}{2}j(w)\frac{dw}{dt}\right)\right)$$

It includes both a translational energy, given by $\frac{1}{2}[C(r)]^t j(w)\frac{dw}{dt}$, and a rotational energy $k^t[C(r)]^t[D(r)]\frac{dr}{dt}$, with a single constant k and K has the dimension of energy.

Planck's constant

The components (w, r) which characterize the spinor are dimensionless. However w is related to the spatial speed, with the universal constant $c : w = \sqrt{\frac{2}{1+\sqrt{1-\frac{\|\vec{v}\|^2}{c^2}}}}\frac{v}{c}$. So it is logical that r would be also related to a measurable physical quantity, through a universal constant. r does not represent a physical rotation, even if its measure is done through rotations of devices, but in the formula above we see that its derivative is related to the kinetic energy. The contribution of the translational motion is $\frac{1}{4}M_p k_0^t \left[C(r)\right]^t j(cw)\frac{dcw}{dt} \sim Mv\frac{dv}{dt}$ and for the rotational motion $\frac{1}{2}M_p c^2 k_0^t \left[C(r)\right]^t \left[D(r)\right]\frac{dr}{dt}$. The matrices $\left[C(r)\right], \left[D(r)\right]$ are dimensionless. $M_p c^2$ has the dimension of energy, so r and the universal constant should have the dimension $Energy \times Time$. Which is just the dimension of the Planck's constant. The SI unit of spin is the joule-second, as the classical angular momentum. In practice, however, for elementary particles it is written as a multiple of the reduced Planck constant $\hbar$.

4.4 SPINOR FIELDS

4.4.1 Definition

The great interest of Spinors is that they sum up the kinematics of a particle in one single, geometric quantity which has a value at any point in a fiber bundle. It is then possible to conceive fields of particles whose world lines are defined by the same vector field, which is an usual case in Physics.

Definition 82 *A **Spinor field** is a section* $\mathbf{S} \in \mathfrak{X}(P_G[E,\gamma C])$ *which represents the kinematics characteristics of a particle or an antiparticle, such that* $\int_\Omega \|S(m)\| \varpi_4 < \infty$

Equivalently a spinor field, denoted $\mathfrak{X}(S_0)$ is defined by a vector $S_0 \in E_0$ and a section $\sigma \in \mathfrak{X}(P_G)$ such that $\mathbf{S}(m) = \gamma C(\sigma(m)) S_0$.

Let $\mathbf{S} \in \mathfrak{X}(P_G[E,\gamma C])$. Then $\langle S(m), S(m)\rangle = i(S_L^* S_R - S_R^* S_L) = y(m)$ defines a function on M. If $\mathbf{S}$ represents the state of a particle, then $\langle S(m), S(m)\rangle = \langle S_0, S_0\rangle$. A necessary condition for a section of $P_G[E,\gamma C]$ to represent the state of a particle is that $\langle S(m), S(m)\rangle = y$ has a fixed, positive, value. Then the set $E(y)$ of vectors $S_0 \in E_0$ such that $\exists \sigma \in \mathfrak{X}(P_G) : S = \gamma C(\sigma) S_0$ is given by :

$E(y) = \left\{ S_0 = \begin{bmatrix} v \\ iv \end{bmatrix}, v^* v = \frac{1}{2} y, v \in \mathbb{C}^2 \right\}$

And for a given vector $S_0 \in E(y)$, at each point m there is a unique $\sigma(m) \in Spin(3,1)$ such that : $S(m) = \gamma C(\sigma(m)) S_0$. It defines a section $\sigma \in \mathfrak{X}(P_G)$, and at each point m, for each value ± 1 of the relativist spin and for a given observer, a vector field which is the tangent to the world line, and a spatial spin.

And similarly for antiparticles.

A spinor field represents particles which have the same inertial behavior. If, in a model, we have several particles, interacting with each others or with force fields, each particle can be assigned to a spinor field, which represents a general solution of the problem. One can also associate a density to spinor fields. All these topics will be seen in more details in the next chapters.

For elementary particles the vector S_0 is one of its fundamental characteristic. For other material bodies S_0 is a kinematic characteristic which, for deformable solids, can be computed.

A Spinor field is defined without any reference to an observer : it has an intrinsic meaning, as it was expected. And the decomposition in translational momentum on one hand, and rotational momentum on the other hand, is relative to each observer. So we have here a new, significant, feature of the relativist momenta.

4.4.2 More on the theory of the representations of groups

Functional Representations

Functional representations are representations on vector spaces of functions or maps. Any locally compact topological group has at least one unitary faithful representation (usually infinite dimensional) of this kind, and they are common in Physics. The principles are the following (Maths.23.2.2).

Let H be a Banach vector space of maps $\varphi : E \to F$ from a topological space E to a vector space F, G a topological group with a continuous left action λ on E : $\lambda : G \times E \to E :: \lambda(g,x)$ such that $\lambda(g \cdot g', x) = \lambda(g, \lambda(g', x)), \lambda(1, x) = x$

Define the left action Λ of G on H : $\Lambda : G \times H \to H :: \Lambda(g,\varphi)(x) = \varphi(\lambda(g^{-1}, x))$

Then (H, Λ) is a representation of G. Thus G acts *on the argument* of φ.

If H is a Hilbert space and G has a Haar measure μ (a measure on G, all the groups that we will encounter have one Maths.22.5) then the representation is unitary with the scalar product :

$\langle \varphi_1, \varphi_2 \rangle = \int_G \langle \Lambda(g, \varphi_1), \Lambda(g, \varphi_2) \rangle_H \mu(g)$

If G is a Lie group and the maps of H and λ are differentiable then $\left(H, \Lambda'_g(1, .)\right)$ is a representation of the Lie algebra $T_1 G$ where $X \in T_1 G$ acts by a differential operator :

$\Lambda'_g(1, \varphi)(X)(x) = -\varphi'(x) \lambda'_g(1, x) X = \frac{d}{dt} \varphi(\lambda(\exp(-tX), x))|_{t=0}$

For a right action $\rho : E \times G \to E :: \rho(g, x)$we have similar results, with

$P : H \times G \to H :: P(\varphi, g)(x) = \varphi(\rho(x, g))$

$P'_g(\varphi, 1)(X)(x) = -\varphi'(x) \rho'_g(x, 1) X = \frac{d}{dt} \varphi(\rho(x, \exp(-tX)))|_{t=0}$

H can be a vector space of sections on a vector bundle. In a functional representation each function is a vector of the representation, so it is usually infinite dimensional. However the representation can be finite dimensional, by taking polynomials as functions, but this is not always possible : the set of polynomials must be algebraically closed under the action of the group.

Isomorphisms of groups

Most of the groups that are encountered in Physics are related to the group $SL(\mathbb{C}, 2)$ of 2×2 complex matrices with determinant 1 (Maths.24).

Any matrix of the Lie algebra $sl(\mathbb{C}, 2)$ reads with $Z = (z_1, z_2, z_3) \in \mathbb{C}^3$

$f(Z) = \begin{bmatrix} iz_3 & z_2 + iz_1 \\ -z_2 + iz_1 & -iz_3 \end{bmatrix} = i\sigma(z) \Rightarrow Trf(Z) = 0$

which is equivalent to take as basis the Dirac matrices.

The exponential is not surjective on $sl(\mathbb{C}, 2)$ and any matrix of $SL(\mathbb{C}, 2)$ reads :

$I \cosh D + \frac{\sinh D}{D} f(Z)$ with $D^2 = -\det f(Z) = -\left(z_1^2 + z_2^2 + z_3^2\right)$

The group $SU(2)$ of 2×2 unitary matrices ($NN^* = I$) is a compact real subgroup of $SL(\mathbb{C}, 2)$. Its Lie algebra is comprised of matrices $f(r)$ with $r \in \mathbb{R}^3$. The exponential is surjective on $SU(2) : \exp f(r) = I \cos \sqrt{r^t r} + \frac{\sin \sqrt{r^t r}}{\sqrt{r^t r}} f(r)$

$T_1 Spin(3, 1)$ is isomorphic to $sl(\mathbb{C}, 2)$ (Math.1959) : $\upsilon(r, w) \to f(r + iw)$

$Spin(3, 1)$ is isomorphic to $SL(\mathbb{C}, 2) : a + \upsilon(r, w) + b\varepsilon_5 \to \exp f(r + iw)$

$T_1 Spin(3)$ is isomorphic to $su(2) : \upsilon(r, 0) \to f(r)$ and $so(3) : \upsilon(r, 0) \to j(r)$

$Spin(3)$ is isomorphic to $SU(2)$:

$a_r + \upsilon(r, 0) \to \exp f(r) = I \cos \sqrt{r^t r} + \frac{\sin \sqrt{r^t r}}{\sqrt{r^t r}} f(r)$

Representations of Spin(3,1), Spin(3) and SO(3)

$SL(\mathbb{C}, 2)$ and $Spin(3, 1)$ have the same representations which are (up to equivalence) :

- a unique, *non unitary*, irreducible representation of dimension n (Maths.1953), which can be seen as the tensorial product of two finite dimensional representations $\left(P^j \otimes P^k, D_j \times D_k\right)$ of $SU(2) \times SU(2)$ (see below).

- the only unitary representations are over spaces of complex functions : they are infinite dimensional and each irreducible representation is parametrized by 2 scalars $z \in \mathbb{Z}, k \in \mathbb{R}$ (Maths.1955).

$SU(2)$ as $Spin(3)$ are compact groups, so their unitary representations are reducible (Math.1960) in a sum of orthogonal, finite dimensional, unitary representations. The only irreducible, finite dimensional, unitary, representations, denoted $\left(P^j, D^j\right)$ are on the space P^j of degree $2j$ homogeneous polynomials with 2 complex variables z_1, z_2, where conventionally j is an integer or half an integer. P^j is $2j+1$ dimensional and the elements of an orthonormal basis are denoted :

$|j,m\rangle = \frac{1}{\sqrt{(j-m)!(j+m)!}} z_1^{j+m} z_2^{j-m}$ with $-j \leq m \leq +j$. And D^j is defined by :

$g \in U(2) : D^j(g) P\left(\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right) = P\left([g]^{-1}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right)$

Thus the functions read : $\varphi(z_1, z_2) = \sum_{j\in\frac{1}{2}\mathbb{Z}} \sum_{m=-j}^{m=+j} \varphi^{jm} |j,m\rangle$ with complex constants φ^{jm}

It induces a representation $\left(P^j, d^j\right)$ of the Lie algebras where d^j is a differential operator acting on the polynomials P :

$X \in su(2) : d^j(X)(P)(z_1, z_2) = \frac{d}{dt} P\left([\exp(-tX)]\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right)|_{t=0}$

which gives, for polynomials, another polynomial.

$d^j(X)$ is a linear map on P^j, which is also linear with respect to X, thus it is convenient to define d^j by the action $d^j(\kappa_a)$ of a basis $(\kappa_a)_{a=1}^3$ of the Lie algebra and the three operators are denoted L_x, L_y, L_z.They are expressed in the orthonormal basis $|j,m\rangle$ by square $2j+1$ matrices (depending on the conventions to represent the Lie algebra). The usage is to denote $L_z|j,m\rangle = m|j,m\rangle$.

The only irreducible, unitary, representations of $SO(3)$ are given by $\left(P^j, D^j\right)$ with j integer.

Casimir element

The **universal enveloping algebra** U of a Lie algebra is actually a vector space, built from *tensorial powers* of the Lie algebra, and whose basis is given by ordered products of elements of the basis $(\kappa_i)_{i\in I}$ of the Lie algebra (Maths.1692). Universal enveloping algebras are necessary when interacting systems are considered (such as in Chemistry), because their representation involve the tensorial product of the variables.

Any representation (E, f) of the Lie algebra can be extended to a representation (E, F) of its universal enveloping algebra (Maths.1891) where the action is :

$F\left(\kappa_{i_1}^{n_1}...\kappa_{i_p}^{n_p}\right) = f(\kappa_{i_1})^{n_1} \circ ... \circ f\left(\kappa_{i_p}\right)^{n_p}$

When the representation (E, f) comes from a functional representation, in the induced representation on U the action of F is represented by differential operators, of the same order than $n_1 + n_2 + ... + n_p$.

In the representation of $T_1Spin(3,1)$ by matrices of $so(3,1)$ the universal enveloping algebra is actually an algebra of matrices.

The **Casimir element** is a special element Ω of U, defined through the Killing form (Maths.1698). In an irreducible representation (E, f) of a semi simple Lie algebra, as $Spin(3,1)$, the image of the Casimir element acts by a non zero fixed scalar $F(\Omega)u = ku$.In functional representations it acts by a differential operator of second order : $F(\Omega)\varphi(x) = D_2\varphi(x) = k\varphi(x) : \varphi$ is an eigen vector of D_2. As a consequence, if there is a scalar product on $E : \langle F(\Omega)u, F(\Omega)u\rangle = \langle ku, ku\rangle = k^2\langle u,u\rangle$.

If $(E_1, f_1), (E_2, f_2)$ are two equivalent representations of the same algebra A :

$\exists\phi : E_1 \to E_2$ such that :

$\forall\kappa \in A : f_1(\kappa) = \phi^{-1} \circ f_2(\kappa) \circ \phi$

$F_1\left(\kappa_{i_1}^{n_1}...\kappa_{i_p}^{n_p}\right) = f_1(\kappa_{i_1})^{n_1} \circ ... \circ f_1\left(\kappa_{i_p}\right)^{n_p}$

$= \left(\phi^{-1} \circ f_2(\kappa_{i_1}) \circ \phi\right)^{n_1} \circ ... \circ \left(\phi^{-1} \circ f_2\left(\kappa_{i_p}\right) \circ \phi\right)^{n_p} = \phi^{-n_1-..n_p} \circ F_2\left(\kappa_{i_1}^{n_1}...\kappa_{i_p}^{n_p}\right) \circ \phi^{n_1+..n_p}$

$F_1(\Omega) = \phi^{-n_1-..n_p} \circ F_2\left(\kappa_{i_1}^{n_1}...\kappa_{i_p}^{n_p}\right) \circ \phi^{n_1+..n_p}(u) = \phi^{-n_1-..n_p} \circ \left(k_2\phi^{n_1+..n_p}(u)\right) = k_2u = k_1u$

Thus the Casimir element acts with the same scalar in all equivalent representations.

The Killing form on $T_1Spin(3,1)$ is :
$B(\upsilon(r,w),\upsilon(r',w')) = 4(w^tw' - r^tr')$
thus the elements
$\kappa_1 = -\frac{1}{8}\varepsilon_3\cdot\varepsilon_2, \kappa_2 = -\frac{1}{8}\varepsilon_1\cdot\varepsilon_3, \kappa_3 = -\frac{1}{8}\varepsilon_2\cdot\varepsilon_1,$
$\kappa_4 = \frac{1}{8}\varepsilon_0\cdot\varepsilon_1, \kappa_5 = \frac{1}{8}\varepsilon_0\cdot\varepsilon_2, \kappa_6 = \frac{1}{8}\varepsilon_0\cdot\varepsilon_3$
constitute an orthonormal basis for B and the Casimir element of $U(T_1Spin(3,1))$ is :
$\Omega = \sum_{i=4}^6(\kappa_i)^2 - \sum_{i=1}^3(\kappa_i)^2$
The action of the Casimir element in the representation $(E,\gamma C)$ of $Spin(3,1)$ is :
$F_E(\Omega)u = \left(\sum_{i=4}^6(\gamma C(\kappa_i))^2 - \sum_{i=1}^3(\gamma C(\kappa_i))^2\right)u = \frac{3}{2}u$
In the representation (P^j,d^j) of $T_1Spin(3)$, if we denote $L_x = f(\kappa_1), L_y = f(\kappa_2), L_z = f(\kappa_3)$ with 3 arbitrary orthogonal axes :
$F(\Omega)|j,m\rangle = L^2|j,m\rangle = (L_x^2 + L_y^2 + L_z^2)|j,m\rangle = j(j+1)|j,m\rangle$
$d^j(\kappa_i)\left(\sum_{m=-j}^{m=+j}X^m|j,m>\right) = \sum_{m=-j}^{m=+j}X^m d^j(\kappa_i)|j,m>$

4.4.3 The Spin of a particle

Definition

The space $\mathfrak{X}(P_G[E,\gamma C])$ of sections is a functional representation of $Spin(3,1)$ with the global action γC and the argument σ. The subspace $\mathfrak{X}(S_0)$ is invariant by the right or left global actions of $Spin(3,1)$: $\gamma C(\sigma(m))S_0 \to \gamma C(s)\gamma C(\sigma(m))S_0$ or $\gamma C(\sigma(m))S_0 \to \gamma C(\sigma((m)))\gamma C(s)S_0$. In particular it is invariant by the action of $Spin(3)$:
$\rho : \mathfrak{X}(S_0)\times Spin(3) \to \rho(S(m),s_r) = \gamma C(\sigma(m)\cdot s_r)S_0$
Moreover the value of $Y(m) = \langle S(m), S(m)\rangle$ is invariant by $Spin(3,1)$.

The spinor fields $S\in\mathfrak{X}(S_0)$ can equivalently be defined by a couple (S_0,σ) where $\sigma\in\mathfrak{X}(P_G)$. For a given observer each $\sigma(m)$ has two decompositions : $\sigma(m) = \epsilon\sigma_w(m)\cdot\epsilon\sigma_r(m)$ so the couple (S_0,σ) defines precisely two Spatial Spinor fields : $S_r(m) = \gamma C(\sigma_r(m))S_0$.

Conversely one can define Spatial Spinor Fields by a couple : (S_0,σ_r) where $\sigma_r\in\mathfrak{X}(P_R)$ and they constitute a set $\mathfrak{X}_r(S_0)$ which is invariant by $Spin(3)$ (but not by $Spin(3,1)$).

Let us denote : $\pi_\epsilon : \mathfrak{X}(S_0) \to \mathfrak{X}_r(S_0)$ the maps which associates, for a given observer, to each Spinor field the Spatial Spinor field with $\epsilon = \pm 1$

On the set : $\mathfrak{X}(S_0)$ we can define the equivalence relation :
$S \sim S' \Leftrightarrow \pi_\epsilon(S) = \pi_\epsilon(S')$

Each class of equivalence is the set of spinor fields which have, for the observer, the same kinematic behavior with regard to a rotation by $Spin(3)$. The value of $\pi_\epsilon(S)$ for a given spinor field is **the Spin of the particle**, in its usual meaning. So to any given spinor field corresponds, for an observer, two Spins, with the Spin up or down. And conversely for a given Spin there can be infinitely many spinor fields, defined by a section of the associated bundle $\sigma_w\in\mathfrak{X}(P_W)$.

The projection π_ϵ depends on the choice of a vector ε_0, so the Spin depends on the observer. The spin can be seen as a rotation, and its derivative as a rotational momentum. The spin is defined by an element $\sigma_r\in Spin(3)$ which reads : $\sigma_r = a_r + \upsilon(r,0)$ and $\frac{dr}{dt}\in\mathbb{R}^3$ can represent an instantaneous rotation. However what is characteristic of the spin is not r but $\upsilon(r,0)$, and we have seen that $\upsilon(r,0)$ does not depend on the choice of a spatial basis. So we have the known paradox : we have a quantity, the spin, which looks like a rotation, which can be measured as a rotation, but is not related to a precise basis, even if its measure is done in one ! The explanation is of course that the spin does not correspond to a rotation in the usual meaning, and we retrieve the distinction between a geometric rotation and a kinematic rotation.

Similarly we have the projection : $\pi_w : Spin(3,1) \to Spin(3,1)/Spin(3)$ and we can define the equivalence relation in $\mathfrak{X}(S_0)$:

$S = \gamma C(\sigma) S_0 \sim S' = \gamma C(\sigma') S_0 \Leftrightarrow \pi_w(\sigma) = \pi_w(\sigma')$

The class of equivalence represents the particles which have the same trajectories. This is also a geometric quantity, invariant by $Spin(3,1)$, but observer dependant. Its derivative is the translational momentum.

Quantization of the Spinor

Theorem 83 *The set $L^1(S_0) = L^1(M, P_G[E,\gamma C], \varpi_4) \cap \mathfrak{X}(S_0)$ of integrable spinor fields associated to a particle is characterized by 2 scalars : $k \in \mathbb{R}, z \in \mathbb{Z}$.*

The Spin, up or down, associated to each section by an observer is characterized by a scalar $j \in \frac{1}{2}\mathbb{N}$ and belongs to a 2j+1 dimensional vector space : $S_r(m) = \sum_{p=-j}^{+j} S_r^p |j,p>$ with the constant components S_r^p and an orthonormal basis $|j,p>$

Proof. i) The space $L^1(M, P_G[E,\gamma C], \varpi_4)$ is a Fréchet space. The Theorem 2 applies and there are a Hilbert space H and an isometry $\Upsilon : L^1 \to H :: \psi = \Upsilon(S)$. This Hilbert space is $L^2\left(M; \widetilde{E}_0; \varpi_4\right)$ with $\widetilde{E}_0 = Span_{\mathbb{C}}(E_0)$ and the scalar product on E.

Moreover $\left(L^1, \gamma C\right)$ is an infinite dimensional representation of $Spin(3,1)$ (the scalar product, thus the norm, is invariant by $Spin(3,1)$, and L^1 is invariant by $Spin(3,1)$). We can apply the theorem 22. $(H, \widehat{\gamma})$ is a unitary representation of $Spin(3,1)$ with $\widehat{\gamma}(\sigma) = \Upsilon \circ \gamma C(\sigma) \circ \Upsilon^{-1}$.

ii) Consider the function : $Y : L^1 \to \mathbb{C} :: Y = \langle S, S\rangle$. For a given section, Y has a value at each point of M and Y is invariant by $Spin(3,1)$. We can implement the theorem 24: to each value y of Y is associated a class of equivalence in L^1 and in H.

If we fix $Y = \langle S_0, S_0\rangle = Ct$ we have two subsets

$L^1(S_0) = L^1 \cap \mathfrak{X}(S_0)$ in L^1 and $H(S_0)$ in H.

$H(S_0)$ is invariant by $\widehat{\gamma}$

iii) The unitary representations of $Spin(3,1)$ read : $H = \oplus_{z,k} H_{z,k} \oplus H_c$ where $H_{z,k}$ are unitary irreducible representations, defined by the parameters $z \in \mathbb{Z}, k \in \mathbb{R}$, and H_c does not contain any irreducible representation, so H_c is not invariant under the action of $Spin(3,1)$ (Maths.1914).

As a consequence $H(S_0)$ is isomorphic to a subset of one of the irreducible representations $H_{z,k}$ and the spinor field is characterized by two scalars $k \in \mathbb{R}, z \in \mathbb{Z}$ linked to S_0.

iv) In L^1, for each section S and a given observer, the Spatial Spinor S_r is a representation of $Spin(3)$. Moreover for S_0, ϵ fixed it belongs to one of the irreducible representations of $Spin(3)$. It is isomorphic to one of the representations $\left(P^j, D^j\right)$ with $j \in \frac{1}{2}\mathbb{N}$. These representations are finite dimensional, so S_r belongs to a $2j+1$ dimensional vector space : $S_r(m) = \sum_{p=-j}^{+j} S_r^p |j,p>$ with the constant components S_r^p. ■

Assume that we study a system comprising of unknown particles $p = 1...N$. The modeling of their kinematic characteristics leads naturally to assume that these particles belong to some spinor fields : $S_p \in \mathfrak{X}(P_G[E,\gamma C])$ with different, unknown, inertial spinor.S_0. Because no value of S_0 is imposed we have a vector space and we can implement the theorem 2.

What the theorem above tells us is that the solutions must be found in maps : $S_p : \Omega \to E$ which can be sorted out :

- by the value of $\langle S_0, S_0\rangle$, that is their mass

- by the value of some integer $z \in \mathbb{Z}$
- and their spin by a half integer $j \in \frac{1}{2}\mathbb{N}$

They correspond to particles which have the same behavior when submitted to a force field (they have the same world lines and spatial spinor for any observer). In other words the spinor is not the only characteristic which determines the behavior of a particle, and these others characteristics can be labeled by a signed integer. This is the starting point to the representation of charged particles.

For elementary particles it is experimentally seen that $j = \frac{1}{2}$,and this is the origin of the name "particles of spin $\frac{1}{2}$". For composite particles or nuclei the spin can be higher.

Even if the set $\mathfrak{X}(S_0)$ is not a vector space, it is a manifold which is embedded in a vector space, so that each of its points (a map S_p) can be written as a fixed linear combination of vectors of a basis. The vector space is always infinite dimensional for the translational momentum, but each spin belongs to a finite dimensional vector space, which is isomorphic to some $\left(P^j, D^j\right)$: $S_r(m) = \sum_{p=-j}^{+j} S_r^p |j,p>$ where S_r^p are fixed scalars and $|j,p>$ are, for a given system, fixed maps $|j,p>: \Omega \to E_0$, images of vectors of the basis of P^j by some isometry. Each vector $|j,p>$ is assimilated to a state of the particle, and j,p are the **quantum numbers** labeling the state. The maps $|j,p>$ are not polynomials (as in P^j), they are used only to define the algebraic structure of the space $H(S_0)$, However they have an interpretation for models of atoms (see below). Under the action of $Spin(3)$ the vectors $S_r(m)$ transform according to the same matrices as in D^j :

$\gamma C(\sigma(m) \cdot s_r) S_0 = \sum_{p=-j}^{+j} S_r^p \left[D^j(f(s_r))\right] |j,p>$ where $f(s_r)$ is the image of σ in $SU(2)$.

By itself the theorem does not provide a solution : a vector of a basis of the vector spaces is itself some map $E_i : \Omega \to E$. But it shows that the solution cannot take any value, even before we implement any physical law relating the fields and the kinematic characteristics. In a given system the solutions that appear follow the same pattern, whatever the initial conditions, or the value of the other variables (notably the fields). In most applications, notably in atomic and molecular physics, there is a special interest for stationary solutions, for which the Spinor does not change with time, or follows a cyclic map. They are special solutions of the PDE which represent the interactions in a system. Then the combination stationary solution of PDE + belonging to a specific vector space quite often restrict to a finite number the possible states, whatever the initial conditions. The states are "quantized". This is specially important when the initial conditions are not known.

There is one important difference in the behavior of the spin, according to the value of j. The Spin is invariant by a rotation by $Spin(3)$, and the scalars $\pm 1 \in Spin(3)$. The actions of $+s$ and $-s$ give opposite results. $Spin(3)$ is the double cover of $SO(3)$: to the same element g of $SO(3)$ are associated two elements $\pm s$ of $Spin(3)$. The representations $\left(P^j, D^j\right)$ with $j \in \mathbb{N}$ are also representations of $SO(3)$. It implies that the vector spaces are invariant by $\pm s$. The fact that j is an integer means that *the particle has a physical specific symmetry* : the rotations $\pm s$ give the same result. And equivalently, if j is half an integer the rotations by $\pm s$ give opposite results.

Measure of the spatial spin of a particle

A particle has, whatever the scale, by definition, no internal structure, so it is impossible to observe its geometric rotation. However it has a spin, its spatial spinor S_r is a variable which can be represented in a finite dimensional space : S_r is an observable. The measure of the derivative of the spatial spinor, similar to a rotational momentum, is done by observing the behavior of the particle when it is submitted to a force field which acts differently according to the value of the spinor. This is similar to the measure of the rotation of a perfectly symmetric ball by observing its trajectory when it is submitted to a dissymmetric initial impulsion (golfers will understand).

Most particles have a magnetic moment, linked to their spinor (more precisely to the vector k). So the usual way to measure the latter is to submit the particles to a non homogeneous magnetic field. This is the Stern-Gerlach analyzer described in all handbooks, where particles have different trajectories according to their magnetic moment. MRI uses a method based on the same principle with oscillating fields whose variation is measured. The process can be modelled as follows.

The spinors of the particles are represented by some section $\mathcal{S} \in \mathfrak{X}(S_0)$. The device operates only on the spin : $S_r(m) = \gamma C(\sigma_r(m)) S_0$ and is parametrized by a spatial rotation $s_r \in Spin(3)$, and usually by a vector $\rho \in \mathbb{R}^3$, corresponding to a rotation s_r.

The first effect is a breakdown of symmetry : s_r has not the same impact for the particles with spin up or down. This manifests by two separate beams in the Stern-Gerlach experiment.

An observable $\Phi(S_r)$ of S_r is a projection on some finite dimensional vector space of maps. Because of the quantization, this vector spaces has for vectors $|j,p>$ which are fixed maps, image of the vectors of basis of P^j which are eigen vectors of the observable. The action of the device can be modelled as an operator $L(\rho)$ acting on this space, and the matrices to go from one orientation ρ_1 to another ρ_2 are the same as in (P^j, d_j) . It reads :

$L(\rho)\Phi(S_r) = \sum_{p=-j}^{+j} S_r^p [d_j(\rho)] |j,p>$

For a given beam we have a breakdown of the measures, corresponding to each of the states labelled by p.

Arbitrary axes x, y, z are chosen for the device, which provide 3 measures $L_x(S_r), L_y(S_r), L_z(S_r)$, such that $L_z(S_r)|j,m\rangle = m|j,m\rangle$.

The Casimir operator Ω is such that $L^2\Phi(S_r) = \left(L_x^2 + L_y^2 + L_z^2\right)(S_r) = j(j+1)\Phi(S_r)$

Atoms and electrons

QM has been developed from the study of atoms, with a basic model (Bohr's atom) in which electrons move around the nucleus. Even if this idea still holds, and this is how atoms are commonly viewed, it had been quickly obvious that a classic model does not work. However using what has been developed previously, we can have another representation.

Let us consider a system comprised of one electron moving around a nucleus. If we consider the atom as a particle, that is without considering its internal structure, it can be represented by a spinor S, and its rotational momentum by a the derivative of a spin S_r. The previous results hold and the spin can be represented in a finite dimensional vector space isomorphic to P^j. However j, which belongs to $\frac{1}{2}\mathbb{N}$, is not necessarily equal to $\frac{1}{2}$.

The polynomials P^j have no physical meaning. However in this case it is usual to provide one. By a purely mathematical computation it is possible to show that the representation (P^j, D^j) is equivalent to a representation on square integrable functions $f(x)$ on $\mathbb{R}^3$, and from there on spherical harmonic polynomials (Maths.1958). It is then assumed that the arguments of the function $f(x)$ are related to the coordinates (in an euclidean frame) of the electron. This is a legacy of the first models of atoms. Actually there is no need for such an assumption to build a consistent model, which would be useless in the GR context, and the image of electrons rotating around a nucleus and spinning has no physical support.

For atoms with several electrons, the model must involve the tensorial products of each spinor. The previous representations of $SU(2)$ are then extended to the tensorial products of P^j, and their derivative to representations of the universal enveloping algebra. It is often possible to rearrange these representations, by combinations using Clebsch-Jordan coefficients (Maths.1960), and in this endeavour the spherical harmonic polynomials are useful because they provide many identities. This is one major application of QM in Chemistry. The same kind of model is used for composite particles in Quantum Theory of Fields.

4.4.4 Material bodies and spinors

Representation of a material body by sections of P_G

A material body B can be defined, from a geometric point of view, by a vector field u whose integral curves are the world lines of its particles. Then the flow $\Phi_u(\tau, a)$ defines the body B itself at each proper time τ as a compact subset $\omega(\tau)$ of a 3 dimensional hypersurface. And there are privileged observers B for whom $\omega(0) \subset \Omega(0)$.

So a material body can be defined with respect to these observers, up to a constant ± 1 by a section σ_w of P_W or, up to a spatial spinor, by a section of P_G and a compact, space like hypersurface $\omega(0)$. Then σ_w provides u, Φ_u and $\omega(0)$ defines $\omega(\tau)$. The section σ_w can be seen as the general definition of B, which can be fitted to any initial conditions $\omega(0)$. This is the most efficient way to define geometrically a material body in physical models.

Spinors representing a solid

The usual concepts of motion of a body over itself (usually a rotation of the body) cannot be easily represented in relativist geometry. This is the main motivation for the introduction of spinors, and any material body whose internal structure can be neglected (at the scale of the study) can be represented, from the kinematic point of view, by a spinor which accounts for its rotation (through the spin) as said above for atoms. If the location of the material body can be represented by a geometric point, then the kinematic representation of B is given by a map : $S_B : \mathbb{R} \to P_G[E, \gamma C]$, such that $S_B(t) = \gamma C(\sigma(t)) S_0$. We do not need more : S_B provides everything, including the rotational momentum. Thus, even if no internal structure or rotation of the body is assumed, eventually it can be accounted for. This is similar to the replacement in Classic Mechanics of a solid by its center of mass.

However this representation assumes that S_0 is known. As in Classic Mechanics for the inertial tensor, the computation of the inertial spinor S_0 is, for a given solid, a separate issue. It can be done through the aggregation of material points (particles) with a specific law giving the shape and the density of the body. And the inertial spinor is not necessarily constant : we can consider deformable solids. Actually we can define *a rigid solid as a material body such that S_0 is constant.*

Proposition 84 *A deformable solid body can be represented by a map :*
$S : \mathbb{R} \to P_G[E, \gamma C]$ *such that* $\langle S(\tau), S(\tau)\rangle > 0$ *or* $\langle S(\tau), S(\tau)\rangle < 0$
A rigid solid body can be represented by a map :
$S : \mathbb{R} \to P_G[E, \gamma C] :: S(\tau) = \gamma C(\sigma(\tau)) S_0$ *for a fixed* $S_0 \in E_0$ *or* $S_0 \in E_0'$
where τ is the proper time of the body

To assume that the material points behave in a coherent way in a solid assumes that there are forces which assure this cohesion. And indeed a material body can be deformed or broken. So we can say that the fact, assumed and which can be checked, that a material body can be represented by a unique spinor incorporates the existence of these internal forces. And ultimately the break of a material body can result in several spinors. So in modelling the evolution of a material body we should include additional assumptions about the laws (which are similar to the phenomenological laws for deformable solids) for the change of S_0.And in a discontinuous process add the laws which rules the splitting in different spinors.

Aggregating matter fields

With these definitions we can consider the task to compute the spinor that we will denote S_B, for a deformable solid, by aggregating material points. This is similar to the computation of the inertial tensor in Classic Mechanics : this is a specific endeavour, done in a separate model, using specific assumptions (about the shape, density, motion of the particles) and the result is then used in a more general model (for instance to compute the motion of different bodies). The single spinor corresponding to the whole body is assigned, in the more general model, to any point : all the material points have then the same location.

The first issue is the definition of the motion of the material points with respect to the body. We need a chart to do it, which is given by an observer B, such that at his proper time $t = 0$ the set $\omega(0)$ is in his present $\Omega(0)$. Then at any given time t the set of particles constituting the solid stays in his present. B uses his standard chart :

$\varphi_B\left(t,\eta^1,\eta^2,\eta^3\right) = \Phi_{\varepsilon_0}\left(t,x\left(\eta^1,\eta^2,\eta^3\right)\right)$ where $x\left(\eta^1,\eta^2,\eta^3\right)$ is a chart on $\omega(0)$ and ε_0 his time like vector field. The chart is arbitrary and fixed.

The particles follow the trajectories given by a vector field V and their location at t is $\Phi_V(t,a) = \Phi_{\varepsilon_0}(t,x(t)) = \varphi_B(t,x(t))$ with $x(0) = a$.

$\omega(t) = \{\Phi_V(t,a), a \in \omega(0)\} = \{\Phi_{\varepsilon_0}(t,x(t)), x(0) \in \omega(0)\}$ represents the location of the body at t and $\Omega = \{\Omega(t), t \in [0,T]\}$

The material points are represented by a section $S(t,x) = \gamma C(\sigma(t,x)) S_0 \in \mathfrak{X}(S_0)$. The choice of S_0 can be arbitrary. S is a geometric quantity which does not depend on a chart, however σ_w provides a vector field of world lines u for the material points with respect to ε_0.

We assume that the observer defines a tetrad $(\varepsilon_i(m))$ from which the metric and the volume form are deduced in the usual way.

The density $\mu(t,x)$ is defined over Ω with respect to the volume form ϖ_4. Because $\omega(t) \subset \Omega(t)$ the unitary, future oriented, normal to $\omega(t)$ is $\varepsilon_0 = \partial\xi_0$ and μ induces a density $\mu_3(t,x)$ over $\omega(t)$ with respect to the volume form ϖ_3 :

$\mu_3(t,x)\varpi_3(t,x) = i_V(\mu(t,x)\varpi_4(t,x))$ which is the flux of matter going through $\omega(t)$.

As noticed before the holonomic basis $e_i(m) = (\mathbf{p}(m), \varepsilon_i)$ of $P_G[E,\gamma C]$ is arbitrary, in that there is no physical reference for the choice of the vectors e_i. It depends only on the observer and we can assume that he keeps the same basis all over the solid, which is legitimate if Ω is not extremely large.

$S(t,x) = \gamma C(\sigma(t,x)) S_0$

Then the integral :

$\int_{\omega(t)} \gamma C(\sigma(t,x)) \mu_3(t,x) S_0 \varpi_3(t,x)$

$= \left[\gamma C\left(\int_{\omega(t)} (\sigma(t,x)) \mu_3(\tau,x) \varpi_3(t,x)\right)\right] S_0$

is well defined on the fixed vector space E.

$\sigma(t,x) = a(t,x) + \upsilon(r(t,x), w(t,x)) + b(t,x)\varepsilon_5$

with the identities :

$a(t,x)^2 - b(t,x)^2 = 1 + \frac{1}{4}(w^t w - r^t r)$

$a(t,x) b(t,x) = -\frac{1}{4} w^t r$

Denote

$\widehat{r}(t) = \int_{\omega(t)} \mu_3(x,t)\, r(t,x)\, \varpi_3(t,x)\,, \widehat{w}(t) = \int_{\omega(t)} \mu_3(x,t)\, w(t,x)\, \varpi_3(t,x)$
$\widehat{a}(t) = \int_{\omega(t)} a(t,x)\, \mu_3(t,x)\, \varpi_3(t,x)\,, \widehat{b}(t) = \int_{\omega(t)} b(t,x)\, \mu_3(t,x)\, \varpi_3(t,x)$
$\int_{\omega(t)} \mu_3(x,t)\, \sigma(t,x)\, \varpi_3(t,x) = \widehat{a}(t) + \upsilon(\widehat{r}(t), \widehat{w}(t)) + \varepsilon_5 \widehat{b}(t)$
We impose, for a deformable solid, that :
$\exists N(t) \in \mathbb{R}, R(t), W(t) \in \mathbb{R}^3$:

$$\widehat{a}(t) + \upsilon(\widehat{r}(t), \widehat{w}(t)) + \varepsilon_5 \widehat{b}(t) = N(t)\,(A(t) + \upsilon(R(t), W(t)) + B(t)\,\varepsilon_5) \tag{4.39}$$

such that $\sigma_B(t) = A(t) + \upsilon(R(t), W(t)) + B(t)\,\varepsilon_5 \in Spin(3,1)$ which requires :
$A^2 - B^2 = 1 + \frac{1}{4}(W^t W - R^t R)$
$AB = -\frac{1}{4} W^t R$
and implies :
$R(t) = \frac{1}{N}\widehat{r}(t), W(t) = \frac{1}{N}\widehat{w}(t)$
$A(t) = \frac{1}{N}\widehat{a}(t), B(t) = \frac{1}{N}\widehat{b}(t)$
Which sums up to the two conditions :

$$\widehat{a}(t)\,\widehat{b}(t) = -\frac{1}{4}\widehat{w}^t \widehat{r} \tag{4.40}$$

$$\widehat{a}^2 - \widehat{b}^2 = N^2 + \frac{1}{4}\left(\widehat{w}^t \widehat{w} - \widehat{r}^t \widehat{r}\right) \tag{4.41}$$

$\Rightarrow N^2 = \widehat{a}^2 - \widehat{b}^2 - \frac{1}{4}(\widehat{w}^t \widehat{w} - \widehat{r}^t \widehat{r}) > 0$
Then the Spinor of the body is : $S_B(t) = N(t)\,\gamma C(\sigma_B(t))\, S_0$
The conditions can be seen as resulting from the forces which keep the cohesion of the body.
The mass of the solid is proportional to
$\langle S_B(t), S_B(t)\rangle = N^2(t)\,\langle S_0, S_0\rangle = \left(\widehat{a}^2 - \widehat{b}^2 - \frac{1}{4}(\widehat{w}^t \widehat{w} - \widehat{r}^t \widehat{r})\right)\langle S_0, S_0\rangle$
and is not necessarily constant. So we may impose the additional condition :
$\frac{d}{dt}\left(\widehat{a}^2 - \widehat{b}^2 - \frac{1}{4}(\widehat{w}^t \widehat{w} - \widehat{r}^t \widehat{r})\right) = 0 \Leftrightarrow \widehat{a}\frac{d\widehat{a}}{dt} - \widehat{b}\frac{d\widehat{b}}{dt} - \frac{1}{4}\left(\widehat{w}^t \frac{d\widehat{w}}{dt} - \widehat{r}^t \frac{d\widehat{r}}{dt}\right) = 0$
See below continuity equation.

In this aggregation the section σ represents the individual state of the constituting material points, with respect to a gauge attached to the solid. The element $\sigma_B(t)$ represents the average motion of these points with respect to the gauge of the observer B in the computation of σ_B. The motion of the solid itself, with respect to the gauge of an observer O (in a different, more general model), is represented by an element $\sigma_o \in P_G$. The total motion (solid + solid on itself) is defined by a change of gauge in P_G and the resulting spinor (as it would be used in a model representing the solid) is then :

$S(t) = N(t)\,\gamma C(\sigma_o(t))\,\gamma C(\sigma_B(t))\, S_0$

Which sums up to replace the fixed inertial spinor S_0 by the variable spinor $S_B(t) = N(t)\,\gamma C(\sigma_B(t))\, S_0$.

The physical meaning of σ_o must be understood with respect to the way the solid is defined : for instance if $\sigma(t,x)$ represents a rotation around an axis, then σ_o will be a rotation of this axis. The vector $r \in \mathbb{R}^3$ in σ_r, which has no geometric meaning for a particle, gets one for a solid, similar to the usual.

Continuity equation

The conservation of the mass of the body means, for the observer B, that :

$\mathcal{M}(t) = \int_{\omega(t)} \mu_3(t,x)\,\varpi_3 = Ct = \int_{\omega(t)} i_V(\mu\varpi_4)$

Consider the manifold $\omega([t_1,t_2])$ with borders $\omega(t_1), \omega(t_2)$:

$\mathcal{M}(t_2) - \mathcal{M}(t_1) = \int_{\partial\omega([t_1,t_2])} i_V(\mu\varpi_4) = \int_{\omega([t_1,t_2])} d(i_V\mu\varpi_4)$

$d(i_V\mu\varpi_4) = \pounds_V(\mu\varpi_4) - i_V d(\mu\varpi_4) = \pounds_V(\mu\varpi_4) - i_V(d\mu\wedge\varpi_4) - i_V\mu d\varpi_4 = \pounds_V(\mu\varpi_4)$

$\mathcal{E}(t_2) - \mathcal{E}(t_1) = \int_{\omega([t_1,t_2])} \pounds_V(\mu\varpi_4)$

with the Lie derivative $\pounds$ (Maths.1517,1587)

The conservation of the mass is equivalent to the condition $\pounds_V(\mu\varpi_4) = 0$.

$\pounds_V\mu\varpi_4$

$= \frac{d\mu}{dt}\varpi_4 + \mu\pounds_V\varpi_4$

$= \frac{d\mu}{dt}\varpi_4 + \mu(divV)\varpi_4$

$= \frac{d\mu}{dt} + \mu(divV)\varpi_4$

and we retrieve the usual continuity equation :

$$\frac{d\mu}{dt} + \mu divV = 0 \tag{4.42}$$

Then $N(t) = \int_{\omega(t)} \mu_3(t,x)\,\varpi_3$.

If $\mu = Ct$ (incompressible solid) the condition becomes : $divV = 0$

Symmetries of a solid

Symmetries are a non issue for particles, but can be considered for the whole body B : under a geometric transformation the body looks the same for an observer. So they are transformations occurring in each $\omega(t)$ and for a privileged observer who can see the whole body along his world line. The only conceivable symmetries are spatial symmetries, occurring in the hypersurfaces $\omega(t)$. In the computations above we have a great freedom, and one can choose a chart of $\omega(0)$ which accounts for this symmetry. There is a symmetry if $S(0, f(x)) = \gamma C(\sigma(0,x))\,S_0$ for some map : $f : \omega(0) \to \omega(0)$. But the main issue is : does the symmetry is kept along its deformation ? and one cannot answer this question without the knowledge of the external forces which are exercised on the body.

This approach can be used at any scale. It can be used to study the deformation of nuclei, atoms or molecules. At the other end it can be useful in Astrophysics, where trajectories of stars systems or galaxies are studied. The spinor S_B can account for the rotational momentum of the bodies, which is significant and contributes to the total kinetic energy of the system.

Chapter 5

FORCE FIELDS

The concept of fields has appeared in the XIX° century, in the wake of the electromagnetism theory, to replace the picture of action at a distance between particles. In the following by force field we mean one of the forces which interact with particles : the strong interaction, the weak and the electromagnetic forces combined in an electroweak interaction, gravitation being in one league by itself.

A force field is one object of Physics, which has distinctive properties :

i) Because particles are localized, a field must be able to act anywhere, that is to be present everywhere. So the first feature of force fields, as opposed to particles, is that, a priori, they are defined all over the universe, even if their action can decrease quickly with the distance.

ii) A force field propagates : the value of the field depends on the location, this propagation occurs when there is no particle, thus it is assumed that it results from the interaction of the force fields with themselves.

iii) Force fields interact with particles, which are themselves seen as the source of the fields. This interaction depend on charges which are carried by the particles.

iv) The interactions, of the fields with themselves or with particles are, in continuous processes, represented in a lagrangian according to the Principle of Least Action.

v) In some cases the force fields can act in discontinuous processes, in which they can be represented as particles (bosons and gravitons).

Thus we need a representation of the charges and of the fields. We will start with a short presentation of the Standard Model, as this is the most comprehensive picture of the force fields.

5.1 THE STANDARD MODEL

In the Standard Model there are 4 force fields which interact with particles (the gravitational field is not included) :

- the electromagnetic field (EM)
- the weak interactions
- the strong interactions
- the Higgs field

and two classes of elementary particles, fermions and bosons[1], in distinct families.

They are the main topic of the Quantum Theory of Fields (QTF).

5.1.1 Fermions and bosons

Fermions

The matter particles, that we will call fermions, are organized in 3 generations, comprised each of 2 leptons and 2 quarks :

- First generation : quarks up and down; leptons : electron, neutrino.
- Second generation : quarks charm and strange; leptons : muon, muon neutrino
- Third generation : quarks top and bottom; leptons : tau and tau neutrino

Their stability decreases with each generation, the first generation constitutes the usual matter. Each type of particle is called a flavor.

Fermions interact with the force fields, according to their charge, which are :

- color (strong interactions) : each type of quark can have one of 3 different colors (blue, green, red) and they are the only fermions which interact with the strong field
- hypercharge (electroweak interaction) : all fermions have an hypercharge (-2,-1,0,1,2) and interact with the weak field
- electric charge (electromagnetic interactions) : except the neutrinos all fermions have an electric charge and interact with the electromagnetic field.

All fermions have a weak isospin T_3, equal to $\pm 1/2$ and there is a relation between the isospin, the electric charge Q and the hypercharge Y :

$Y = 2\left(Q - T_3\right)$

The total sum of weak isospin is conserved in interactions.

Each fermion (as it seems also true for the neutrinos) has a mass and so interacts also with the gravitational field. These kinematic properties are represented in the Standard Model by a spinor with 4 components[2], and in weak and strong interactions the left and right components interact differently with the fields (this is the chirality effect noticed previously).

Each fermion has an associated antiparticle, which is represented by conjugation of the particle. In the process the charge changes (color becomes anticolor which are different, hypercharge takes the opposite sign), left handed spinors are exchanged with right handed spinors, but the mass is the same.

Elementary particles can be combined together to give other particles, which have mass, spin, charge,... and behave as a single particle. Quarks cannot be observed individually and

[1] Actually the words fermions and bosons are also used for particles, which are not necessarily elementary, that follow the Fermi or the Bose rules in stattistics related to many interacting particles. Here we are concerned only with elementary particles. So fermions mean elementary fermions and bosons elementary bosons or gauge bosons..

[2] Because the right and left part are related, usually only one of them is used in computations, and we have two components Weyl spinors.

group together to form a meson (a quark and anti-quark) or a baryon (3 quarks) : a proton is composed of 3 quarks *udd* and a neutron of 3 quarks *uud*. A particle can transformed itself into another one, it can also disintegrate in other particles, and conversely particles can be created in discontinuous process, notably through collisions. The weak interaction is the only field which can change the flavor in a spontaneous, discontinuous, process, and is responsible for natural radioactivity.

Bosons

Besides the fermions, the Standard Model involves other objects, called gauge bosons, linked to the force fields, which share some of the characteristics of particles. They are :

- 8 gluons linked to the strong interactions : they have no electric charge but each of them carries a color and an anticolor, and are massless. They are their own antiparticles.

- 3 bosons W^j linked with the electroweak field, which carry weak hypercharge and have a mass.

- 1 boson B, specific to the electromagnetic field, which carries a hypercharge and a mass.

- 1 Higgs boson, which has two bonded components, is its own antiparticle and has a mass but no charge or color

The bosons W, B combine to give the photon, the neutral boson Z and the charged bosons $W^{\pm}$. The photon and Z are their antiparticle, $W^{\pm}$ are the antiparticle of each other. So in the Standard Model photons are not elementary particles (at least when electroweak interactions are considered).

5.1.2 The group representation

To put some order in the zoo of the many particles which were discovered a natural starting point is QM : since states of particles can be represented in Hilbert space, it seems logical to assign to each (truly) elementary fermion a vector of a basis of this Hilbert space F. Then the combinations which appear are represented by vectors ϕ, which are linear combinations (or in some case tensorial products) of these basis vectors, and the process of creation / annihilation are transitions between given states, following probability laws. The fact that there are three distinct generations of fermions, which interact together and appear in distinctive patterns, leads to the idea that they correspond to different representations of a group U. Indeed the representations of compact groups can be decomposed in sum of finite dimensional irreducible representations, thus one can have in the same way one group and several distinct but related Hilbert spaces. The problem was then to identify both the group U, and its representations. A given group has not always a representation of a given dimension, and representations can be combined together. Experiments lead to the choice of the direct product SU(3)×SU(2)×U(1) as the group, and to precise the representations (whose definition is technical and complicated, but does not involve high dimensions). Actually the range and the strength of the force fields are different : the range is very short for the strong and weak interactions, infinite for the electromagnetic field, moreover all fermions interact with the weak force and, except for the neutrinos, with the electromagnetic field. So this leads to associate more specifically a group to each force field :

- $SU(3)$ for the strong force

- $U(1) \times SU(2)$ for the electroweak force (when the weak force is involved, the electromagnetic field is necessarily involved)

- $U(1)$ for the electromagnetic force

and to consider three layers : $U(1)$, $U(1) \times SU(2)$, $U(1) \times SU(2) \times SU(3)$ according to the forces that are involved in a problem.

On the other hand it was necessary to find a representation of the force fields, if possible which fits with the representation of the fermions. The first satisfying expression of the Maxwell's laws is relativist and leads to the introduction of the potential $\grave{A}$, which is a 1-form, and of the strength of the field $\mathcal{F}$, which is a two-form, to replace the electric and magnetic fields. It was soon shown that the Maxwell's equations can be expressed elegantly in the fiber bundle formalism, with the group $U(1)$. In the attempt to give a covariant (in the SR context) expression of the Schrodinger's equation including the electromagnetic field it was seen that this formalism was necessary. Later Yang and Mills introduced the same formalism for the weak interactions, which was extended to the strong interactions, and it became commonly accepted in what is called the gauge theories. The key object in this representation is a connection, coming from a potential, acting on a vector bundle, where ϕ lives, which corresponds to the representation of the group U.

5.1.3 The Standard Model

The Standard Model is a version of the Yang-Mills model, adapted to the Special Relativity geometry :

i) Each of the groups or product of groups defines a principal bundle over the Minkovski affine space (which is $\mathbb{R}^4$ with the Lorentz metric).

ii) The physical characteristics (the charges) of the particles are vectors ϕ of a vector bundle associated to a principal bundle modelled on U.

iii) The state of the particles is then represented in a tensorial bundle, combining the spinor S (for the kinematic characteristics) and the physical characteristics ϕ.

iv) The masses are defined separately, because it is necessary to distinguish the proper mass and an apparent mass resulting from the screening by virtual particles.

v) Linear combination of these fermions give resonances which have usually a very short life. Stable elementary particles (such as the proton and the neutron) are bound states of elementary particles, represented as tensorial combinations of these fermions.

vi) The fields are represented by principal connections, which act on the vector bundles through ϕ. The Higgs field is represented through a complex valued function. The electroweak field acts differently on the chiral parts of fermions.

vii) The lagrangian is built from scalar products and the Dirac's operator.

viii) The bosons correspond to vectors of bases of the Lie algebras of each of the groups : the 8 gluons to $su(3)$, the 3 bosons W^j to $su(2)$, 1 boson B to $u(1)$.

5.1.4 The issues

The Standard Model does not sum up all of QTF, which encompasses many other aspects of the interactions between fields and particles. However there are several open issues in the Standard Model.

1. The Standard Model, built in the Special Relativity geometry, ignores gravitation. Considering the discrepancy between the forces at play, this is not really a problem for a model dedicated to the study of elementary particles. QTF is rooted in the Poincaré's algebra, and the localized state vectors, so it has no tool to handle trajectories, which are a key component of differential geometry.

2. The Higgs boson, celebrated recently, raises almost as many questions as it gives answers. It has been introduced in what can be considered as a patch, needed to solve the issue of masses for fermions and bosons. The Dirac's operator, as it is used for the fermions, does not give a definite positive scalar product and is null (and so their mass) whenever the particles are chiral. And as for the bosons, the equivariance in a change of gauge forbids the explicit introduction of

the potential, which is assumed to be their correct representation, in the lagrangian. The Higgs boson solves these problems, but at the cost of many additional parameters, and the introduction of a fifth force which it should carry.

3. From a semi-classic lagrangian, actually most of the practical implementation of the Standard Model relies on particles to particles interactions, detailed by Feynmann's diagram and computed through perturbative methods. Force fields are actually localized operators acting on the states of particles, which is consistent with a dual vision of particles and fields, and with a discrete representation of the physical world, but in the process the mechanism of propagation vanishes.

4. The range of the weak and strong interactions is not well understood. Formally it is represented by the introduction of a Yukawa potential (which appears as a "constant coupling" in the Standard Model), proportional to $\frac{1}{r}\exp(-km)$ which implies that if the mass m of the carrier boson is not null the range decreases quickly with the distance r. Practically, as far as the system which is studied is limited to few particles, this is not a big issue.

5. We could wish to incorporate the three groups in a single one, meanwhile encompassing the gravitational field and explaining the hierarchy between the forces. This is the main topic of the Great Unification Theories (GUT) (see Sehbatu for a review of the subject). The latest, undergone by Garrett Lisi, invokes the exceptional Lie group $E8$. Its sheer size (its dimension is 248) enables to account for everything, but also requires the introduction of as many parameters.

An option, which has been studied by Trayling and Lisi, would be to start, not from Lie groups, but from Clifford algebras as we have done for the Spinors. The real dimension of $SU(3)\times SU(2)\times U(1)$ is $12 = 8 + 3 + 1$ which implies to involve at least a Clifford algebra (dimension 2^n) on a four dimensional vector space and it makes sense to look at its complexified. The groups would then be Spin subgroups of the Clifford algebra. We have the following isomorphisms :

$U(1) \sim Spin(\mathbb{R}, 2)$

$SU(2) \sim Spin(\mathbb{R}, 3)$

but there is no simple isomorphism for $SU(3)$.

Albeit all together they are part of $Cl(\mathbb{R}, 10)$.

In the next sections we will see how the states of particles, force fields, including gravitation, and their interactions can be represented, in the geometrical context of GR. In the next chapter we will review the requirements that these representations impose to Lagrangians and continuous models. Two kinds of continuous models, simplified but similar to the Standard Model, will then be studied. They do not pretend to replace the Standard Model, but to help to understand the mechanisms at play, notably the motivation to use the mathematical tools in the representation of physical phenomena. So we will not insist on the many technical details of the Standard Model, heavily loaded with historical notations, and keep the formalism to a minimum.

5.2 STATES OF PARTICLES

Spinor fields can be characterized, beyond the inertial spinor, by an integer, which defines families of particles with similar behavior. Particles can then be differentiated, in addition to their kinematic characteristics summarized in the spinor, by a charge which accounts for their interaction with force fields. A particle can be seen as a system in itself. Its state is then a combination of its kinematic characteristics, represented by the spinor, and of its charge, which represents its interaction with the force fields. Using the description of elementary particles given by the Standard Model, it is then possible to set up a representation of elementary particles. From there the representation can be extended to composite particles and matter fields.

5.2.1 The space of representation of the states

The Law of Equivalence

We can follow some guidelines :

i) For any particle there are intrinsic characteristics ψ_0, which do not change with the fields or the motion. If we assume that ψ belongs to a vector space V, then there is a set of vectors $\{\psi_{0p}\}_{p=1}^{N}$ such that ψ_{0p} characterizes a family of particles which have the same behavior.

ii) Motion is one of the features of the state of particles. It is represented by the action of $Spin(3,1)$ on the space V, as we have done in the previous chapter.

iii) The intrinsic kinematic characteristics of particles are represented in the vector spaces E_0, E_0' : each family of particles is associated to one vector of these spaces. Particles and anti-particles are distinguished by their inertial spinor.

iv) In the Newton's law of gravitation $F = G\frac{MM'}{r^2}$ and his law of Mechanics : $F = \mu\gamma$ the scalars M, μ represent respectively the gravitational charge and the inertial mass, and there is no reason why they should be equal. However this fact has been verified with great accuracy (two bodies fall in the vacuum at the same speed). This has lead Einstein to state the fundamental Law of Equivalence "Gravitational charge and inertial mass are identical". From which he built the Theory of General Relativity. This Law leads us to take as gravitational charge of particles the inertial spinor $\frac{1}{c}S_0$, where c is necessary to keep the dimension of mass to the gravitational charge.

Proposition 85 *The Gravitational charge of a particle is represented by its inertial spinor* $\frac{1}{c}S_0$.

So, if we stay only with the gravitational field, the space E and the representation $(E, \gamma C)$ suffice to represent the state of particles. The kinematic characteristics of particles of the same flavor (quarks, leptons) are not differentiated according to their other charges. So we have $\psi_{0p} = S_{0p}$.

In the previous chapter we assumed that :

- there is, along the world line of a particle, a privileged frame $\varphi_G(m, \sigma(m))$ such that the spinor of the particle is $(\varphi_G(m, \sigma(m)), S_0)$ with $S_0 = Ct$

- the observer measures the spinor $S(m)$ in his gauge : $\varphi_G(m, 1)$ and $(\varphi_G(m, \sigma(m)), S_0) \sim (\varphi_G(m, 1), \gamma C(\sigma(m)) S_0) = (\varphi_G(m, 1), S(m))$ thus : $S(m) = \gamma C(\sigma(m)) S_0$

We have now to consider an interpretation which is mathematically equivalent, but physically different :

- the observer measures the spinor $S(m)$ with $(\varphi_G(m, 1), \gamma C(\sigma(m)) S_0)$

- in presence of gravity this spinor is equivalent to : $(\varphi_G(m, \sigma(m)), S_0)$

The privileged gauge (for the particle) is provided by the gravitational field. And *the action of the motion, that is of the inertial forces, is equivalent to the action of gravity on the state of* *the particle*, which is the meaning of the Law of Equivalence. This is a key point to understand the gravitational and the other fields : *particles have intrinsic properties, that they keep all over their travel on their world lines but, because of the existence of the field, their measure by an observer is distinct from this intrinsic value.* This leads to see the fields as the value of the element of the group ($\sigma \in Spin(3,1)$ for the gravitational field) but, as we will see, the action of the field goes through a special derivative because it manifests itself in the motion of the particle.

Representation of the charges for the other fields

For the other fields :

i) Bosons give the structure of the fields, in accordance with the dimension of the groups : 8 for the strong force ($SU(3)$ dimension 8), 3 for the weak force ($SU(2)$ dimension 3), 1 for the electromagnetic force ($U(1)$ dimension 1). In QTF the action of fields is represented by operators acting on V, in the representation of the Lie algebra of the groups. Because the exponential is surjective on compact groups it sums up to associate the fields to an action of the groups on V.

ii) The action depends on the charges - accounting for the possible combinations of charges, there all together 24 kinds of fermions - but also on the inertial spinors : particles and antiparticles do not behave the same way, and weak forces act differently according to the left or right chiral parts

Assuming that V is a vector space, and the actions of the fields are linear, the solution is to take V as the tensorial product $V = E \otimes F$ where F is a vector space such that (F, ϱ) is a representation of the group U corresponding to the forces other than gravity ($U = SU(3) \times SU(2) \times U(1)$ in the Standard Model).

That we sum up by :

Proposition 86 *There is a compact, connected, real Lie group U which characterizes the force fields other than gravitation.*

There is a n dimensional complex vector space F, endowed with a definite positive scalar product denoted $\langle\rangle_F$ and (F, ϱ) is a unitary representation of U

The states of particles are vectors of the tensorial product $E \otimes F$

The intrinsic characteristics of each type of particles are represented by a tensor $\psi_0 \in E \otimes F$,that we call a ***fundamental state****, and all particles sharing the same characteristics behave identically under the actions of all the fields.*

Notation 87 $(f_i)_{i=1}^n$ *is a basis of F. We will assume that it is orthonormal.*

$\left(\overrightarrow{\theta}_a\right)_{a=1}^m$ *is a basis of the Lie algebra T_1U*

$[\theta_a]$ *is the matrix of* $\varrho'(1)\left(\overrightarrow{\theta}_a\right)$ *expressed in the basis* $(f_i)_{i=1}^n$.

As a consequence :

i) Because (F, ϱ) is a unitary representation, the scalar product is preserved by ϱ : $\langle \varrho(g)\phi, \varrho(g)\phi' \rangle_F = \langle \phi, \phi' \rangle_F$

ii) $(F, \varrho'(1))$ is a representation of the Lie algebra T_1U

iii) The derivative $\varrho'(1)$ is anti-unitary and the matrices $\left[\varrho'(1)\overrightarrow{\theta}_a\right] = [\theta_a]$ are anti-hermitian :

$$[\theta_a] = -[\theta_a]^* \tag{5.1}$$

F must be a complex vector space to account for the electromagnetic field. F is actually organized as different representations of the group U, and the representation is not irreducible, to account for the generations effect. Composite particles (such as the proton or the neutron) are represented by tensorial product of vectors of $E \otimes F$.

A basis of $E \otimes F$ is $(e_i \otimes f_j)_{i=0..3}^{j=1...n}$

The state of a particle is expressed as a tensor :

$\psi = \sum_{i=1}^{4} \sum_{j=1}^{n} \psi^{ij} e_i \otimes f_j$ that we will usually denote in the matrix form : $[\psi]$ with 4 rows and n columns.

which reads :

$\psi = \sum_{j=1}^{n} \left(\sum_{i=1}^{4} \psi^{ij} e_i \right) \otimes f_j = \sum_{j=1}^{n} S^j \otimes f_j$ where $S^j \in E$

So, when gravity alone is involved, the particles such as $\sum_{i=1}^{4} \psi_0^{ij} e_i = S_0^j$ have the same behavior and can be seen as n particles, differentiated by their inertial spinor, and thus by their mass. At an elementary level the different values of the inertial spinors characterize the kinematics of each elementary particle.

The experimental fact that the action of the force fields depends also of the spinor part implies that the tensor is not necessarily decomposable (it cannot be written as the tensorial product of two vectors). However one can attribute a charge to a particle, but it is not expressed as a scalar quantity. There is no natural unit for the charges (except, for historical reasons, for the electric charge), and, indeed, what could be the unit for the colors of the strong force ? The set $\mathfrak{F}$ of existing vectors ψ_0 is just an organized map of all the known combinations of spinors and charges. The formalism with the group representation is built on the experimental facts, but it does not answer the question : why is it so ?

The direct product group $Spin(3,1) \times U$ has an action denoted ϑ on $E \otimes F$

$\vartheta : Spin(3,1) \times U \to \mathcal{L}(E \otimes F; E \otimes F)$

defined by linear extension of γC and ϱ :

$\vartheta(\sigma, \varkappa)(\psi) = \sum_{i,k=1}^{4} \sum_{j,l=1}^{n} [\gamma C(\sigma)]_k^i [\varrho(\varkappa)]_l^j \psi^{kl} e_i \otimes f_j$

that we will denote in matrices :

Notation 88

$$\vartheta : Spin(3,1) \times U \to \mathcal{L}(E \otimes F; E \otimes F) :: \vartheta(\sigma, \varkappa)[\psi] = [\gamma C(\sigma)][\psi][\varrho(\varkappa)] \quad (5.2)$$

One can extend the action of the Spin group to the action of the Clifford algebra. We define the action ϑ of $Cl(\mathbb{R}, 3, 1) \times U$ on $E \otimes F$ by the unique linear extension of :

$\vartheta : Cl(\mathbb{R}, 3, 1) \times U \to \mathcal{L}(E \otimes F; E \otimes F) ::$

$\vartheta(s, g)(S \otimes \phi) = \gamma C(s)(S) \otimes \varrho(g)(\phi)$

to all tensors on $E \otimes F$

This is a morphism from $Cl(\mathbb{R}, 3, 1)$ on $L(E \otimes F; E \otimes F)$: ϑ is linear and preserves the Clifford product.

$\vartheta(\sigma, 1)\psi = \gamma C(\sigma)\psi = \sum_{jkl} [\gamma C(\sigma)]_k^j \psi^{kl} e_j \otimes f_l$

So the map ϑ defines a representation of $Cl(\mathbb{R}, 3, 1) \times U$ on $E \otimes F$.

Scalar product on the space $E \otimes F$

The scalar product on $E \otimes F$ is necessarily defined as :

$\langle \psi, \psi' \rangle = \sum_{ijq} [\gamma_0]_k^i \delta_{jq} \overline{\psi}^{ij} \psi'^{kq} = \sum_{ijk} [\gamma_0]_k^i \overline{\psi}^{ij} \psi'^{kj} = Tr\left([\psi]^* [\gamma_0][\psi']\right)$

because the basis $(f_j)_{j=1}^{n}$ is orthonormal.

$$\langle \psi, \psi' \rangle = Tr \left([\psi]^* [\gamma_0] [\psi'] \right) \quad (5.3)$$

Theorem 89 *The scalar product on $E \otimes F$ is preserved by ϑ :*
$\langle \vartheta (\sigma, \varkappa) \psi, \vartheta (\sigma, \varkappa) \psi' \rangle = \langle \psi, \psi' \rangle$

Proof. $\widetilde{\psi}^{ij} = \sum_{k=1}^{4} \sum_{l=1}^{n} [\gamma C (\sigma)]^i_k [\varrho (\varkappa)]^j_l \psi^{kl}$
$\left\langle \widetilde{\psi}, \widetilde{\psi}' \right\rangle = \sum [\gamma_0]^i_k \overline{[\gamma C (\sigma)]}^i_p \overline{[\varrho (\varkappa)]}^j_q \overline{\psi}^{pq} [\gamma C (\sigma)]^k_r [\varrho (\varkappa)]^j_s \psi^{rs}$
$= \sum \left([\gamma C (\sigma)]^* [\gamma_0] [\gamma C (\sigma)] \right)^p_r \left([\varrho (\varkappa)]^* [\varrho (\varkappa)] \right)^q_s \overline{\psi}^{pq} \psi'^{rs}$
$= \sum [\gamma_0]^p_r \overline{\psi}^{pq} \psi'^{rq}$ ■

The scalar product is not definite, positive or negative, on $E \otimes F$. However there is a norm $\|\|_E$ on the space E and a norm on the space F, the latter defined by the scalar product. They define a norm on $E \otimes F$ by taking $\|e_i \otimes f_j\| = \|e_i\|_E \|f_j\|_F$. Moreover this norm is invariant by ϑ. So that $E \otimes F$ is a Banach vector space.

Physical states of elementary particles

$(E \otimes F, \vartheta)$ is a unitary representation of the Lie group $Spin(3,1) \otimes U$.

For any $\psi \in E \otimes F$ the set $\{\vartheta (\sigma, \varkappa) \psi, (\sigma, \varkappa) \in Spin(3,1) \otimes U\}$ is the orbit of ψ. The orbits are the set of states corresponding to the same type of particles.

The relation of equivalence $\psi \sim \psi' \Leftrightarrow \exists (\sigma, \varkappa) \in Spin(3,1) \otimes U : \psi' = \vartheta (\sigma, \varkappa) \psi$ defines a partition of $E \otimes F$ corresponding to the orbits. And each class of equivalence can be identified with a fundamental state ψ_0.

All particles of the same type ψ_0 have the same behavior with the same fields $\varkappa$: so for $\psi_0, \varkappa$ fixed, σ then ψ are fixed uniquely

The measures of fields is done by measuring the motion σ of known particles ψ_0 subjected to fields $\varkappa$: so from ψ, ψ_0 and σ one can compute a unique value $\varkappa$ of the field.

Which sums up to, if $\mathfrak{F}$ is the set of possible states of elementary particles :

Proposition 90 *The action of $Spin(3,1) \times U$ on $\mathfrak{F}$ is free and faithful : $\forall \psi \in \mathfrak{F} : \vartheta (\sigma, \varkappa) \psi = \psi \Leftrightarrow (\sigma, \varkappa) = (1,1)$*

Then $\vartheta (\sigma, \varkappa) \psi = \vartheta (\sigma', \varkappa') \psi \Leftrightarrow (\sigma, \varkappa) = (\sigma', \varkappa')$

This is the case for the spinor and is extended to the states of particles.

The orbits are not vector subspaces :

Theorem 91 *For any fundamental state ψ_0, the orbit $(E \otimes F)(\psi_0)$ of ψ_0 is a real finite dimensional Riemannian manifold, embedded in $E \otimes F$*

Proof. $Spin(3,1)$ and U are real Lie groups, thus manifolds, take a chart in each

The vector spaces tangent at any point to the manifold are subspaces of the vector space $E \otimes F$

The metric on the tangent bundle is given by the scalar product, which is definite, positive or negative. ■

Particles and antiparticles

Chirality was introduced in the choice of the representation $(E, \gamma C)$ because it is significant in the behavior with the force fields. We will similarly distinguish in the matrix $[\psi]$ a right part, with the first 2 rows, and a left part,with the last 2 rows, so that in matrix form $[\psi] = \begin{bmatrix} \psi_R \\ \psi_L \end{bmatrix}$. In QTF this is called a Dirac's spinor, and ψ_R, ψ_L are Weyl's spinors.

The difference between particles and antiparticles was based on the sign of the scalar product $\langle S_0, S_0 \rangle$. So it is legitimate to discriminate particles and antiparticles in a similar way. More precisely, we look for the subsets of $E \otimes F$ such that :

i) the scalar product is definite either positive or negative : $\langle \psi_0, \psi_0 \rangle = 0 \Rightarrow \psi_0 = 0$

ii) this is still true whenever ψ_0 is the tensorial product $\psi_0 = S_0 \otimes F_0$

iii) the populations of antiparticles and particles are preserved by space reversal, and exchanged by time reversal, as we know that this is still true for particles in the Standard Model.

Theorem 92 *The only vector subspaces of* $E \otimes F$ *which meet these conditions are such that* $\psi_L = \epsilon i \psi_R$ *with* $\epsilon = \pm 1$

Proof. i) $\langle \psi, \psi \rangle = Tr\left([\psi]^* [\gamma_0] [\psi]\right) = iTr\left(-\psi_R^* \psi_L + \psi_L^* \psi_R\right)$

$Tr\left(\psi_L^* \psi_R\right) = Tr\left(\psi_L^* \psi_R\right)^t = Tr\left(\psi_R^t \overline{\psi_L}\right) = \overline{Tr\left(\psi_R^* \psi_L\right)}$

Thus : $Tr\left(-\psi_R^* \psi_L + \psi_L^* \psi_R\right) = \overline{Tr\left(\psi_R^* \psi_L\right)} - Tr\left(\psi_R^* \psi_L\right)$

$= -2i \operatorname{Im} Tr\left(\psi_R^* \psi_L\right) \in i\mathbb{R}$

and $\langle \psi, \psi \rangle = 2 \operatorname{Im} Tr\left(\psi_R^* \psi_L\right) \in \mathbb{R}$

For $\psi = S \otimes F$ the matrix $[\psi]$ reads : $[\psi] = [S] [F]^t = \begin{bmatrix} S_R F^t \\ S_L F^t \end{bmatrix}$

and $\langle \psi, \psi \rangle = 2 \operatorname{Im} Tr\left(\overline{[F]} [S_R]^* [S_L] [F]^t\right) = 2 \operatorname{Im} [S_R]^* [S_L] Tr\left(\overline{[F]} [F]^t\right)$

It will be non degenerate iff : $S_L = \epsilon i S_R$ as seen previously and so we can generalize to $\psi_L = \epsilon i \psi_R$:

$\langle \psi, \psi \rangle = 2 \operatorname{Im} Tr\left(\epsilon i \psi_R^* \psi_R\right) = 2\epsilon Tr\left(\psi_R^* \psi_R\right)$

ii) Time reversal is an operator on $E \otimes F$, represented by the matrix (see the section Spinor Model above) :

$T = \begin{bmatrix} 0 & i\sigma_0 \\ i\sigma_0 & 0 \end{bmatrix}$ with signature (3,1)

$T \begin{bmatrix} \psi_R \\ \epsilon i \psi_R \end{bmatrix} = \begin{bmatrix} 0 & i\sigma_0 \\ i\sigma_0 & 0 \end{bmatrix} \begin{bmatrix} \psi_R \\ \epsilon i \psi_R \end{bmatrix} = \begin{bmatrix} -\epsilon \psi_R \\ i\psi_R \end{bmatrix} = \begin{bmatrix} -\epsilon \psi_R \\ -\epsilon i \left(-\epsilon \psi_R\right) \end{bmatrix}$

$T = \begin{bmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{bmatrix}$ with signature (1,3)

$T \begin{bmatrix} \psi_R \\ \epsilon i \psi_R \end{bmatrix} = \begin{bmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{bmatrix} \begin{bmatrix} \psi_R \\ \epsilon i \psi_R \end{bmatrix} = \begin{bmatrix} i\epsilon \psi_R \\ \psi_R \end{bmatrix} = \begin{bmatrix} i\epsilon \psi_R \\ -\epsilon i \left(i\epsilon \psi_R\right) \end{bmatrix}$

iii) Space reversal is an operator on $E \otimes F$, represented by the matrix :

$S = i\gamma_0 = \begin{bmatrix} 0 & \sigma_0 \\ -\sigma_0 & 0 \end{bmatrix}$ with signature (3,1)

$S \begin{bmatrix} \psi_R \\ \epsilon i \psi_R \end{bmatrix} = \begin{bmatrix} 0 & \sigma_0 \\ -\sigma_0 & 0 \end{bmatrix} \begin{bmatrix} \psi_R \\ \epsilon i \psi_R \end{bmatrix} = \begin{bmatrix} i\epsilon \psi_R \\ -\psi_R \end{bmatrix} = \begin{bmatrix} i\epsilon \psi_R \\ \epsilon i \left(i\epsilon \psi_R\right) \end{bmatrix}$

$S = \begin{bmatrix} 0 & -i\sigma_0 \\ i\sigma_0 & 0 \end{bmatrix}$ with signature (1,3)

$S \begin{bmatrix} \psi_R \\ \epsilon i \psi_R \end{bmatrix} = \begin{bmatrix} 0 & -i\sigma_0 \\ i\sigma_0 & 0 \end{bmatrix} \begin{bmatrix} \psi_R \\ \epsilon i \psi_R \end{bmatrix} = \begin{bmatrix} \epsilon \psi_R \\ i\psi_R \end{bmatrix} = \begin{bmatrix} \epsilon \psi_R \\ \epsilon i \left(\epsilon \psi_R\right) \end{bmatrix}$ ■

And we can state :

Proposition 93 *The fundamental states ψ_0 of particles (fermions) are such that :*
$\psi_L = i\psi_R$ *for particles, their mass M_p is such that*
$\langle \psi_0, \psi_0 \rangle = 2Tr\left(\psi_R^* \psi_R\right) = M_p^2 c^2$
$\psi_L = -i\psi_R$ *for antiparticles, their mass is*
$\langle \psi_0, \psi_0 \rangle = -2Tr\left(\psi_R^* \psi_R\right) = -M_p^2 c^2$
To each fermion is associated an antiparticle which has the same mass.

As ϑ preserves the scalar product : $\langle \vartheta(\sigma,\varkappa)\psi_0, \vartheta(\sigma,\varkappa)\psi_0 \rangle = \langle \psi_0, \psi_0 \rangle$ the scalar product is definite positive or negative on the sets :
$(E \otimes F)(\psi_0) = \{\vartheta(\sigma,\varkappa)\psi_0, \sigma \in Spin(3,1), \varkappa \in U\}$ for a fixed ψ_0 such that $\psi_L = \epsilon i \psi_R$
But these sets are not vector spaces. The expressions are :
$[\psi] = [\gamma C(\sigma)][\psi_0][\varrho(\varkappa)]$ and $\sigma = \sigma_w \cdot \sigma_r$
$\gamma C\left(a + \upsilon(r,w) + b\varepsilon_5\right) \begin{bmatrix} \psi_{0R} \\ \epsilon i \psi_{0R} \end{bmatrix} [\varrho(\varkappa)]$
$= \begin{bmatrix} (a+b) + \frac{1}{2}\sigma(w - ir)\psi_{0R}[\varrho(\varkappa)] \\ \epsilon i(a-b) - \frac{1}{2}\sigma(w+ir)\psi_{0R}[\varrho(\varkappa)] \end{bmatrix}$
$\gamma C(\sigma_r)$ preserves E_0,and similarly the chiral relation $\psi_L = i\psi_R$

CPT Conservation Principle

It is acknowledged that physical laws are invariant by CPT operations. We have already seen the P (space inversion) and T (time inversion).

The C (Charge inversion) operation transforms a charge into its opposite.

We have seen the action of the operators P, T, and from $CPT = I$ we can deduce C :

$$P : \begin{bmatrix} 0 & \sigma_0 \\ -\sigma_0 & 0 \end{bmatrix} \begin{bmatrix} \psi_R \\ \epsilon i \psi_R \end{bmatrix}$$
$$T : i \begin{bmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{bmatrix} \begin{bmatrix} \psi_R \\ \epsilon i \psi_R \end{bmatrix}$$
$$PT : i \begin{bmatrix} -\sigma_0 & 0 \\ 0 & \sigma_0 \end{bmatrix}$$
$$PTC = I = i \begin{bmatrix} -\sigma_0 & 0 \\ 0 & \sigma_0 \end{bmatrix} \begin{bmatrix} \psi_R \\ \epsilon i \psi_R \end{bmatrix} C = \begin{bmatrix} -i\psi_R \\ -\epsilon \psi_R \end{bmatrix} C \equiv \begin{bmatrix} \psi_R \\ \epsilon i \psi_R \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} \psi_R \\ \epsilon i \psi_R \end{bmatrix} C = \begin{bmatrix} -i\psi_R \\ -\epsilon \psi_R \end{bmatrix}$$

As CPT keeps everything, this means that the set of possible values of the fundamental states ψ_0 is organized : antiparticles have charges opposite to the particles. All particles have an associated antiparticle, and there is no particle which is its own antiparticle (but bosons can be their own antibosons), so the dimension of F is necessarily even (each basis vector corresponds to a combination of charges).

The gravitational charge is the inertial spinor S_0, so the operation C would be :
$C \begin{bmatrix} S_R \\ \epsilon i S_R \end{bmatrix} = \begin{bmatrix} -iS_R \\ -\epsilon S_R \end{bmatrix}$

The Electromagnetic field (EM)

In the Standard Model the Electromagnetic field (EM) is represented by the group $U(1)$, the set of complex numbers with module 1 ($uu^* = 1$). It is a compact abelian group. Its irreducible representations are unidimensional, that is multiple of a given vector.

For any given arbitrary vector f there are 3 possible irreducible non equivalent representations :

- the standard one : $(F, \varrho) : \varrho\left(e^{i\phi}\right) f = e^{i\phi} f$ and $F = \left\{e^{i\phi} f, \phi \in \mathbb{R}\right\}$
- the contragredient : $(F, \overline{\varrho}) : \overline{\varrho}\left(e^{i\phi}\right) f = e^{-i\phi} f$ and $F = \left\{e^{i\phi} f, \phi \in \mathbb{R}\right\}$ (Maths.23.1.2)
- the trivial representation : $(F, \varrho) : \varrho\left(e^{i\phi}\right) f = f$ and $F = \{f\}$

The standard representation corresponds to negative charge, the contragredient representation to positive charge and the trivial one to neutral charge. The choice positive / negative is arbitrary.

The EM field interacts similarly with the left and right part of a spinor, so the space of states of the particles is the sum of tensorial products : $S \otimes f$. The theory can be fully expressed this way. However it is legitimate to choose the vectors f in E, which is a 4 dimensional complex vector space. For *elementary particles*, then :

i) Let $\{S_p \in E_0, p = 1...N\}$ be N vectors representing inertial spinors of particles. Then for each of them their states are represented by $\left\{e^{i\phi} S_p, \phi \in \mathbb{R}\right\}$ with the standard representation;

ii) Let $\left\{S'_p \in E'_0, p = 1...N\right\}$ be N vectors representing inertial spinors of particles with the charge opposite to S_p.Then for each of them their states are represented by $\left\{e^{i\phi} S'_p, \phi \in \mathbb{R}\right\}$ with the contragredient representation;

iii) Neutral particles $\{S_q \in E_0 \oplus E'_0, q = 1...N'\}$ correspond to the trivial representation : their states is just one vector $S \in E_0$ or E'_0.

Under the CPT principle the vectors S_p representing the elementary particles, and S'_p representing the particles with opposite charge would be deduced by :

$S_p = \begin{bmatrix} S_R \\ \epsilon i S_R \end{bmatrix} \rightarrow S'_p = \begin{bmatrix} -iS_R \\ -\epsilon S_R \end{bmatrix} = \begin{bmatrix} S'_R \\ -\epsilon i S'_R \end{bmatrix}$

so we have couples particles / anti-particles and particles have a charge opposite to their anti-particle. But this does not solve the problem of the representation of neutral particles (the only known are the neutrinos) which, anyway, do not interact with the EM field.

Notice that the maps $\varrho, \overline{\varrho}$ are distinct from $\gamma C(\sigma)$ for any $\sigma \in Spin, T_1 Spin$, so the state of a particle can change only by the action of $U(1)$.

A basis of E, for elementary particles, is then $\left\{\left(S_p, S'_p\right)_{p=1}^{N}, \left(S_q\right)_{q=1}^{N'}\right\}$. Each vector defines the mass, the charge and the type of the particle. In this picture there is no unit for the electric charge.

In the case of the EM field the structure brought by the charges is then built in the space E_0 or E'_0. The state of an elementary particle is then :

$$\psi = e^{i\phi} S \tag{5.4}$$

where ϕ is a phase factor which must be considered as variable for charged particles, and $\phi = 0$ for neutral particles. As we will see actually the phase can usually be ignored : particles whose states differ by a phase factor have the same behavior, with regard to the electromagnetic field and they have the same mass : $\langle(\exp i\phi) S_0, (\exp i\phi') S_0\rangle = \langle S_0, S_0\rangle$ and the same EM charge.

This is the origin of the introduction of rays in QM. Two particles such that their states differ by a phase factor $e^{i\phi}$ behave the same way, for the gravitational field or the EM field, so they can be deemed representing the same state.

Composite material bodies

The picture that we have just drawn corresponds to elementary particles : the set $\mathfrak{F} \subset E \otimes F$ of possible states comes from the known combinations of charges and kinematic characteristics of the particles which are deemed elementary. But elementary particles are assumed to be the building blocks of material bodies, and they constitute more or less stable composites which can themselves be considered as particles. Several solutions can be considered to represent these composite bodies.

i) Whenever the weak or strong interactions are involved the state of composite bodies is represented by tensors in $\otimes_n \mathfrak{F}$,according to the theorem 28 of QM. In particular they are the only forces which can change the flavor of a particle.

ii) When only the EM field is involved the states of elementary particles can be represented in E. Mathematically the tensorial product of non equivalent representations is well defined. The action of $U(1)$ on the tensor $S_1 \otimes S_2$ is $e^{i\phi(\epsilon_1+\epsilon_2)} S_1 \otimes S_2$ with $\epsilon_k = \pm 1$ depending on the representation. However a basis of the tensorial product is comprised of tensorial products of all the vectors of the basis, which would not have the same behavior under the action of $U(1)$ (the basis of E has positive, negative and neutral particles). So actually the only combinations which are acceptable are made of particles of the same kind (positive, negative or neutral) and the action is then $e^{iq\phi\epsilon}$ where q is the number of particles, where q can be a positive, a negative or a null integer. As a consequence :

- when only the EM field is involved composite particles are comprised of particles with the same type of charge (this does not hold when the weak and strong interactions are considered)
- the electric charge of particles must be an integer multiple of an elementary charge.

Such tensorial products of spinors can be used for nuclei, atoms or molecules. The associated EM charge is an integer multiple of the elementary charge. Of course this does not matter for neutral particles but, as they do not interact with the EM field, we are actually in the next case.

iii) When only the gravitational field is involved the states of particles or material bodies can be represented either by a tensorial product of spinors, or by a single spinor in E.

Material bodies can be represented by a single spinor. There is no scale limit for these solutions, however the representation of a composite body by a single spinor assumes that the body meets the conditions stated in the previous chapter for deformable solids. The interactions between its particles are such that an equilibrium is reached, which preserves its cohesion. And the laws which rule its deformation account for these interactions. As noticed before the aggregation of particles in a deformable solid is done in a separate model, accounting for the specific characteristics of the body. It provides a single spinor which is an approximation of the collection of particles whose state can always be represented by tensors. Such spinors can be considered with the EM field, because they are the combination of an integer multiple of particles with the same charge, the inertial spinor of such composite must account for the total charge, expressed in an integer multiple of the elementary charge. For instance a nucleus can be represented by a single spinor with an EM charge.

Classic Mechanics provides efficient and simpler tools and the use of spinors would be just pedantic in common problems. However there is a domain where spinors could be useful : Astrophysics. It cannot ignore General Relativity, gravity is the main force at work, the common presentation of GR is awfully complicated and does not allow to account (at least in a simple way) for the rotation of stars or star systems, and there are some disturbing facts (nowadays linked to the existence of a dark matter) that are not explained, but about which spinors could provide some answer. In this case the introduction of density besides spinors is the natural path.

In the following we will keep the definition of particles of the previous sections : they are material bodies whose state can be represented by a single vector $\psi \in E \otimes F$, which corresponds either to an elementary particle or a material body whose structure is fully represented in ψ. And we will assume, at least formally, that the state of the particle belong to $E \otimes F$, even if it can be represented by a single spinor in E.

5.2.2 The fiber bundle representation

The action ϑ of the groups gives the value of ψ for any fundamental state ψ_0 :

$\psi : (E \otimes F) \times (Spin(3,1) \times U) \to \psi = \vartheta(\sigma, \varkappa)\psi_0$

The discrepancy between the fundamental state ψ_0 and the measured value ψ is assumed to come from the existence of the fields. So we are lead to assume that there is a principal bundle $Q(M, Spin(3,1) \times U, \pi_U)$ with fiber $Spin(3,1) \times U$, and the fields are represented by a section of this principal bundle : $\mathbf{q}_f \in \mathfrak{X}(Q) :: \mathbf{q}_f(m) = \varphi_Q(m, (\sigma(m), \varkappa(m)))$.

The observer is assumed to use a standard gauge : $\mathbf{q} \in \mathfrak{X}(Q) :: \mathbf{q}(m) = \varphi_Q(m, (1,1))$.

Then the state of the particle is represented by a vector ψ of the associated vector bundle $Q[E \otimes F, \vartheta]$ with fiber $E \otimes F$. This is a geometric quantity, which is intrinsic to the particle and does not depend on a gauge. It adds the charges to the spinor S.

In the presence of fields the state of the particle is

$(\mathbf{q}_f(m), \psi_0) = (\varphi_Q(m, \vartheta(\sigma(m), \varkappa(m))), \psi_0)$

which is equivalent to :

$(\varphi_Q(m, \vartheta(1,1)), \vartheta(\sigma(m), \varkappa(m))\psi_0) = (\mathbf{q}(m), \vartheta(\sigma(m), \varkappa(m))\psi_0)$

and the observer measures $\psi(m) = \vartheta(\sigma(m), \varkappa(m))\psi_0$ in his gauge $\varphi_Q(m, \vartheta(1,1))$.

The measure of the state depends on the observer.

Notice that, as a consequence of this representation, the conservation of the characteristics ψ_0 of the particle entails that of its charge and mass during its motion. It is built in the formalism. And, meanwhile spinor and charge are entangled in the tensorial product $E \otimes F$, the gravitational field and the other fields keep their originality : Q has for fiber $Spin(3,1) \times U$ and not $Spin(3,1) \otimes U$.

That we sum up in :

Proposition 94 *There is a principal bundle $Q(M, Spin(3,1) \times U, \pi_U)$ with trivialization $\varphi_Q(m, (\sigma, \varkappa))$ and the fields are represented by sections $\mathbf{q}_f \in \mathfrak{X}(Q)$ of the principal bundle.*

The state of the particles is represented as vectors of the associated bundle $Q[E \otimes F, \vartheta]$

The observers measure the state in a standard gauge $\mathbf{q}(m) = \varphi_Q(m, (1,1)) \in \mathfrak{X}(Q)$ and the measured states of particles in this gauge are $\psi(m) = \vartheta(\sigma(m), \varkappa(m))\psi_0$

$Q[E \otimes F, \vartheta]$ has for trivialization :

$(\varphi_Q(m, (1,1)), \psi) \sim (\varphi_Q(m, (s^{-1}, g^{-1})), \vartheta(s,g)\psi)$

and holonomic basis:

$(\mathbf{e}_i(m) \otimes \mathbf{f}_j(m))_{i=0..3}^{j=1...n} = (\varphi_Q(m, (1,1)), e_i \otimes f_j)$

$$\psi(m) = \sum_{i=1}^{4}\sum_{j=1}^{n} [\gamma C(\sigma(m))]_k^i [\varrho(\varkappa(m))]_l^j \psi_0^{kl}(m) \mathbf{e}_i(m) \otimes \mathbf{f}_j(m) \tag{5.5}$$

$$\text{in matrix form : } [\psi]_{4\times n} = [\gamma C(\sigma)][\psi][\rho(\varkappa)] \tag{5.6}$$

A change of trivialization with a section $\chi(m) \in \mathfrak{X}(Q)$ induces a change of gauge :

$$\begin{bmatrix} \mathbf{q}(m) = \varphi_Q(m,(1,1)) \to \widetilde{\mathbf{q}}(m) = \widetilde{\varphi}_Q(m,(1,1)) = \mathbf{q}(m) \cdot \chi(m)^{-1} \\ (\sigma(m), \varkappa(m)) = \varphi_Q(m,(\sigma,\varkappa)) = \widetilde{\varphi}_Q(m,(\widetilde{\sigma},\widetilde{\varkappa})) : (\widetilde{\sigma},\widetilde{\varkappa}) = \chi(m) \cdot (\sigma,\varkappa) \\ \mathbf{e}_i(m) \otimes \mathbf{f}_j(m) = (\mathbf{p}(m), e_i \otimes f_j) \to \widetilde{\mathbf{e}}_i(m) \otimes \widetilde{\mathbf{f}}_j(m) = \vartheta\left(\chi(m)^{-1}\right)(\mathbf{e}_i(m) \otimes \mathbf{f}_j(m)) \\ [\psi(m)] \to \left[\widetilde{\psi}(m)\right] = \vartheta(\chi(m))[\psi(m)] = [\gamma C(s)][\psi][\varrho(g)] \end{bmatrix} \tag{5.7}$$

$\widetilde{\psi}^{ij} = \sum_{k=1}^{4} \sum_{l=1}^{n} [\gamma C(s)]_k^i [\varrho(g)]_l^j \psi^{kl}$

The scalar product on $E \otimes F$ extends pointwise to $Q[E \otimes F, \vartheta]$:
$\langle \psi(m), \psi'(m) \rangle = Tr\left([\psi(m)]^* [\gamma_0][\psi'(m)]\right)$
It is preserved by ϑ .
The state of a particle along its world line is then represented by a path on the vector bundle :
$\psi(\tau) = \vartheta(\tau)\psi_0$ with $\vartheta(\tau) = \gamma C(\sigma(\tau)), \rho(\varkappa(\tau)))$ and $\psi_0 \in \widehat{E}_0 \otimes F$

$$\langle \psi, \psi \rangle = \langle \psi_0, \psi_0 \rangle = Ct \Leftrightarrow Tr\left([\psi]^* [\gamma_0][\psi]\right) = Tr\left([\psi_0]^* [\gamma_0][\psi_0]\right) \tag{5.8}$$

We will use the following bundles, which can be seen as restrictions of the previous ones :
By restriction to $\sigma = 1$ the principal bundle $Q(M, Spin(3,1) \times U, \pi_U)$ is a principal bundle with fiber U, that we denote P_U with trivialization $\varphi_U(m, \varkappa)$.
A change of trivialization with a section $\chi(m) \in \mathfrak{X}(P_U)$ induces a change of gauge, and of basis $\mathbf{f}_j(m) = (\mathbf{p}_U(m), f_j)$ in the associated vector bundle $P_U[F, \varrho]$:

$$\begin{bmatrix} \mathbf{p}_U(m) = \varphi_U(m,1) \to \widetilde{\mathbf{p}}_U(m) = \widetilde{\varphi}_U(m,1) = \mathbf{p}_U(m) \cdot \chi(m)^{-1} \\ \varkappa(m) = \varphi_U(m, \varkappa(m)) = \widetilde{\varphi}_U(m, \chi(m) \cdot \varkappa(m)) \\ \mathbf{f}_j(m) = (\mathbf{p}(m), f_j) \to \widetilde{\mathbf{f}}_j(m) = \varrho\left(\chi(m)^{-1}\right)(\mathbf{f}_j(m)) \\ \phi(m) \to \widetilde{\phi}(m) = \varrho(\chi(m))\phi(m) \end{bmatrix} \tag{5.9}$$

Momentum

The state of a particle does not represent its momentum. We can expand what we have done for spinors. The derivative reads :
$\partial_\alpha [\psi] = \partial_\alpha ([\gamma C(\sigma)][\psi_0][\varrho(g)]) = ([\gamma C(\partial_\alpha \sigma)][\psi_0][\varrho(g)]) + ([\gamma C(\sigma)][\psi_0][\varrho(\partial_\alpha g)])$
$\vartheta\left(\sigma^{-1}, g^{-1}\right)\partial_\alpha[\psi] = \left[\gamma C\left(\sigma^{-1}\partial_\alpha \sigma\right)\right][\psi_0] + [\psi_0]\left[\varrho'(1)\left(g^{-1}\partial_\alpha g\right)\right] \in E \otimes F$
and :

$$\mathcal{M} = \sum_{\alpha=0}^{3} \left(\left[\gamma C\left(\sigma^{-1}\partial_\alpha \sigma\right)\right][\psi_0] + [\psi_0]\left[\varrho'(1)\left(g^{-1}\partial_\alpha g\right)\right]\right) \otimes d\xi^\alpha \in \Lambda_1(M; E \otimes F) \tag{5.10}$$

Because $\sigma^{-1}\partial_\alpha \sigma \in T_1 Spin(3,1), g^{-1}\partial_\alpha g \in T_1 U$ the momenta can be linearly combined in $E \otimes F$.
For a photon :
$\partial_\alpha [\psi] = \partial_\alpha \left(e^{iq\phi}[\gamma C(\sigma)][S_0]\right)$
$\vartheta\left(\sigma^{-1}, e^{-iq\phi}\right)\partial_\alpha[\psi] = iq(\partial_\alpha \phi)[\gamma C(\sigma)][S_0] + \left[\gamma C\left(\sigma^{-1}\partial_\alpha \sigma\right)\right][S_0]$

5.2.3 Matter fields

The quantity ψ sums up everything (motion, kinematic, charge) about the particle. When particles are considered in a model they are naturally represented by ψ whose value can be measured at each point of its trajectory. So the most natural way to represent the particle is by a map : $\psi : [0, T] \rightarrow Q[E \otimes F, \vartheta]$ which can be parametrized either by the proper time or the time of the observer.

It is usual to consider models involving particles of the same type, submitted to similar conditions in a given area. Then, because they have the same behavior, one can assume that their trajectories can be represented by a unique vector field, and it is natural to represent their relativist momentum as sections of the fiber bundle $\mathfrak{X}(Q[E \otimes F, \vartheta])$.

Definition

Definition 95 *A **matter field** is a section $\psi \in \mathfrak{X}(Q[E \otimes F, \vartheta])$ which, at each point, represents the relativist momentum of the same particle (or antiparticle). More precisely we will assume :*

$\exists(\sigma, \varkappa) \in \mathfrak{X}(Q), \exists \psi_0 \in E \otimes F : \psi_L = \epsilon i \psi_R :: \psi(m) = \vartheta(\sigma, \varkappa)\psi_0$

$\int_\Omega \|\psi(m)\| \varpi_4(m) < \infty$

Notation 96 *$\mathfrak{X}(M)$ is the set of matter fields, $\mathfrak{X}(\psi_0)$ the set of matter fields corresponding to $\psi_0 \in E \otimes F$.*

For a matter field representing an elementary particle $\psi \in \mathfrak{F}$. For a composite, as said before, the assumption that it can be represented as a particle implies then the existence of such a fundamental state, with a variable inertial spinor for a deformable solid (but there is still some fixed S_0).

A necessary condition to be a matter field is : $\langle\psi(m), \psi(m)\rangle = Ct$.

The set of matter fields is a subset of $\mathfrak{X}(Q[E \otimes F, \vartheta])$. This is not a vector space but, as seen for the spinors, it belongs to a Hilbert space, subspace of $E \otimes F$: the scalar product is definite (positive or negative) on the vector space spanned by $\mathfrak{F}$.

Mass, Spin and Charge of a matter field

We can proceed as for spinor fields.

The space $\mathfrak{X}(Q[E \otimes F, \vartheta])$ is a functional representation of $Spin(3,1) \times U$ with the global action ϑ. The subset $\mathfrak{X}(\psi_0)$ is invariant by the right or left global actions of $Spin(3,1) \times U$. Moreover the value of $Y(m) = \langle\psi(m), \psi(m)\rangle$ is invariant by $Spin(3,1) \times U$.

The **mass of the particle** is defined as : $M_p = \frac{1}{c^2}\sqrt{|\langle\psi, \psi\rangle|}$ for $\psi \in \mathfrak{X}(\psi_0)$

The matter fields $\psi \in \mathfrak{X}(\psi_0)$ can equivalently be defined by a couple $(\psi_0, \sigma \times g)$ where $(\sigma \times g) \in \mathfrak{X}(Q)$. The representation is faithful : for given values of $\psi_0, \psi(m)$ there is a unique couple $(\sigma(m) \times g(m))$ and thus a unique $\sigma(m)$.

For a given observer $\sigma(m)$ admits two decompositions $\sigma(m) = \epsilon\sigma_w(m) \cdot \sigma_r(m)$. ϵ defines the relativist spin of the particle.

Define on $\mathfrak{X}(\psi_0)$ the equivalence relation :

$\psi \sim \psi' \Leftrightarrow \forall m \in M : \sigma_r(m) = \sigma'_r(m)$

Each class of equivalence is invariant by $Spin(3)$. The common value of $\psi = \gamma C(\sigma_r)\psi_0$ is **the Spin of the particle**, in his usual meaning. So to any given matter field correspond two Spins, with the Spin up or down.

Define on $\mathfrak{X}(\psi_0)$ the equivalence relation :

$\psi \sim \psi' \Leftrightarrow \forall m \in M : \sigma_w(m) = \sigma'_w(m)$

Each class of equivalence defines with the observer the same trajectories

Define on $\mathfrak{X}(\psi_0)$ the equivalence relation :

$\psi \sim \psi' \Leftrightarrow \forall m \in M : g(m) = g'(m)$

Each class of equivalence is invariant by U. The common value of $\psi = \varrho(g(m))\psi_0$ is **the charge of the particle**. So the charge is not expressed by a scalar.

Quantization of the Spin and Charge

The set of sections $\mathfrak{X}(\psi_0)$ is not a vector space. However there is a norm on $E \otimes F$ invariant by ϑ. The space :

$L^1 = L^1(M, Q[E \otimes F, \vartheta]) = \left\{\psi \in \mathfrak{X}(Q[E \otimes F, \vartheta]) : \int_\Omega \|\psi(m)\| \varpi_4(m) < \infty\right\}$ is a separable Fréchet vector space. And we have the following :

Theorem 97 *The set $L^1(\psi_0) = L^1(M, Q[E \otimes U, \gamma C], \varpi_4) \cap \mathfrak{X}(\psi_0)$ of integrable matter fields associated to a particle is characterized by 2 scalars : $k \in \mathbb{R}, z \in \mathbb{Z}$.*

The Spin, up or down, associated to each section is characterized by a scalar $j \in \frac{1}{2}\mathbb{N}$ and belongs to a 2j+1 dimensional vector space isomorphic to $\left(P^j, D^j\right)$

The Charge of a matter field is characterized by a scalar and belongs to a finite dimensional vector space which is invariant by U

If the section ψ is continuous then the evaluation map : $\mathcal{E}(m) : L^1(\psi_0) \to E \otimes F :: \mathcal{E}(m)\psi = \psi(m)$ is continuous

Proof. i) For the first part the proof is the same as 83.

ii) For the charge, we add to the variable $\psi \in L^1(M, Q[E \otimes U, \gamma C], \varpi_4)$ the quantity $Z = \varrho(g(m))\psi_0$. For each value of Z we have a subset of the Hilbert space which is invariant by $\widehat{U} = \Upsilon \circ U \circ \Upsilon^{-1}$, so corresponds to an irreducible representation of U. U is compact, so Z belongs to one of the irreducible representation, which is a finite dimensional vector space characterized by scalars, and characteristic of ψ_0.

iii) The space of continuous, compactly supported maps is dense in $L^1(M, E \otimes F, \varpi_4)$ (Maths.2292)

Let be ψ_n such a sequence converging to ψ in L^2

$\langle \psi - \psi_n, \psi - \psi_n \rangle(m)$ is continuous, ≥ 0 on the open Ω so there are

$A_n = \min_{m \in \Omega} \langle \psi - \psi_n, \psi - \psi_n \rangle(m)$

$\int_\Omega A_n \varpi_4 \leq \int_\Omega \langle \psi - \psi_n, \psi - \psi_n \rangle \varpi_4$

$\Rightarrow A_n \to 0$

$\Rightarrow \psi_n(m) \to \psi(m)$ ■

Experiments show that the Spin of an elementary particle is $j = \frac{1}{2}$.

The scalar z in the representation of Spin(3,1) corresponds to the charge of the particle. In each irreducible representation of U we can choose an orthonormal basis and the collection of these vectors is a basis of F, and z can label the irreducible representations. The Charge of a matter field is, as the Spin, a class of equivalence, which is represented as a section of $\mathfrak{X}(Q[E \otimes F, \vartheta])$ but the value of ψ itself is not simply a combination of the Spin and the charge. However the class of equivalence, and thus the charge, is constant.

When only the EM and gravitational fields are involved, then $E \otimes F$ is reduced to E and $z \in \mathbb{Z}$ is the charge of the particle, expressed in a suitable unit.

Density of particles

A section $\psi \in \mathfrak{X}(Q[E \otimes F, \vartheta])$ or a map : $\psi : \mathbb{R} \to Q[E \otimes F, \vartheta]$ such that $\langle \psi, \psi \rangle$ is not constant but has a constant sign can, formally, represent the state of a material body, however it is clear that the aggregation of particles has a physical meaning only for spinors (valued in E).

We can similarly consider a matter field with density μ. It has been introduced in the previous chapter to address the aggregation of particles in a solid body, but here we will consider the case of beams of particles, that is of a collection of particles of the same type which follow trajectories which do not cross. Their state can be represented by a matter field (with a constant ψ_0) and the density is then the number of particles by unit of volume. A solid body is characterized by the existence of a proper time, where the particles are localized on the space like hypersurface. A beam of particles will be defined as a collection of particles, continuously present over a period $[0, T]$ of an observer, so the time is measured with respect to an observer.

The chart which is used is the chart of the observer $\varphi_o(t, x), \varepsilon_0(m) = \partial\xi_0 = c\partial t$ and the volume is $\varpi_4 = \det[P']\, d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$.

The density is the number of particles by unit of 4 dimensional volume : on any area $\omega \subset \Omega$ this number is : $\int_\omega \mu(m)\, \varpi_4(m)$.

$\mu(m) = \sqrt{\frac{\langle\psi(m), \psi(m)\rangle}{\langle\psi_0, \psi_0\rangle}}$

The velocity of the particles, with respect to the observer, is :

$V = \frac{dp}{dt} = \overrightarrow{v} + c\varepsilon_0(m) = c\left(\left(2a_w^2 - 1\right)\varepsilon_0 + \epsilon\frac{a_w}{2a_w^2 - 1}\sum_{i=1}^3 w_i\varepsilon_i\right)$

The number of particles is constant, and the continuity equation is expressed by the Lie derivative of the volume form $\mu\varpi_4$ along the vector field V : $\pounds_V \mu\varpi_4 = 0$ (see Deformable Solids in the previous chapter).

$\pounds_V \mu\varpi_4 = 0 = \mu\pounds_V\varpi_4 + \varpi_4\pounds_V\mu = \mu(divV)\varpi_4 + \varpi_4(\mu'(V)) = \mu(divV)\varpi_4 + \frac{d\mu}{dV}\varpi_4$

$$\mu divV + \frac{d\mu}{dt} = 0 \tag{5.11}$$

Which is similar to the classic continuity equation in a fluid.

Schrödinger equation for the particles

Whenever there is a fundamental state ψ reads :

$\psi(t) = \vartheta(\sigma(t), \varkappa(t))\, \psi_0$

for a particle on its trajectory. For $t = 0 : \psi(0) = \vartheta(\sigma(0), \varkappa(0))\, \psi_0$ with known values and

$\psi(t) = \vartheta\left(\sigma(t) \cdot \sigma^{-1}(0), \varkappa(t) \cdot \varkappa(0)^{-1}\right)\psi(0)$

So : $\psi(t) = \Theta(t)\, \psi(0)$ with $\Theta(t) = \vartheta\left(\sigma(t) \cdot \sigma^{-1}(0), \varkappa(t) \cdot \varkappa(0)^{-1}\right)$

Moreover ψ belongs to a Hilbert space.

This is the GR formulation of the Schrödinger equation for particles.

In all common computations in QM, the wave function related to particles can be replaced by the state vector ψ, and the variables $r(t), w(t)$, replace the linear and rotational momentum operators. The lagrangian, notably under its perturbative form, replaces the Hamiltonian, with parameters the potential of the fields $\widehat{G}, \widehat{\grave{A}}$ along the trajectory. We have a clear mathematical framework in which the usual problems can be addressed.

The measure of observables follow the usual rules of QM, as exposed in the 2nd Chapter. In particular *stationary states* correspond to eigen vectors of the Schrödinger equation, that is $\sigma(t) = Ct$, $\varkappa(t) = Ct$. In continuous models, that we will see in the following, $r(t), w(t)$ enter into differential equations, and the "quantized states" of the particles are given by constant functions. This holds also for any observable.

With a matter field, and a density of particles :

$\psi(m) = \vartheta(\sigma(m), \varkappa(m))\, \psi_0$

the matter field itself defines a vector field for the trajectories, with respect to an observer, thus we have natural maps :

$[0,T] \to Q[E\otimes F,\vartheta] :: \psi(\Phi_V(t,x)) = \vartheta(\sigma(\Phi_V(t,x)), \varkappa(\Phi_V(t,x)))\psi_0$

which represent the evolution of the state of a given particle (labeled by x on $\Omega_3(0)$) along its trajectory, and indeed the implementation of the Principle of Least Action provides differential equations along the trajectory. So we still have :

$\psi(\Phi_V(t,x)) = \Theta(t)(\Phi_V(0,x))$

The stronger version of the theorem 27 requires that ψ is defined all over $\mathbb{R}$: the evolution of the particle is in a continuous process, with no beginning or ending, and this assumption is crucial. An important special case is of bonded particles in a regular environment, such as a crystal : it can then be assumed that $\psi(m)$ is a periodic map over a lattice defined by the geometric structure of the medium. The observer is then defined with respect to this lattice (which sums up to choose a suitable chart of $\Omega(0)$). The value of the potentials is defined in this chart.

Wave function

Usually one has a collection of particles of different types observed in a domain Ω, the goal of the experiment is to know the type and the motion of the particles. The states of the particles are represented by a unique section : $\psi \in L^1(M, Q[E\otimes F,\vartheta])$ and a primary observable is a linear map $\Phi : L^1(M, Q[E\otimes F,\vartheta]) \to V :: \Phi(\psi) = Y$ where V is a finite dimensional vector space. The observable can address some features of the particles only (such as the nature of the particles, their spin or charge,...).

There is a Hilbert space H associated to $L^1(M, Q[E\otimes F,\vartheta])$. This is an infinite dimensional, normed and separable vector space, and $E\otimes F$ is finite dimensional. The evaluation map $\mathcal{E}(m) : L^1(\psi_0) \to E\otimes F :: \mathcal{E}(m)\psi = \psi(m)$ is continuous. To Φ is associated the self adjoint operator $\widehat{\Phi} = \Upsilon \circ \Phi \circ \Upsilon^{-1}$ on H.

We can apply the theorem 19. For any state ψ of the system there is a function : $W : M \times E\otimes F \to \mathbb{R}$ such that $W(m,Y) = \Pr(\Phi(\psi)(m) = y|\psi)$ is the probability that the measure of the value of the observable $\Phi(\psi)$ of ψ at m is y. It is given by :

$\Pr(\Phi(\psi)(m) = y|\psi) = \frac{1}{\|\Upsilon(\psi)\|_H^2}\int_{Y\in\varpi(m,y)} \left\|\widehat{\Phi}(\Upsilon(Y))\right\|_H^2 \pi(Y) = W(m,y)$

This can be seen as a density of probability, corresponding to the square of a wave function.

Of particular interest is the observable $\Phi(\psi) = \langle\psi,\psi\rangle$ which can be seen as the identification of the particles. The choice of the observable cannot be seen any longer as random. However one can assume that the choice of the point m is random. L^1 is partitioned in subsets $L^1(\psi_0)$ and any section ψ can be written as : $\psi(m) = \sum_j \varpi_j(m)\psi_j(m)$ where $\psi_j \in L^1(\psi_{0j})$ and $\varpi_j(m)$ is the characteristic function of the domain of ψ_j. Then the probability : $\Pr(\langle\psi(m),\psi(m)\rangle = \langle j,j\rangle|\psi) = \left(\int_\Omega \varpi_4\right)^{-1}\int_\Omega \varpi_j\varpi_4$

Difference with the classic QTF interpretation

So our picture provides, in a classic interpretation, the tools used in Quantum Physics : to each particle is associated a state, valued in a vector space, a spin, a mass and a charge, which follows the Schrödinger equation and can be measured through observables and wave functions. A matter field is the representation of a population of virtual particles and can be used in any model. The difference with the usual picture of the Quantum Theory of Fields (QTF) comes from the interpretation of "virtuality".

In our picture, let be N particles with the same fundamental state ψ_0 for which the state $\psi_p(\varphi_o(0,x_p))$ at $t=0$ is known. A necessary condition for a section $\mathbf{\Psi}$ of $\mathfrak{X}(\psi_0)$ to represent the states of the N particles is that $\mathbf{\Psi}(\varphi_o(0,x_p)) = \psi_p(\varphi_o(0,x_p))$. The section $\mathbf{\Psi}$ defines a vector field corresponding to the possible trajectories, and the trajectory of each particle is given by the integral curve of this vector field passing by $\varphi_o(0, x_p)$ and the state of the particle is given by $\psi_p(\varphi_o(t, x_p(t))) = \mathbf{\Psi}(\varphi_o(t, x_p(t)))$. A section meeting the necessary condition is not unique, and the evolution laws of the system including the particles select a general solution, which must then meets the initial conditions. However it is convenient to use $\mathbf{\Psi}$ as variable in a model : it provides a specification for the solution. A matter field can be considered as a collection over Ω of tests particles, or as **virtual particles**. But a *single* section represents any number of particles. In this picture operators are linear maps which act on the section $\mathbf{\Psi}$, and not locally at each point. They provide wave functions as shown above, if necessary, which have a mathematical, statistical interpretation, without physical meaning. And a virtual particle is just a blueprint for the representation of real particles which, alone, interact.

In a strict interpretation of standard QM a physical object has no property until a measure of this property has been done. This is true of any property, including the location. This is a bit awkward because the observable usually associated to the location (a spatial or temporal coordinate) is not compact : its spectrum is continuous and cannot provide a precise answer (meanwhile any quantity related to the SO(3) group, which is compact, provides a set of fixed solutions). There are many subtle or less subtle (such as the recurring usage of Dirac's function which are, usually, nothing more than a mathematical tautology) solutions to circumvent the problem, but the main consequence is that to *each* particle is associated a different section of $\mathfrak{X}(\psi_0)$: a given particle can be present everywhere. Then an observable becomes an operator which acts locally in the local Hilbert space in which can live the state of the particle (and not on the maps as in our picture). A complication arises from the fact that now many particles can potentially be at the same location. There are some restrictions but, as a consequence the Hilbert space to consider is the tensorial product of Hilbert spaces, and as the number of particles is not fixed, the structure involved is a Fock space (the sum $\oplus_{n=0}^{\infty} \otimes^n H$) . This is actually the tool to study the creation and annihilation of particles in discontinuous processes, but is, as one can guess, inappropriate for continuous processes. Virtual particles become even more virtual : they are just collections of tests functions used to define distributions.

So, overall, our picture provides a representation which is consistent with Classic Physics and account for the usual features of Quantum Physics.

5.3 CONNECTIONS

The field changes the representation of the states of particles, but its action is not static and limited to ϑ. The propagation of the fields as well as their interaction with particles are, as usually, seen through infinitesimal changes and differential equations. So we need a mathematical way to define the derivative of $\mathbf{q}_f(m) = \varphi_Q(m, (\sigma(m), \varkappa(m)))$. Because of the anisotropy of the universe, the value of the derivative will depend on the direction on M, represented by a vector, so we are looking for a map : $M \to \Lambda_1(T_1Spin(3,1) \times T_1U)$,that is a one form valued in the Lie Algebra. This derivative is the **covariant derivative**. The action of a field on a particle, usually represented by a force $\overrightarrow{F} = \frac{d\overrightarrow{p}}{dt}$ which is the derivative of the momentum, is replaced by the covariant derivative $\nabla_V\psi$ of the state along the direction V.

Covariant derivatives are built from more general mathematical objects, called connections. It will be necessary to refer to them, so it is better to introduce them from the start. This can be done in two complementary ways, geometric or through differential operators, and we need to remind some more mathematics.

5.3.1 Connections in Mathematics

Geometric connections

A fiber bundle $P(M, V, \pi)$ is a manifold, and its tangent space is split in two parts, related to its two manifolds components. By differentiation of the trivialization :

$\varphi : M \times V \to P :: p = \varphi(m, u)$

$\varphi' : T_mM \times T_uV \to T_pP :: v_p = \varphi'_m(m, g)v_m + \varphi'_u(m, g)v_u$

$\pi(p) = m \Rightarrow \pi'(p)v_p = v_m$ and the vector subspace $V_pP = \{\pi'(p)v_p = 0\}$ of T_pP called the **vertical space** does not depend on the trivialization. It is isomorphic to the tangent space of V.

Our purpose is to look for a way to define a derivative of p, and the decomposition of the vector v_p shows that it requires two components : one linked to a motion in M, and another to a change in V. However, even if $\pi'(p)v_p = v_m$,this is not sufficient to define a decomposition which would be independent on the choice of a trivialization. A **connection** is just this tool : it is a projection of v_p on the vertical space V_pP. It is a one form on P valued in the vertical bundle VP. So it enables us to distinguish in a variation of p what can be imputed to a change of m and what can be imputed to a change of u. A section of P depends only on m : $\mathbf{p}(m) = \varphi(m, u(m))$ so by differentiation with respect to m this is a map from TM to TP and the value of a connection at each $\mathbf{p}(m)$ is a one form over M, valued in VP, called the **covariant derivative**. So it meets our purpose. Moreover because the vertical space is isomorphic to the tangent space on V, the value of the connection can be expressed in a simpler vector space.

The covariant derivative issued from a linear connection on a vector bundle $P(M, V, \pi)$ reads:

$\nabla X = \sum_{\alpha=0}^{3}\sum_{a=1}^{m}\left(\partial_\alpha X^i(m) + \Gamma_{\alpha i}^j(m)X^j(m)\right)e_i \otimes d\xi^\alpha$

where $\Gamma_{\alpha i}^j(m)$ is the Christoffel symbol of the connection. Its meaning is clear : the covariant derivative adds a linear combination of X to the derivative of X, this is the simplest form for the definition of a derivative on a fiber bundle. Readers who are familiar with GR are used to Christoffel symbols, and their definition through the metric. We will see how it works.

This holds for any kind of fiber bundle (Maths.27), but the connection takes different forms according to the kind of fiber bundle. With a principal bundle one can define many others fiber bundles by association and similarly a connection on a principal bundle defines a connection on any associated bundle. So connections on principal bundles have a special importance.

The second way to define a connection is through differential operators, but for this we need to say a bit about r-jets.

r-jet extensions

In Differential Geometry one tries to avoid as much as possible the coordinates expressions. But when one deals with derivatives of orders higher than one this becomes impossible. It is always difficult to deal with partial derivatives, and notably to define the set in which they belong. The r-jet formalism provides a convenient solution. See Maths.26 for more.

For any r differentiable map $f \in C_r(M;N)$ between manifolds, the partial derivatives $\frac{\partial^s f}{\partial \xi^{\alpha_1} ... \partial \xi^{\alpha_s}}$ are s symmetric linear maps from the tangent space $T_m M$ to the tangent space $T_p N$. Their expression in holonomic bases is a set of components $f^i_{\alpha_1 ... \alpha_s}$ symmetric in the indices $\alpha_1, ..\alpha_s$. It defines fully the map $\frac{\partial^s f}{\partial \xi^{\alpha_1} ... \partial \xi^{\alpha_s}}$ at m, and the set

$\left(p, f(p), f^i_{\alpha_1 ... \alpha_s}, \alpha_k = 1.. \dim M, i = 1.. \dim N\right)$ is denoted $J^s_m(M,N)_{f(m)}$. Conversely one can conceive the same set of scalars and points of M, N, without any reference to the map f. It defines the class of equivalence of all maps which have the same derivatives up to order r, and is called the r-jet prolongation of $C_r(M;N)$. When one forgets the origin m and target $f(m)$ the set $J^s(M,N)$ is a vector space, and an affine space if one forgets only $f(m)$.

The formalism can be extended to fiber bundles, by replacing the maps f with sections on a fiber bundle. The **r-jet extension** $J^r P$ of the fiber bundle $P(M,V,\pi)$ is the r-jet prolongation of its sections. This is a closed manifold of the vector space $J^r(M,V)$. A map which associates to each $m \in M$ a set of values $Z = \left(m, z^i_{\alpha_1 ... \alpha_s}, i = 1...n, s = 0, ..., r\right)$ (called coordinates of the jet) in $J^r P$ is a **r-jet**, denoted j^r_m. A key point is that *there is a priori no relation between the* $z^i_{\alpha_1 ... \alpha_s}(m)$: they do not correspond necessarily to the derivatives of the same map. But conversely a given section S of P provides a r-jet, denoted $J^r S$. A change of trivialization on the fiber bundle entails relations, usually complicated, between the elements $z^i_{\alpha_1 ... \alpha_s}$, so they are not totally arbitrary.

The r-jet prolongation of a principal bundle is a principal bundle (but with a more complicated group), and the r-jet prolongation of an associated vector bundle is an associated vector bundle. Its elements have the coordinated expressions : $Z = \left(z^i_{\alpha_1 ... \alpha_s}, i = 1...n, s = 0, ..., r\right)$. The index i refers to the component in the vector space V, the indices $\alpha_1, ...\alpha_s$ to the partial derivatives with respect to the coordinates in M.

The principal application of the r-jet formalism is in Differential Equations and Differential Operators.

Covariant derivatives as differential operators

A r differential operator is a base preserving morphism $D : \mathfrak{X}(J^r E_1) \rightarrow \mathfrak{X}(E_2)$ between two fiber bundles (Maths.32). It maps fiberwise $Z(m)$ in $J^r E_1$ to $Y(m)$ in E_2.It is local : its computation involves only the values at m, and provides a result at m. By itself D does not involve any differentiation (it is defined for any section of the r-jet bundle $J^r E_1$). Combined with the map : $J^r : \mathfrak{X}(E_1) \rightarrow \mathfrak{X}(J^r E_1)$, $D \circ J^r$ maps sections on E_1, to sections on E_2. This distinction is useful : differential equations are equations between sections of $J^r E_1$ (the partial derivatives are, a priori, independent variables) and solutions are sections such that the components of the r-jet are deduced from the same map by differentiation. The most general mathematical definition of a set of differential equations of order r between variables defined on some fiber bundle E is that it is a closed subbundle of $J^r E$. The solutions are sections of E such that their r-jet extensions belong to this closed subbundle.

A linear r-differential operator is a linear, base preserving morphism, between two vector bundles (associated or not to a principal bundle, this does not matter here) : $E_1(M, V_1, \pi_1), E_2(M, V_2, \pi_2)$. The coordinates of a section $Z \in \mathfrak{X}(J^r E_1)$ read : $Z = \left(m, z^i_{\alpha_1...\alpha_s}, i = 1...n, s = 0, ..., r\right)$ and DZ reads :

$DZ = \sum_{s=0}^r \sum_{\alpha_i=1}^m \sum_{i=1}^n \sum_{j=1}^p A(m)_i^{j,\alpha_1...\alpha_s} z^i_{\alpha_1...\alpha_s}(m) e_{2j}(m)$

with a basis $(e_{2j}(m))_{j=1}^p$of E_2 , scalars $A(m)_i^{j,\alpha_1...\alpha_s}$, and for a section $Z \in \mathfrak{X}(E_1)$: $z^i_{\alpha_1...\alpha_s}(m) = \frac{\partial^s z^i}{\partial\xi^{\alpha_1}...\partial\xi^{\alpha_s}}$

In this framework it is easy to study the properties of Differential Operators such as action on distributions, adjoint of an operator, symbol, Fourier transform... We are concerned here with 1st order operators on a principal bundle.

A covariant derivative can then be considered as a differential operator (Maths.32.2.7) and indeed this is the way it is introduced in many books. For a vector bundle :

$DZ = \sum_{\alpha=1}^m \sum_{i=1}^n \left(z^i_\alpha + \sum_{j=1}^p \Gamma^j_{\alpha i} z^i(m)\right) e_{2j}(m) \otimes d\xi^\alpha$

and similarly for an associated vector bundle.

The Differential Operator approach is useful when one considers higher order derivatives (one can define higher order covariant derivative)

5.3.2 Connection for the force fields other than Gravity

Connection on the principal bundle P_U

Tangent space

Its tangent space is given by vectors :

$v_p = \varphi'_{Um}(m,g) v_m + \varphi'_{Ug}(m,g) v_g$

The vertical space $VP_U = \ker \pi'_U = \left\{\varphi'_{Ug}(m,g) v_g, v_g \in T_\chi U\right\}$ is isomorphic to the Lie algebra.

The map :

$\zeta : T_1U \to VP :: \zeta(\theta)(\varphi_U(m,g)) = \varphi'_{Ug}(m,g) L'_{g^{-1}}g(\theta)$ where $L'_g(1)$ is the derivative of the left translation : $L : G \times G \to G :: L(g,h) = g \cdot h$

is linear, does not depend of the trivialization and has the property :

$\zeta(\theta)(\rho(p,g)) = \rho'_p(p,g) \zeta(Ad_g\theta)(p)$

For a given $\theta \in T_1U$ the vector field $\zeta(\theta)(p) = \varphi'_{Ug}(m,g) L'_g(1)\theta \in V_pP_U$ is called a **fundamental vector.**

If is convenient to write vectors as :

$v_p = \varphi'_{Gm}(m,g) v_m + \varphi'_{G\varkappa}(m,g) v_g = \sum_{\alpha=0}^3 v_m^\alpha \partial m_\alpha + \zeta(\theta)(p)$ with $\theta = L'_{g^{-1}}g(v_g)$

Connection on the Principal bundle

A **connection**is a tensor, a one form $\grave{\mathbf{A}} \in \mathbf{\Lambda}_1(TP_U; VP_U)$ on TP valued in VP :

$\grave{\mathbf{A}}(p)\left(\sum_{\alpha=0}^3 v_m^\alpha \partial m_\alpha + \zeta(\theta)(p)\right) = \zeta\left(\theta + L'_{g-1}g\left(\sum_{\alpha=0}^3 v_m^\alpha \grave{A}_\alpha(p)\right)\right)(p)$

A connection $\grave{\mathbf{A}} \in \mathbf{\Lambda}_1(TP_U; VP_U)$ is **principal** if it is equivariant by the right action :

$\forall p, g : \rho(p,g)^* \grave{\mathbf{A}}(p) = \grave{\mathbf{A}}(\rho(p,g)) \rho'_p(p,g) = \rho'_p(p,g) \grave{\mathbf{A}}(p)$

where $\rho(p,g)^*$ is the pull-back of **.**[3]

Its value for any gauge of P_U can be defined through its value for $\mathbf{p} = \varphi_U(m,1)$

$\grave{\mathbf{A}}(\mathbf{p}(m))(\varphi'_m(m,1) v_m + \zeta(\theta)(\mathbf{p}(m))) = \zeta\left(\theta + \sum_\alpha \grave{A}_\alpha(m) v_m^\alpha\right)(\mathbf{p}(m))$

where $\grave{A}$, the **potential** of the connection, is a map valued in the fixed vector space T_1U. :

$$\grave{A} \in \Lambda_1 (M; T_1U) : TM \to T_1U :: \grave{A}(m) = \sum_{\alpha=0}^{3} \sum_{a=1}^{m} \grave{A}_\alpha^a (m) \overrightarrow{\theta}_a \otimes d\xi^\alpha \tag{5.12}$$

In a change of gauge
$\mathbf{p}_U (m) \to \widetilde{\mathbf{p}}_U (m) = \widetilde{\varphi}_U (m, 1) = \mathbf{p}_U (m) \cdot \chi (m)^{-1}$:
$\grave{A}$ changes with an *affine law*, which involves the derivative $\chi' (m)$ of the change of gauge :
$\grave{A}(m) \to \widetilde{\grave{A}}(m) = Ad_\chi \left(\grave{A}(m) - L'_{\chi^{-1}} (\chi) \chi' (m) \right)$
and this feature is at the origin of many specificities (and complications, such as the Higgs boson...).

Covariant derivative on P_U

The covariant derivative of a section $\mathbf{p}_g = \varphi_U (m, g(m)) \in \mathfrak{X}(P_U)$ is then :

$$\nabla^U \mathbf{p}_g = \left(L'_{g^{-1}} g \right) (g'(m)) + \sum_{\alpha=0}^{3} Ad_{g^{-1}} \grave{A}_\alpha (m) d\xi^\alpha \in \Lambda_1 (M, T_1U) \tag{5.13}$$

which can also be written : $\mathbf{S}^* \mathbf{p}_g = \zeta \left(\nabla^U \mathbf{p}_g \right) (\mathbf{p}_g (m))$
and for the holonomic gauge : $\mathbf{p}_U = \varphi_U (m, 1) : \nabla^U \mathbf{p}_U = \sum_{\alpha=0}^{3} \grave{A}_\alpha (m) d\xi^\alpha$
The covariant derivative is invariant in a change of gauge :
$\mathbf{p}_U (m) \to \widetilde{\mathbf{p}}_U (m) = \widetilde{\varphi}_U (m, 1) = \mathbf{p}_U (m) \cdot \chi (m)^{-1}$
$\grave{A}(m) \to \widetilde{\grave{A}}(m) = Ad_\chi \left(\grave{A}(m) - L'_{\chi^{-1}} (\chi) \chi' (m) \right)$
$\nabla^U \mathbf{p}_g \to \widetilde{\nabla^U \mathbf{p}_g} = \nabla^U \mathbf{p}_g$

Connection on the associated bundles

With the connection on P_U it is possible to define a linear connection and a covariant derivative ∇^F, 1 form on M acting on sections $\phi (m) = \sum_{j=1}^{n} \phi^j (m) \mathbf{f}_j (m)$ of the associated vector bundle $P_U [F, \varrho]$ (Maths.27.4.2) :

$$\nabla^F \phi = \sum_{\alpha=0}^{3} \left(\partial_\alpha \phi^i + \sum_{i=1}^{n} \left[\grave{A}_\alpha \right]_j^i \phi^j \right) \mathbf{f}_i (m) \otimes d\xi^\alpha \in \Lambda_1 (M, P_U [F, \varrho]) \tag{5.14}$$

with the

Notation 98 $\left[\grave{A}_\alpha \right] = \sum_{a=1}^{m} \grave{A}_\alpha^a [\theta_a]$ *is a* $n \times n$ *matrix representing* $\varrho' (1) \grave{A}_\alpha^a \in \mathcal{L}(F; F)$

and $\left[\grave{A}_\alpha \right]_j^i$ has the same meaning as the Christoffel symbol Γ of a linear connection.

A covariant derivative, when acting on a vector field $u \in TM$, becomes a section of the vector bundle $P_U [F, \rho]$, and transforms as such in a change of trivialization, so we have a map : $\mathfrak{X}(P_U [F, \rho]) \times \mathfrak{X}(TM) \to \mathfrak{X}(P_U [F, \rho])$. It meets our goal, and it can be proven than this is the only way to achieve it.

Practically this is the potential which represents the field. There has been some questions about the physical meaning of the potential. However some experiments such as Aharonov-Bohm's shows that, at least for the electromagnetic field, the potential is more than a simple formalism.

In QTF, because the groups are comprised of matrices with complex coefficients, and the elements of the Lie algebra T_1U are operators in the Hilbert spaces, it is usual to introduce the imaginary i everywhere, and to consider the complexified of the Lie algebra T_1U. However it is clear that the potential $\grave{A}_\alpha$ *belongs to the real algebra, so it is a real quantity.* And there are as many force carriers bosons (12) as the dimension of U.

The electromagnetic field

The Lie algebra of $U(1)$ is $\mathbb{R}$. So the potential $\grave{A}$ of the connection is a real valued one form on M : $\grave{A} = \sum_{\alpha=0}^{3} \grave{A}_\alpha d\xi^\alpha \in \Lambda_1(M;\mathbb{R})$ which is usually represented as a vector field and not a form.

$\psi = e^{iq\varphi}\gamma C(\sigma) S_0$ can be represented through the spinor and a charge q expressed as a signed integer multiple of the negative elementary charge, and $q = 0$ for neutral particles.

The action of $U(1)$ depends on the representation, thus on the charge of the particle :

- negative charge : $\varrho(\varkappa)\psi = (\exp i\varkappa)\psi \Rightarrow \varrho'(1)\overrightarrow{\theta} = i\overrightarrow{\theta}$
- positive charge : $\varrho(\varkappa)\psi = (\exp(-i\varkappa))\psi \Rightarrow \varrho'(1)\overrightarrow{\theta} = -i\overrightarrow{\theta}$
- neutral : $\varrho(\varkappa)\psi = \psi \Rightarrow \varrho'(1) = 0$

So the covariant derivative reads :

$$\nabla_\alpha^F \psi = \partial_\alpha \psi + qi\grave{A}_\alpha \psi \tag{5.15}$$

5.3.3 The connection of the gravitational field

Potential

The principles are similar. The vertical bundle VP_G of the principal bundle $P_G(M, Spin(3,1), \pi_G)$ is isomorphic to the Lie algebra $T_1Spin(3,1)$.

The potential G of a principal connection $\mathbf{G}$ on P_G is a map : $G \in \Lambda_1(M; T_1Spin(3,1))$.

Using the Clifford algebra to represent the Lie algebra, G reads :

$$\left[\begin{array}{l} G \in \Lambda_1(M; T_1Spin(3,1)) : TM \to T_1Spin(3,1) :: \\ G(m) = \sum_{a=1}^{6}\sum_{\alpha=0}^{3} G_\alpha^a(m)\overrightarrow{\kappa}_a \otimes d\xi^\alpha = \sum_{\alpha=0}^{3} \upsilon(G_{r\alpha}(m), G_{w\alpha}(m)) d\xi^\alpha \end{array}\right] \tag{5.16}$$

$G_{r\alpha}(m), G_{w\alpha}(m)$ are two vectors $\in \mathbb{R}^3$. So the *gravitational field has a transversal* $(G_{w\alpha})$ *and a rotational* $(G_{r\alpha})$ *component.* This is the unavoidable consequence of the gauge group.

$G_r(m) = \sum_{\alpha=0}^{3} \upsilon(G_{r\alpha}(m), 0) d\xi^\alpha$ is a map $G \in \Lambda_1(M; T_1Spin(3)) : TM \to T_1Spin(3)$

In a change of gauge the potential transforms by an affine map :

$\mathbf{p}(m) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m) \cdot \chi(m)^{-1} : G(m) \to \widetilde{G}(m) = \mathbf{Ad}_\chi\left(G(m) - L'_{\chi^{-1}}(\chi)\chi'(m)\right)$

We introduce the convenient notation that will be used in the following :

Notation 99 $\upsilon\left(\widehat{G}_r(\tau), \widehat{G}_w(\tau)\right)$ *is the value of the potential of the gravitational field along the integral curve* $m(\tau) = \Phi_V(\tau, x)$ *of any vector field* V

$\upsilon\left(\widehat{G}_r(\tau), \widehat{G}_w(\tau)\right) = \sum_{\alpha=0}^{3} V^\alpha \upsilon(G_{r\alpha}(\tau), G_{\alpha w}(\tau))$

And similarly : $\widehat{\grave{A}} = \sum_{\alpha=0}^{3} V^\alpha \grave{A}_\alpha^a \overrightarrow{\theta}_a$

There are several covariant derivatives deduced from this connection.

Covariant derivative on P_G

The connection acts on sections of the principal bundle, and the covariant derivative of $\sigma = \varphi_G(m, \sigma(m)) \in \mathfrak{X}(P_G)$ is (see general formula in the previous section) :

$$\left[\begin{array}{l} \nabla^G : \mathfrak{X}(P_G) \to \Lambda_1(M; T_1 Spin) :: \\ \nabla^G \sigma = \sigma^{-1} \cdot \sigma' + \mathbf{Ad}_{\sigma^{-1}} G = \left(\sum_{\alpha=0}^{3} \sigma^{-1} \cdot \partial_\alpha \sigma + \mathbf{Ad}_{\sigma^{-1}} G_\alpha\right) d\xi^\alpha \end{array}\right] \tag{5.17}$$

The covariant derivative is invariant in a change of gauge :
$\mathbf{p}(m) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m) \cdot \chi(m)^{-1}$
$\nabla^G \sigma \to \widehat{\nabla^G} \sigma = \nabla^G \sigma$
The explicit formula is the following :
$\nabla^G \sigma = = \upsilon(X_\alpha, Y_\alpha)$ with
$X_\alpha = [C(r)]^t \left(([D(r)] \partial_\alpha r + \frac{1}{2} j(w) \partial_\alpha w + [A(w)] G_{r\alpha} + [B(w)] G_{w\alpha}\right)$
$Y_\alpha = [C(r)]^t \left(\frac{1}{4a_w}[4 - j(w) j(w)] \partial_\alpha w - [B(w)] G_{r\alpha} + [A(w)] G_{w\alpha}\right)$
$[C(r)] = \left[1 + a_r j(r) + \frac{1}{2} j(r) j(r)\right]$
$[D(r)] = \left[\frac{1}{a_r} + \frac{1}{2} j(r) + \frac{1}{4a_r} j(r) j(r)\right]$ and $[D(r)] = [C(r)][D(r)]^t$
$[A(w)] = \left[1 - \frac{1}{2} j(w) j(w)\right]$
$[B(w)] = a_w [j(w)]$

Proof. $\sigma = \sigma_w \cdot \sigma_r$
We have seen the formula previously :
$\sigma^{-1} \cdot \partial_\alpha \sigma = \upsilon\left(\left([C(r)]^t \left(\frac{1}{2} j(w) \partial_\alpha w\right) + [D(r)]^t \partial_\alpha r\right), [C(r)]^t \left(\frac{1}{4a_w}[4 - j(w) j(w)] \partial_\alpha w\right)\right) S_0$
$\mathbf{Ad}_{\sigma^{-1}} G_\alpha = \mathbf{Ad}_{\sigma^{-1}} \upsilon(G_{r\alpha}, G_{w\alpha}) = \mathbf{Ad}_{\sigma_r^{-1}} \mathbf{Ad}_{\sigma_w^{-1}} \upsilon(G_{r\alpha}, G_{w\alpha})$
$\left[\mathbf{Ad}_{\sigma_r^{-1}}\right] = [\mathbf{Ad}_{\sigma_r}]^t = \begin{bmatrix} [C(r)]^t & 0 \\ 0 & [C(r)]^t \end{bmatrix}$
$\left[\mathbf{Ad}_{\sigma_w^{-1}}\right] = [\mathbf{Ad}_{\sigma_w}]^t = \begin{bmatrix} A & B \\ -B & A \end{bmatrix}$
$\mathbf{Ad}_{\sigma^{-1}} G_\alpha = \begin{bmatrix} C^t & 0 \\ 0 & C^t \end{bmatrix} \begin{bmatrix} A & B \\ -B & A \end{bmatrix} \begin{bmatrix} G_{r\alpha} \\ G_{w\alpha} \end{bmatrix} = \begin{bmatrix} C^t A G_{r\alpha} + C^t B G_{w\alpha} \\ -C^t B G_{r\alpha} + C^t A G_{w\alpha} \end{bmatrix}$
Thus :
$\nabla_\alpha \sigma = \left(\sigma^{-1} \partial_\alpha \sigma + \mathbf{Ad}_{\sigma^{-1}} G_\alpha\right) = \upsilon(X_\alpha, Y_\alpha)$
$X_\alpha = [C(r)]^t \left(\left(\left[\frac{1}{a_r} + \frac{1}{2} j(r) + \frac{1}{4a_r} j(r) j(r)\right] \partial_\alpha r + \frac{1}{2} j(w) \partial_\alpha w + [A(w)] G_{r\alpha} + [B(w)] G_{w\alpha}\right)$
$Y_\alpha = [C(r)]^t \left(\frac{1}{4a_w}[4 - j(w) j(w)] \partial_\alpha w - [B(w)] G_{r\alpha} + [A(w)] G_{w\alpha}\right)$ ■

Covariant derivative for spinors

The covariant derivative reads for a section $\mathbf{S} \in \mathfrak{X}(P_G[E, \gamma C])$:

$$\nabla^S S = \sum_{\alpha=0}^{3} (\partial_\alpha S + \gamma C(G_\alpha) S) d\xi^\alpha = \sum_{\alpha=0}^{3} (\partial_\alpha S + \gamma C(\upsilon(G_{r\alpha}, G_{w\alpha})) S) d\xi^\alpha \tag{5.18}$$

With the signature (3,1) :

$$\gamma C\left(\upsilon\left(G_{r\beta},G_{w\beta}\right)\right)=-i\frac{1}{2}\sum_{a=1}^{3}\left(G_{w\beta}^{a}\gamma_{a}\gamma_{0}+G_{r\beta}^{a}\widetilde{\gamma}_{a}\right) \tag{5.19}$$

With the signature (1,3) :

$$\gamma C'\left(\upsilon\left(G_{r\beta},G_{w\beta}\right)\right)=-i\frac{1}{2}\sum_{a=1}^{3}\left(G_{w\beta}^{a}\gamma_{a}\gamma_{0}-G_{r\beta}^{a}\widetilde{\gamma}_{a}\right) \tag{5.20}$$

So we go from the signature (3,1) to (1,3) by a change of the sign of $G_{r\alpha}$.

G_α being valued in $T_1Spin\left(3,1\right)$ and γC being a representation of the Clifford algebra the expression makes sense. Its coordinates expression is with right and left chiral parts:

$\nabla^S S=\sum_{\alpha=0}^{3}\left[\begin{array}{c}\partial_\alpha S_R+\frac{1}{2}\sigma\left(G_{w\alpha}-iG_{r\alpha}\right)S_R\\ \partial_\alpha S_L-\frac{1}{2}\sigma\left(G_{w\alpha}+iG_{r\alpha}\right)S_L\end{array}\right]d\xi^\alpha$

It preserves the chirality.

In a change of gauge :

$\mathbf{p}\left(m\right)\rightarrow\widetilde{\mathbf{p}}\left(m\right)=\mathbf{p}\left(m\right)\cdot\chi\left(m\right)^{-1}$

a section on $\mathfrak{X}\left(P_G\left[E,\gamma C\right]\right)$ transforms as : $\widetilde{S}\left(m\right)=\gamma C\left(\chi\left(m\right)\right)S\left(m\right)$

The covariant derivative transforms as a section of $P_G\left[E,\gamma C\right]:\nabla^S S\rightarrow\widetilde{\nabla^S S}=\gamma C\left(\chi\right)\nabla^S S$ so the operator reads: $\nabla^S:\mathfrak{X}\left(P_G\left[E,\gamma C\right]\right)\rightarrow *_1\left(M;\mathfrak{X}\left(P_G\left[E,\gamma C\right]\right)\right)$

Covariant derivatives for vector fields on M

The connection on P_G induces a linear connection ∇^M on the associated vector bundle $P_G\left[\mathbb{R}^4,\mathbf{Ad}\right]$, which is TM with orthonormal bases, with Christoffel symbols :

$\Gamma_M\left(m\right)=\left(\mathbf{Ad}_s\right)'_{s=1}\left(G\left(m\right)\right)$

with the product of vectors in $Cl(m)$:

$v=\sum_{j=0}^{3}v^j\varepsilon_j\left(m\right)\rightarrow$

$\sum_{i,j=0}^{3}\left[\Gamma_G\left(m\right)\right]_i^j v^i\varepsilon_j\left(m\right)=\upsilon\left(G_{r\alpha},G_{w\alpha}\right)\cdot v-v\cdot\upsilon\left(G_{r\alpha},G_{w\alpha}\right)$

It is then more convenient to use the representation of $T_1Spin(3,1)$ by matrices of $so(3,1)$:

$\left[\Gamma_{M\alpha}\right]=\sum_{a=1}^{6}G_\alpha^a\left[\kappa_a\right]=\begin{bmatrix}0&G_{w\alpha}^1&G_{w\alpha}^2&G_{w\alpha}^3\\G_{w\alpha}^1&0&-G_{r\alpha}^3&G_{r\alpha}^2\\G_{w\alpha}^2&G_{r\alpha}^3&0&-G_\alpha^1\\G_{w\alpha}^3&-G_{r\alpha}^2&G_{r\alpha}^1&0\end{bmatrix}$

In a change of gauge :

$G\left(m\right)\rightarrow\widetilde{G}\left(m\right)=Ad_\chi\left(G\left(m\right)-L'_{\chi^{-1}}\left(\chi\right)\chi'\left(m\right)\right)$

$\left[\widetilde{\Gamma}_{M\alpha}\right]=\left[h\left(s\right)\right]\left(\left[\Gamma_{M\alpha}\right]-\left[h\left(s^{-1}\right)\right]\left[h\left(s'\right)\right]\right)$

The covariant derivative of a section $V\in\mathfrak{X}\left(P_G\left[\mathbb{R}^4,\mathbf{Ad}\right]\right)$ is then :

$$\nabla^M V=\sum_{\alpha i=0}^{3}\left(\partial_\alpha V^i+\sum_{j=0}^{3}\left[\Gamma_{M\alpha}\left(m\right)\right]_j^i V^j\right)\varepsilon_i\left(m\right)\otimes d\xi^\alpha \tag{5.21}$$

For any vector field $W:\nabla_W^M:\mathfrak{X}\left(TM\right)\rightarrow\mathfrak{X}\left(TM\right)$ is a linear map which preserves the scalar product of vectors (Maths.2205):

$\left\langle\nabla_W^M U,\nabla_W^M V\right\rangle=\left\langle U,V\right\rangle$

$P_G\left[E,\gamma C\right]$ is a spin bundle, and we have the identity between the derivatives :

$\forall V\in\mathfrak{X}\left(P_G\left[\mathbb{R}^4,\mathbf{Ad}\right]\right),S\in\mathfrak{X}\left(P_G\left[E,\gamma C\right]\right):$

$\nabla\left(\gamma C\left(V\right)S\right)=\gamma C\left(\nabla^{M}V\right)S+\gamma C\left(V\right)\nabla S$
which makes of **G** a Clifford connection (Maths.2207).

5.3.4 Geodesics

There are several definitions of Geodesics, which, in different formulations, mean the curves of minimum length between two points. In Euclidean Geometry they are straight lines, in GR they are usually curves, and they play an important role because free particles move along geodesics. Moreover there is a unique geodesic passing through a point with a given tangent vector.

A connection enables to define the **parallel transport** of a vector (or a basis) along a curve (or a vector field).

Let C be a curve defined by a path $p:\mathbb{R}\rightarrow M:p(\tau)$ with $p(0)=a$, and a vector $v\in T_aM$. The parallel transported vector is given by a map :

$V:\mathbb{R}\rightarrow T_{p(\tau)}M:V(\tau)$ such that : $\nabla^{M}_{\frac{dp}{d\tau}}V\left(\tau\right)=0,V\left(0\right)=v$

thus we have the differential equation with $V\left(\tau\right)=\sum_{i=0}^{3}V^{i}\left(\tau\right)\varepsilon_{i}\left(p\left(\tau\right)\right)$

$\nabla^{M}_{\frac{dp}{d\tau}}V\left(\tau\right)=\sum_{\alpha i=0}^{3}\left(\partial_{\alpha}V^{i}+\sum_{j=0}^{3}\Gamma_{M}\left(p\left(\tau\right)\right)^{i}_{\alpha j}V^{j}\right)\left(\frac{dp}{d\tau}\right)^{\alpha}\varepsilon_{i}\left(p(t)\right)=0$

$\frac{dV^{i}}{d\tau}+\sum_{\alpha j=0}^{3}\Gamma_{M}\left(p\left(\tau\right)\right)^{i}_{\alpha j}V^{j}\left(\frac{dp}{d\tau}\right)^{\alpha}=0$

A **geodesic** is a path such that its tangent is parallel transported by the connection :
$p:\mathbb{R}\rightarrow M:p(\tau)$ with $p(0)=a$
$V\left(\tau\right)=\frac{dp}{d\tau}=\sum_{i=0}^{3}V^{i}\left(\tau\right)\varepsilon_{i}\left(p\left(\tau\right)\right)=\sum_{k\alpha=0}^{3}V^{k}\left(\tau\right)P'^{\alpha}_{k}\left(p\left(\tau\right)\right)\partial\xi_{\alpha}$
$\frac{dV^{i}}{d\tau}+\sum_{\alpha jk=0}^{3}\Gamma_{M}\left(p\left(\tau\right)\right)^{i}_{\alpha j}V^{j}\left(\tau\right)V^{k}\left(\tau\right)P'^{\alpha}_{k}\left(p\left(\tau\right)\right)=0$
or in matrix form :
$\left[\frac{dV}{d\tau}\right]+\sum_{\alpha}\left(\left[\Gamma_{M\alpha}\right]\left[V\right]\right)\left(\left[P'\right]\left[V\right]\right)^{\alpha}=0$
The scalar product $\langle V,V\rangle$ is constant :
$\frac{d}{d\tau}\langle V,V\rangle=\frac{d}{d\tau}\left(\left[V\right]^{t}\left[\eta\right]\left[V\right]\right)$
$=-\sum_{\alpha}\left(\left[P'\right]\left[V\right]\right)^{\alpha}\left[V\right]^{t}\left[\Gamma_{M\alpha}\right]^{t}\left[\eta\right]\left[V\right]-\sum_{\alpha}\left(\left[P'\right]\left[V\right]\right)^{\alpha}\left[V\right]^{t}\left[\eta\right]\left(\left[\Gamma_{M\alpha}\right]\left[V\right]\right)$
$=-\sum_{\alpha}\left(\left[P'\right]\left[V\right]\right)^{\alpha}\left[V\right]^{t}\left(\left[\Gamma_{M\alpha}\right]^{t}\left[\eta\right]+\left[\eta\right]\left[\Gamma_{M\alpha}\right]\right)\left[V\right]=0$

A **field of geodesics** is a vector field U such that it is parallel transported along its integral curves $p\left(\tau\right)=\Phi_{U}\left(\tau,x\right)$.

As $\langle U,U\rangle$ is constant, for a time like geodesic field we can take $\langle U,U\rangle=-1$ and, for a given observer, associate a section $\sigma_w\in P_W$:
$U=\left(\left(2a_w^2-1\right)\varepsilon_0\left(m\right)+a_w\sum_{j=1}^{3}w_j\varepsilon_j\left(m\right)\right)$
and $\mathbf{U}\left(m\right)=\mathbf{Ad}_{\sigma_w}\varepsilon_0\left(m\right)$

The formalism of vector bundles enables us to give a useful description of these geodesics, through the value of σ_w.

Theorem 100 *For a given observer, fields of geodesics are represented by sections* $\sigma_w\in\mathfrak{X}\left(P_G\right)$ *such that* $\nabla^{G}_{U}\sigma_w\in T_1Spin\left(3\right)$.

They are solutions of the differential equation :

$$\frac{dw}{dt}=\left[j\left(w\right)\right]\widehat{G}_r+\left(-a_w+\frac{1}{4a_w}j\left(w\right)j\left(w\right)\right)\widehat{G}_w \tag{5.22}$$

where $\upsilon\left(\widehat{G}_r,\widehat{G}_w\right)$ *is the value of the potential of the gravitational field along the geodesic*

Proof. i) In the standard basis and with the Clifford algebra formalism :

$\nabla_V^M U = \frac{dU}{d\tau} + \sum_{\alpha=0}^{3} \left(V \left(\upsilon \left(G_{r\alpha}, G_{w\alpha} \right) \cdot U - U \cdot \upsilon \left(G_{r\alpha}, G_{w\alpha} \right) \right) \right)$

$= \frac{d}{d\tau} \mathbf{Ad}_{\sigma_w} \varepsilon_0 + \upsilon \left(\widehat{G}_r, \widehat{G}_w \right) \cdot \mathbf{Ad}_{\sigma_w} \varepsilon_0 - \mathbf{Ad}_{\sigma_w} \varepsilon_0 \cdot \upsilon \left(\widehat{G}_r, \widehat{G}_w \right)$

with $V^\alpha = \sum_i P_i^\alpha U^i, \widehat{G}_r = \sum_{\alpha=0}^{3} G_{r\alpha} V^\alpha; \widehat{G}_w = \sum_{\alpha=0}^{3} G_{w\alpha} V^\alpha, U(m) = \mathbf{Ad}_{\sigma_w} \varepsilon_0 (m)$

$\frac{d}{d\tau} \mathbf{Ad}_{\sigma_w} \varepsilon_0 = \frac{d\sigma_w}{d\tau} \cdot \varepsilon_0 \cdot \sigma_w^{-1} - \sigma_w \cdot \varepsilon_0 \cdot \sigma_w^{-1} \cdot \frac{d\sigma_w}{d\tau} \cdot \sigma_w^{-1}$

$= \left(\sigma_w \cdot \nabla_U^G \sigma_w - \upsilon \left(\widehat{G}_r, \widehat{G}_w \right) \cdot \sigma_w \right) \cdot \varepsilon_0 \cdot \sigma_w^{-1} - \sigma_w \cdot \varepsilon_0 \cdot \sigma_w^{-1} \cdot \left(\sigma_w \cdot \nabla_U^G \sigma_w - \upsilon \left(\widehat{G}_r, \widehat{G}_w \right) \cdot \sigma_w \right) \cdot \sigma_w^{-1}$

$= \sigma_w \cdot \nabla_U^G \sigma_w \cdot \varepsilon_0 \cdot \sigma_w^{-1} - \upsilon \left(\widehat{G}_r, \widehat{G}_w \right) \cdot \sigma_w \cdot \varepsilon_0 \cdot \sigma_w^{-1} - \sigma_w \cdot \varepsilon_0 \cdot \nabla_U^G \sigma_w \cdot \sigma_w^{-1} + \sigma_w \cdot \varepsilon_0 \cdot \sigma_w^{-1} \cdot \upsilon \left(\widehat{G}_r, \widehat{G}_w \right)$

with $\frac{d\sigma_w}{d\tau} = \sigma_w \cdot \nabla_U^G \sigma_w - \upsilon \left(\widehat{G}_r, \widehat{G}_w \right) \cdot \sigma_w$

$\nabla_V^M U = \sigma_w \cdot \nabla_U^G \sigma_w \cdot \varepsilon_0 \cdot \sigma_w^{-1} - \sigma_w \cdot \varepsilon_0 \cdot \nabla_U^G \sigma_w \cdot \sigma_w^{-1}$

$-\upsilon \left(\widehat{G}_r, \widehat{G}_w \right) \cdot \sigma_w \cdot \varepsilon_0 \cdot \sigma_w^{-1} + \sigma_w \cdot \varepsilon_0 \cdot \sigma_w^{-1} \cdot \upsilon \left(\widehat{G}_r, \widehat{G}_w \right)$

$+\upsilon \left(\widehat{G}_r, \widehat{G}_w \right) \cdot \sigma_w \cdot \varepsilon_0 \cdot \sigma_w^{-1} - \sigma_w \cdot \varepsilon_0 \cdot \sigma_w^{-1} \cdot \upsilon \left(\widehat{G}_r, \widehat{G}_w \right)$

$= \sigma_w \cdot \left(\nabla_U^G \sigma_w \cdot \varepsilon_0 - \varepsilon_0 \cdot \nabla_U^G \sigma_w \right) \cdot \sigma_w^{-1}$

So, with the covariant derivative on the principal bundle P_G we have a geodesic iff :

$\nabla_U^G \sigma_w \cdot \varepsilon_0 - \varepsilon_0 \cdot \nabla_U^G \sigma_w = 0$

that is iff $\nabla_u^G \sigma_w$ commutes with ε_0.

For any element $\upsilon(r, w)$ of $T_1 Spin(3,1)$ we have the identity :

$\upsilon(r, w) \cdot \varepsilon_0 - \varepsilon_0 \cdot \upsilon(r, w) = w$ (see Annex for the proof)

So : $\upsilon(r, w) \in T_1 Spin(3) \Leftrightarrow \upsilon(r, w) \cdot \varepsilon_0 - \varepsilon_0 \cdot \upsilon(r, w) = 0 \Leftrightarrow w = 0$

And the geodesics are represented by sections such that $\nabla_U^G \sigma_w \in T_1 Spin(3)$.

ii) The sections

$\nabla_\alpha^G \sigma_w = \sigma_w^{-1} \cdot \left(\partial_\alpha \sigma_w + \upsilon \left(G_{r\alpha}, G_{w\alpha} \right) \cdot \sigma_w \right)$ read :

$\sigma_w^{-1} \cdot \partial_\alpha \sigma_w = \upsilon\left(\frac{1}{2} j(w) \partial_\alpha w, \frac{1}{4a_w} \left(-j(w) j(w) + 4 \right) \partial_\alpha w\right)$

$\sigma_w^{-1} \cdot \upsilon \left(G_{r\alpha}, G_{w\alpha} \right) \cdot \sigma_w = \mathbf{Ad}_{\sigma_w^{-1}} \upsilon \left(G_{r\alpha}, G_{w\alpha} \right) =$

$= \upsilon \left(\left[1 - \frac{1}{2} j(w) j(w) \right] G_{r\alpha} + \left[a_w j(w) \right] G_{w\alpha}, - \left[a_w j(w) \right] G_{r\alpha} + \left[1 - \frac{1}{2} j(w) j(w) \right] G_{w\alpha} \right)$

$\nabla_\alpha^G \sigma_w = \upsilon\left(\frac{1}{2} j(w) \partial_\alpha w + \left[1 - \frac{1}{2} j(w) j(w) \right] G_{r\alpha} + \left[a_w j(w) \right] G_{w\alpha},\right.$

$\left.\frac{1}{4a_w} \left(-j(w) j(w) + 4 \right) \partial_\alpha w - \left[a_w j(w) \right] G_{r\alpha} + \left[1 - \frac{1}{2} j(w) j(w) \right] G_{w\alpha} \right)$

So geodesic fields are associated to the sections such that :

$\sum_\alpha V^\alpha \frac{1}{4a_w} \left(-j(w) j(w) + 4 \right) \partial_\alpha w - \left[a_w j(w) \right] G_{r\alpha} + \left[1 - \frac{1}{2} j(w) j(w) \right] G_{w\alpha} = 0$

$\left(-j(w) j(w) + 4 \right) \frac{dw}{dt} - 4a_w^2 \left[j(w) \right] \widehat{G}_r + 4a_w \left[1 - \frac{1}{2} j(w) j(w) \right] \widehat{G}_w = 0$

By left multiplication with w^t :

$w^t \frac{dw}{dt} + a_w w^t \widehat{G}_w = 0$

$w^t \frac{dw}{dt} = 4a_w \frac{da_w}{dt} = -a_w w^t \widehat{G}_w$

$\frac{da_w}{dt} = -w^t \widehat{G}_w$

Moreover the equation reads :

$\left(-ww^t + w^t w + 4 \right) \frac{dw}{dt} - 4a_w^2 \left[j(w) \right] \widehat{G}_r + 4a_w \left[1 - \frac{1}{2} \left(ww^t - w^t w \right) \right] \widehat{G}_w = 0$

$-w \left(w^t \frac{dw}{dt} \right) + \frac{dw}{dt} w^t w + 4 \frac{dw}{dt} - 4a_w^2 \left[j(w) \right] \widehat{G}_r + 4a_w \left(1 + \frac{1}{2} w^t w \right) \widehat{G}_w - 2a_w ww^t \widehat{G}_w = 0$

$-\left(2a_w w^t \widehat{G}_w + w^t \frac{dw}{dt} \right) w + \left(4 \left(a_w^2 - 1 \right) + 4 \right) \frac{dw}{dt} - 4a_w^2 j(w) \widehat{G}_r + 4a_w \left(1 + 2 \left(a_w^2 - 1 \right) \right) \widehat{G}_w = 0$

$-\left(a_w w^t \widehat{G}_w \right) w + 4a_w^2 \frac{dw}{dt} - 4a_w^2 \left[j(w) \right] \widehat{G}_r + 4a_w \left(2a_w^2 - 1 \right) \widehat{G}_w = 0$

$\left(-\frac{1}{4} w^t \widehat{G}_w \right) w + a_w \frac{dw}{dt} - a_w \left[j(w) \right] \widehat{G}_r + \left(2a_w^2 - 1 \right) \widehat{G}_w = 0$

$a_w \frac{dw}{dt} = a_w \left[j(w) \right] \widehat{G}_r + \left(1 - 2a_w^2 + \frac{1}{4} ww^t \right) \widehat{G}_w$

$a_w \frac{dw}{dt} = a_w \left[j(w) \right] \widehat{G}_r + \left(1 - 2a_w^2 + \frac{1}{4} \left(j(w) j(w) + 4 \left(a_w^2 - 1 \right) \right) \right) \widehat{G}_w$

$\frac{dw}{dt} = [j(w)] \widehat{G}_r + \left(-a_w + \frac{1}{4a_w} j(w) j(w)\right) \widehat{G}_w$ ■

We have a first order ODE with parameters G. Fields of geodesics are not unique, but we can assume that usually there would be a unique geodesic passing through a point and with an initial tangent vector.

There is nothing equivalent for the null curves, such that their tangent vector u has a null scalar product : $\langle u, u \rangle = 0$. But the definition of the flow of a vector field, which does not involve the metric, still holds.

Remark : there are other definitions of geodesic curves, in particular as curve with an extremal length, using a metric. A classic demonstration proves that a curve of extremal length is necessarily a curve along which the tangent is transported, but this proof uses explicitly the Levi-Civita connection and some of its specific properties and does not hold any longer for a general affine connection. Using the principle of least action, a free particle moves along a path of extremal length, but which is not necessarily a geodesic as understood here.

5.3.5 The Levi-Civita connection

In Differential Geometry one defines affine connections (Maths.1537), which are bilinear operators acting on vector fields (sections of the tangent bundle) $\nabla \in \mathcal{L}^2 (\mathfrak{X}(TM), \mathfrak{X}(TM); \mathfrak{X}(TM))$. They read in holonomic basis of a chart :

$\nabla_\alpha V = \sum_\beta \left(\partial_\beta V^\alpha + \sum_\gamma \Gamma^\alpha_{\beta\gamma} V^\gamma\right) \partial\xi^\beta \otimes d\xi_\alpha$

with Christoffel symbols $\Gamma^\alpha_{\beta\gamma}(m)$ which change in a change of chart in a complicated way. So an affine connection is a covariant derivative, defined in the tangent bundle, and acting on sections of the tangent bundle, which are vector fields, or tensors. There can be many different affine connections.

An affine connection is said to be symmetric if $\Gamma^\alpha_{\beta\gamma} = \Gamma^\alpha_{\gamma\beta}$

When there is a metric (Riemannian or not) defined by a tensor g on a manifold, an affine connection is said to be metric if $\nabla_\alpha g = 0$: it preserves the scalar product of two vectors. Then one can define a unique, metric, symmetric connection, called the Levi-Civita connection. It reads (Maths.1626) :

$\Gamma^\alpha_{\beta\gamma} = \frac{1}{2} \sum_\eta g^{\alpha\eta} (\partial_\beta g_{\gamma\eta} + \partial_\gamma g_{\beta\eta} - \partial_\eta g_{\beta\gamma})$

And this has been the bread and butter of workers on GR for decenniums, in a formalism where the metric is at the core of the model.

With a principal bundle, and a principal connection, one can define covariant derivatives in any associated vector bundle, including of course the tangent bundle to M. And it has all the properties of the usual covariant derivative of affine connections. Connections on fiber bundles are a more general tool than usual affine connections which are strictly limited to the tangent bundle. We have seen that the connection $\mathbf{G}$ on P_G induces a linear connection on $P_G[\mathbb{R}^4, \mathbf{Ad}]$, which is nothing more than TM with an orthonormal basis, and a covariant derivative ∇^M with Christoffel symbol Γ_M. By translating the orthonormal basis $(\varepsilon_i)_{i=0}^3$ into the holonomic basis $(\partial\xi_\alpha)_{\alpha=0}^3$ of any chart using the tetrad, a straightforward computation (Maths.2005) gives the Christoffel coefficients $\widehat{\Gamma}^\gamma_{\alpha\beta}$ of the affine connection Γ_M, expressed in the basis of the chart :

$\widehat{\Gamma}^\gamma_{\alpha\beta} = P^\gamma_i \left(\partial_\alpha P'^i_\beta + \Gamma^i_{M\alpha j} P'^j_\beta\right)$

In matrix form :

$\widehat{\Gamma}^\gamma_{\alpha\beta} = \left[\widehat{\Gamma}_\alpha\right]^\gamma_\beta, \Gamma^i_{M\alpha j} = [\Gamma_{M\alpha}]^i_j,$

$[\Gamma_{M\alpha}] = \sum_{a=1}^6 G_{a\alpha} [\kappa_a]$

$\left[\widehat{\Gamma}_\alpha\right] = [P]\left([\partial_\alpha P'] + [\Gamma_{M\alpha}][P']\right) \Leftrightarrow [\Gamma_{M\alpha}] = \left([P']\left[\widehat{\Gamma}_\alpha\right] - [\partial_\alpha P']\right)[P]$

Any affine connection deduced this way from a principal connection is necessarily metric, but it is not necessarily symmetric.

To sum up :

- affine connections are defined in the strict framework of the tangent bundle, and the Levi-Civita connection is one of these connections, with specific properties (it is metric and symmetric); the covariant derivative which is deduced acts only on vectors fields (or tensors) of the tangent bundle.

- connections on principal bundle define connections on any associated vector bundle and act on sections of these bundles. So one can compute a covariant derivative acting on vectors fields of the tangent bundle, which is necessarily metric but not necessarily symmetric.

So, using the formalism of fiber bundles we do not miss anything, we can get the usual results, but in a more elegant and simple way. One can require from the principal connection $\mathbf{G}$ on P_G that the induced connection on TM is symmetric, which will then be identical to the Levi-Civita connection. This requests :

$\forall \alpha, \beta, \gamma$:

$\left[\widehat{\Gamma}_\alpha\right]^\gamma_\beta = \left([P]\left([\partial_\alpha P'] + [\Gamma_{M\alpha}][P']\right)\right)^\gamma_\beta = \left[\widehat{\Gamma}_\beta\right]^\gamma_\alpha = \left([P]\left([\partial_\beta P'] + [\Gamma_{M\beta}][P']\right)\right)^\gamma_\alpha$

which has no obvious meaning for Γ_M.

Actually the Levi-Civita connection is traditionally used because it is the natural mathematical choice when one starts from the metric. Moreover it is assumed that the gravitational field (whose action goes through the connection) acts symmetrically, in the meaning that it has no torsion (or no torque). But actually this assumption has not been verified (which is difficult), and different theories have been proposed, notably by Einstein, Cartan and Eisenhart (the so called "fernparallelism"), which consider connections with torsion, that is connections other than the Levi-Civita connection. However, when starting from the metric, they lead mostly to more complicated computations, in what is already a dreadful endeavour. In the fiber bundle framework there is no such problem and actually it would be the requirement of symmetry, always possible at any point, which would introduce a complication. Moreover the introduction of spinors and the distinction of the components G_r, G_w of the connection, are a more efficient way to deal with rotation and torque so it is justified that we keep the more general connection. An additional argument is that the Levi-Civita connection does not make any distinction between the bases, which can be induced by any chart. But, as we have seen, there is always a privileged chart, that of the observer, and the use of an orthogonal basis, in the fiber bundle formalism, is a useful reminder of this feature.

5.3.6 The total connection

Definition

The covariant derivative is computed from the action of the Lie algebra of the group $Spin(3,1) \times U$, that is by derivation of the action $\vartheta(\sigma, \varkappa)\psi$.And by combining the previous result we have :

Proposition 101 *There are on Q a principal connection defined by the potentials*

$G \in \Lambda_1(M; T_1 Spin(3,1)) : TM \to T_1 Spin(3,1) ::$

$G(m) = \sum_{\alpha=0}^{3} \upsilon(G_{r\alpha}(m), G_{w\alpha}(m))\, d\xi^\alpha$

$\grave{A} \in \Lambda_1(M; T_1 U) : TM \to T_1 U :: \grave{A}(m) = \sum_{\alpha=0}^{3}\sum_{a=1}^{m} \grave{A}^a_\alpha(m)\, \theta_a \otimes dm^\alpha$

The action of the fields on the state of a particle is given by the covariant derivative.

It reads for $\psi \in \mathfrak{X}(Q[E\otimes F,\vartheta])$:
$\nabla_\alpha \psi = \sum_{\alpha=0}^{3}\left(\partial_\alpha \psi^{ij} + \sum_{a=1}^{6}[\gamma C(G_\alpha^a)][\psi] + \sum_{a=1}^{m}\grave{A}_\alpha^a[\psi][\theta_a]\right)^{ij}\mathbf{e}_i(m)\otimes\mathbf{f}_i(m)\otimes d\xi^\alpha$
In matrix form :

$$\nabla_\alpha \psi \in \Lambda_1(M, Q[E\otimes F,\vartheta]) : [\nabla_\alpha \psi] = \sum_{\alpha=0}^{3}[\partial_\alpha \psi] + [\gamma C(G_\alpha)][\psi] + [\psi]\left[\grave{A}_a\right] \quad (5.23)$$

with $[\gamma C(G_\alpha)] = [\gamma C(\upsilon(G_{r\alpha}, G_{w\alpha}))]$
$[\nabla_\alpha \psi] = [\partial_\alpha \psi] + \sum_{a=1}^{m}\grave{A}_\alpha^a[\psi][\theta_a] - \frac{i}{2}\left(\sum_{a=1}^{3}G_{w\alpha}^a[\gamma_a][\gamma_0] + G_{r\alpha}[\widetilde{\gamma}_a]\right)[\psi]$

We have another formula :

$$\nabla_\alpha \psi = \vartheta(\sigma,\varkappa)\left(\left[\gamma C\left(\sigma^{-1}\partial_\alpha \sigma + \mathbf{Ad}_{\sigma^{-1}}G_\alpha\right)\right][\psi_0] + [\psi_0]\left[Ad_\varkappa\left(\grave{A}_\alpha\right)\right]\right) \quad (5.24)$$

Proof. $[\partial_\alpha \psi] = \partial_\alpha([\gamma C(\sigma)][\psi_0][\varrho(\varkappa)])$
$= [\gamma C(\partial_\alpha \sigma)][\psi_0][\varrho(\varkappa)] + [\gamma C(\sigma)][\psi_0][\varrho'(\varkappa)\partial_\alpha \varkappa]$
$[\varrho'(\varkappa)\partial_\alpha \varkappa] = \left[\varrho(\varkappa)\varrho'(1)\left(L'_{\varkappa^{-1}}\varkappa\right)\partial_\alpha \varkappa\right] = [\varrho(\varkappa)]\sum_{a=1}^{m}\partial_\alpha \varkappa^a[\theta_a] = [\varrho(\varkappa)][\partial_\alpha \varkappa]$ (Maths.1900)
$\nabla_\alpha \psi = ([\gamma C(\partial_\alpha \sigma)][\psi_0][\varrho(\varkappa)]) + [\gamma C(\sigma)][\psi_0][\varrho(\varkappa)][\partial_\alpha \varkappa]$
$+ [\gamma C(G_\alpha)][\gamma C(\sigma)][\psi_0][\varrho(\varkappa)] + [\gamma C(\sigma)][\psi_0][\varrho(\varkappa)]\left[\grave{A}_\alpha\right]$
$= [\gamma C(\sigma)]\{\left[\gamma C\left(\sigma^{-1}\cdot\partial_\alpha \sigma\right)\right][\psi_0] + \left[\gamma C\left(\sigma^{-1}\cdot G_\alpha\cdot\sigma\right)\right][\psi_0]$
$+ [\psi_0][\varrho(\varkappa)][\partial_\alpha \varkappa][\varrho(\varkappa)]^{-1} + [\psi_0][\varrho(\varkappa)]\left[\grave{A}_\alpha\right][\varrho(\varkappa)]^{-1}\}[\varrho(\varkappa)]$
$\nabla_\alpha \psi = [\gamma C(\sigma)]\{\left[\gamma C\left(\sigma^{-1}\cdot\partial_\alpha \sigma + \mathbf{Ad}_{\sigma^{-1}}G_\alpha\right)\right][\psi_0] + [\psi_0]\left(Ad_\varkappa[\partial_\alpha \varkappa] + \left[\grave{A}_\alpha\right]\right)\}[\varrho(\varkappa)]$
$\sigma^{-1}\partial_\alpha \sigma + \mathbf{Ad}_{\sigma^{-1}}G_\alpha \in T_1 Spin(3,1), Ad_\varkappa\left(\partial_\alpha \varkappa + \grave{A}_\alpha\right) \in T_1 U$ ■

Physical meaning

Action of a force field

The Newton's law gives the relation between the force F exercised by a field and the variation of the momentum P of a particle. It reads : $F = \frac{dP}{dt} = u\nabla P$ with the operator $\nabla = \partial_\alpha$

The operator ∇ is replaced by the covariant derivative :
$\nabla_u \psi = \frac{d\psi}{dt} + \left[\gamma C\left(\widehat{G}\right)\right][\psi] + [\psi]\left[\widehat{\grave{A}}\right]$

$\left[\gamma C\left(\widehat{G}\right)\right][\psi], [\psi]\left[\widehat{\grave{A}}\right]$ are the action of the force fields, expressed as usual with the potential. They involve the charge and the mass, but also r, w. The action of the force depends on the geometric disposition of the particle with respect to the field.

Systemic energy of a particle

We have an important result :

Theorem 102 *The scalar product $\langle\psi, \nabla_\alpha \psi\rangle$ is purely imaginary : $\langle\psi, \nabla_\alpha \psi\rangle = i\operatorname{Im}\langle\psi, \nabla_\alpha \psi\rangle$*

Proof. $[\nabla_\alpha \psi] = [\partial_\alpha \psi] + [\psi]\left[\grave{A}_\alpha\right] - \frac{i}{2}\sum_{a=1}^{3}G_{w\alpha}^a\gamma_a\gamma_0[\psi] + G_{r\alpha}\widetilde{\gamma}_a[\psi]$
$\langle\psi, \nabla_\alpha \psi\rangle = Tr[\psi]^*\gamma_0[\partial_\alpha \psi] + Tr[\psi]^*\gamma_0[\psi]\left[\grave{A}_a\right]$
$-\frac{i}{2}\sum_{a=1}^{3}G_{w\alpha}^a Tr[\psi]^*\gamma_0\gamma_a\gamma_0[\psi] + G_{r\alpha}Tr[\psi]^*\gamma_0\widetilde{\gamma}_a[\psi]$
$= Tr[\psi]^*\gamma_0[\partial_\alpha \psi] + Tr[\psi]^*\gamma_0[\psi]\left[\grave{A}_a\right] - \frac{i}{2}\sum_{a=1}^{3}\left(-G_{w\alpha}^a Tr[\psi]^*\gamma_a[\psi] + G_{r\alpha}Tr[\psi]^*\gamma_0\widetilde{\gamma}_a[\psi]\right)$

$Tr\left[\psi\right]^{*}\gamma_{a}\left[\psi\right], Tr\left[\psi\right]^{*}\gamma_{0}\widetilde{\gamma}_{a}\left[\psi\right]$ are real, $Tr\left[\psi\right]^{*}\gamma_{0}\left[\psi\right]\left[\theta_{a}\right], Tr\left[\psi\right]^{*}\gamma_{0}\left[\partial_{\alpha}\psi\right]$ are imaginary :
$\overline{\left(Tr\left[\psi\right]^{*}\gamma_{a}\left[\psi\right]\right)} = Tr\left(\left[\psi\right]^{*}\gamma_{a}\left[\psi\right]\right)^{*} = Tr\left[\psi\right]^{*}\gamma_{a}\left[\psi\right]$
$\overline{Tr\left[\psi\right]^{*}\gamma_{0}\widetilde{\gamma}_{a}\left[\psi\right]} = Tr\left(\left[\psi\right]^{*}\gamma_{0}\widetilde{\gamma}_{a}\left[\psi\right]\right)^{*} = Tr\left[\psi\right]^{*}\left(\gamma_{0}\widetilde{\gamma}_{a}\right)^{*}\left[\psi\right] = Tr\left[\psi\right]^{*}\gamma_{0}\widetilde{\gamma}_{a}\left[\psi\right]$
$\overline{Tr\left[\psi\right]^{*}\gamma_{0}\left[\psi\right]\left[\theta_{a}\right]} = Tr\left(\left[\psi\right]^{*}\gamma_{0}\left[\psi\right]\left[\theta_{a}\right]\right)^{*} = Tr\left[\theta_{a}\right]^{*}\left[\psi\right]^{*}\left[\gamma_{0}\right]\left[\psi\right]$
$= -Tr\left[\theta_{a}\right]\left[\psi\right]^{*}\left[\gamma_{0}\right]\left[\psi\right] = -Tr\left[\psi\right]^{*}\left[\gamma_{0}\right]\left[\psi\right]\left[\theta_{a}\right]$
$\left\langle\psi,\psi\right\rangle = \left\langle\psi_{0},\psi_{0}\right\rangle \Rightarrow \left\langle\psi,\partial_{\alpha}\psi\right\rangle + \left\langle\partial_{\alpha}\psi,\psi\right\rangle = 0$ ■

$$\operatorname{Im}\left\langle\psi,\nabla_{\alpha}\psi\right\rangle = \frac{1}{i}\left(\left\langle\psi,\partial_{\alpha}\psi\right\rangle + \left\langle\psi,\left[\psi\right]\left[\grave{A}_{\alpha}\right]\right\rangle + \left\langle\psi,\left[\gamma C\left(G_{\alpha}\right)\right]\psi\right\rangle\right) \quad (5.25)$$

$\operatorname{Im}\left\langle\psi,\nabla_{V}\psi\right\rangle = \frac{1}{i}\left(\left\langle\psi,\frac{d\psi}{dt}\right\rangle + \left\langle\psi,\left[\psi\right]\left[\widehat{A}\right]\right\rangle + \left\langle\psi,\gamma C\left(\widehat{G}\right)\psi\right\rangle\right)$

This quantity is defined through a derivative and the value of the potential for a vector V : it clearly measures a variation : the change of the energy of the particle along its trajectory, given by V, in a system which includes forces fields. The energy depends on the choice of an observer : V is defined by σ_{w} with respect to ε_{0} and the time t.

$\frac{1}{i}\left\langle\psi,\frac{d\psi}{dt}\right\rangle$ is the change in the kinetic energy of the particle with respect to the observer. This is the extension of the expression $\frac{1}{i}\left\langle S,\frac{dS}{d\tau}\right\rangle$ for the spinor.

$\frac{1}{i}\left\langle\psi,\left[\psi\right]\left[\widehat{A}\right]\right\rangle + \left\langle\psi,\gamma C\left(\widehat{G}\right)\psi\right\rangle$ represents the change of the energy of the particle in its interaction with the fields, as measured by an observer.

$\frac{1}{i}\left\langle S,\frac{dS}{dt}\right\rangle$ can be expressed with one vector k. The result can be extended.

By combination with the previous result :
$\operatorname{Im}\left\langle\psi,\nabla_{\alpha}\psi\right\rangle$
$= \frac{1}{i}\left\langle\vartheta\left(\sigma,\varkappa\right)\psi_{0},\vartheta\left(\sigma,\varkappa\right)\left(\left[\gamma C\left(\sigma^{-1}\partial_{\alpha}\sigma + \mathbf{Ad}_{\sigma^{-1}}G_{\alpha}\right)\right]\left[\psi_{0}\right] + \left[\psi_{0}\right]\left[Ad_{\varkappa}\left(\grave{A}_{\alpha}\right)\right]\right)\right\rangle$
$= \frac{1}{i}\left\langle\psi_{0},\left[\gamma C\left(\sigma^{-1}\partial_{\alpha}\sigma + \mathbf{Ad}_{\sigma^{-1}}G_{\alpha}\right)\right]\left[\psi_{0}\right] + \left[\psi_{0}\right]\left[Ad_{\varkappa}\left(\grave{A}_{\alpha}\right)\right]\right\rangle$
$= \frac{1}{i}\left\langle\psi_{0},\left[\gamma C\left(\sigma^{-1}\partial_{\alpha}\sigma + \mathbf{Ad}_{\sigma^{-1}}G_{\alpha}\right)\right]\left[\psi_{0}\right]\right\rangle + \frac{1}{i}\left\langle\psi_{0},\left[\psi_{0}\right]\left[Ad_{\varkappa}\left(\grave{A}_{\alpha}\right)\right]\right\rangle$

We have an explicit formula for the kinetic energy and the action of the gravitational field.

Theorem 103 *There are 3 scalars $\left(k_{a}\right)_{a=1}^{3}$ defined by the fundamental state ψ_{0} such that :*

$$\left\langle\psi_{0},\gamma C\left(\sigma^{-1}\partial_{\alpha}\sigma + \mathbf{Ad}_{\sigma^{-1}}G_{\alpha}\right)\psi_{0}\right\rangle = i\sum_{a=1}^{3}X_{\alpha}^{a}k_{a}$$

where $X_{\alpha} = \left[C\left(r\right)\right]^{t}\left(\left[D\left(r\right)\right]\partial_{\alpha}r + \frac{1}{2}j\left(w\right)\partial_{\alpha}w + \left[A\left(w\right)\right]G_{r\alpha} + \left[B\left(w\right)\right]G_{w\alpha}\right)$
with :
$\left[A\left(w\right)\right] = \left[1 - \frac{1}{2}j\left(w\right)j\left(w\right)\right]$
$\left[B\left(w\right)\right] = a_{w}\left[j\left(w\right)\right]$
$\left[C\left(r\right)\right] = \left[1 + a_{r}j\left(r\right) + \frac{1}{2}j\left(r\right)j\left(r\right)\right]$
$\left[D\left(r\right)\right] = \left[\frac{1}{a_{r}} + \frac{1}{2}j\left(r\right) + \frac{1}{4a_{r}}j\left(r\right)j\left(r\right)\right]$
$\left[C\left(r\right)\right]^{t}\left[D\left(r\right)\right] = \left[D\left(r\right)\right]^{t}$

Proof. Without the other fields :
$\operatorname{Im}\left\langle\psi,\nabla_{\alpha}\psi\right\rangle = \frac{1}{i}\left\langle\psi_{0},\left[\gamma C\left(\sigma^{-1}\partial_{\alpha}\sigma + \mathbf{Ad}_{\sigma^{-1}}G_{\alpha}\right)\right]\left[\psi_{0}\right]\right\rangle$
We have already proven that $\frac{1}{i}\left\langle S,\frac{dS}{d\tau}\right\rangle = k^{t}X$. The demonstration is very similar.
i) $\left(\sigma^{-1}\partial_{\alpha}\sigma + \mathbf{Ad}_{\sigma^{-1}}G_{\alpha}\right) = \upsilon\left(X_{\alpha},Y_{\alpha}\right) \in T_{1}Spin\left(3,1\right)$

$\gamma C(\upsilon(X_\alpha, Y_\alpha)) = \frac{1}{2}\begin{bmatrix} \sum_{a=1}^{3}(Y_\alpha^a - iX_\alpha^a)\sigma_a & 0 \\ 0 & -\sum_{a=1}^{3}(Y_\alpha^a + iX_\alpha^a)\sigma_a \end{bmatrix}$

Let us denote $[\psi_0] = \begin{bmatrix} \psi_R \\ \epsilon i \psi_R \end{bmatrix}$ with ψ_R a $2 \times n$ matrix and $\epsilon = +1$ for particles, and $\epsilon = -1$ for antiparticles.

$\gamma C(\upsilon(X_\alpha, Y_\alpha))[\psi_0] = \frac{1}{2}\begin{bmatrix} \sum_{a=1}^{3}(Y_\alpha^a - iX_\alpha^a)\sigma_a\psi_R \\ -\epsilon i \sum_{a=1}^{3}(Y_\alpha^a + iX_\alpha^a)\sigma_a\psi_R \end{bmatrix}$

$\left\langle \psi_0, \gamma C\left(\sigma^{-1}\partial_\alpha\sigma + \mathbf{Ad}_{\sigma^{-1}}G_\alpha\right)\psi_0 \right\rangle$

$= \frac{1}{2}Tr\left(\psi_R^*\left(-\epsilon\sum_{a=1}^{3}(Y_\alpha^a + iX_\alpha^a)\sigma_a\psi_R\right) - \epsilon i\psi_R^*\left(i\sum_{a=1}^{3}(Y_\alpha^a - iX_\alpha^a)\sigma_a\psi_R\right)\right)$

$= \epsilon\frac{1}{2}\left(-\sum_{a=1}^{3}(Y_\alpha^a + iX_\alpha^a)Tr(\psi_R^*\sigma_a\psi_R) + \sum_{a=1}^{3}(Y_\alpha^a - iX_\alpha^a)Tr(\psi_R^*\sigma_a\psi_R)\right)$

$= \frac{1}{2}\epsilon\sum_{a=1}^{3}\left(-(Y_\alpha^a + iX_\alpha^a) + (Y_\alpha^a - iX_\alpha^a)\right)Tr(\psi_R^*\sigma_a\psi_R)$

$= -\epsilon i\sum_{a=1}^{3}X_\alpha^a Tr(\psi_R^*\sigma_a\psi_R)$

$= i\sum_{a=1}^{3}X_\alpha^a k^a = ik^t X_\alpha$

where $k^a = -\epsilon(Tr(\psi_R^*\sigma_a\psi_R))$ are fixed scalars.

For spinors $k = -\epsilon(S_R^*\sigma_a S_R)$ is the inertial vector.

ii) $\sigma^{-1}\partial_\alpha\sigma = \upsilon([C(r)]^t(\frac{1}{2}j(w)\partial_\alpha w + [D(r)]\partial_\alpha r, [C]^t\left(\frac{1}{4a_w}[4 - j(w)j(w)]\partial_\alpha w\right)$

$\mathbf{Ad}_{\sigma^{-1}}G_\alpha = \begin{bmatrix} C^t & 0 \\ 0 & C^t \end{bmatrix}\begin{bmatrix} A & B \\ -B & A \end{bmatrix}\begin{bmatrix} G_{r\alpha} \\ G_{w\alpha} \end{bmatrix} = \begin{bmatrix} C^tAG_{r\alpha} + C^tBG_{w\alpha} \\ -C^tBG_{r\alpha} + C^tAG_{w\alpha} \end{bmatrix}$

Thus :

$\nabla_\alpha\sigma = \left(\sigma^{-1}\partial_\alpha\sigma + \mathbf{Ad}_{\sigma^{-1}}G_\alpha\right) = \upsilon(X_\alpha, Y_\alpha)$

$X_\alpha = [C(r)]^t\left(([D(r)]\partial_\alpha r + \frac{1}{2}j(w)\partial_\alpha w + [A(w)]G_{r\alpha} + [B(w)]G_{w\alpha}\right)$ ■

Remark : the definition of the matrices $[A], [B], [C]$ is the same as in [**Ad**] which is used in other parts of the book so there should be no confusion.

We will denote k, X_α as column matrices 3×1 and :

$$\begin{bmatrix} k^t X_\alpha = \sum_{a=1}^{3}X_\alpha^a k_a = -\epsilon\sum_{a=1}^{3}X_\alpha^a(Tr(\psi_R^*\sigma_a\psi_R)) \\ \widehat{X} = \sum_{\alpha=0}^{3}V^\alpha X_\alpha = \sum_{\alpha=0}^{3}\sum_{a=1}^{3}V^\alpha X_\alpha^a \overrightarrow{\kappa}^a \\ \widehat{X} = [C(r)]^t\left([D(r)]\frac{dr}{dt} + \frac{1}{2}j(w)\frac{dw}{dt} + [A(w)]\widehat{G}_r + [B(w)]\widehat{G}_w\right) \end{bmatrix} \tag{5.26}$$

The variation of energy due to the gravitational field is :

$\frac{1}{i}\langle\psi, [\gamma C(G_\alpha)]\psi\rangle = k^t[C(r)]^t\left([A(w)]\widehat{G}_r + [B(w)]\widehat{G}_w\right)$

So *the action of the gravitational field has two components* :

$k^t[C(r)]^t[A(w)]\widehat{G}_r \simeq k^t\left[1 - a_r j(r) + \frac{1}{2}j(r)j(r)\right]\widehat{G}_r$

$k^t[C(r)]^t[B(w)]\widehat{G}_w \simeq k^t\left[1 - a_r j(r) + \frac{1}{2}j(r)j(r)\right]j\left(\frac{v}{c}\right)\widehat{G}_w$

The second component, depending of the speed v, is not related to a "tidal effect" : the particle has no internal structure. We can assume that usually r is small and the first component is the most important, and practically depends only on G_r as usually experimented. But we have also another component depending on G_w and the spatial speed v. It is usually very small, but can be significant in Astro-Physics (the speed of stars in galaxies can reach 200 km / s). The sign of k depends on ϵ : particles and anti-particles should have opposite behaviors.

The variation of energy due to the other fields is :

$\frac{1}{i}\left\langle\psi, [\psi]\left[\widehat{A}\right]\right\rangle = \frac{1}{i}\left\langle\psi_0, [\psi_0]\left[Ad_\varkappa\widehat{A}\right]\right\rangle$

For the electromagnetic field the electric charge is represented in the inertial spinor. Elementary particles correspond to negative charge, and antiparticles to positive charge.
$\frac{1}{i}\left\langle \psi, [\psi]\left[\widehat{\grave{A}}\right]\right\rangle = q\widehat{\grave{A}}$

If only the EM field and gravity are present, $\psi = S$ the total covariant derivative reads :
$\nabla_\alpha S = \partial_\alpha S + [\gamma C(G_\alpha)] S + iq\grave{A}_\alpha S$
$\langle S_0, S_0\rangle = \epsilon M_p^2 c^2$
$k = -\epsilon \frac{1}{2} M_p^2 c^4 k_0$
and the systemic energy is :
$\frac{1}{i}\langle \psi, \nabla_V \psi\rangle = k^t \widehat{X} + q\widehat{\grave{A}}\langle S_0, S_0\rangle = -\epsilon\frac{1}{2}M_p^2 c^2 k_0^t \widehat{X} + q\epsilon\widehat{\grave{A}} M_p^2 c^2$

$$\frac{1}{i}\langle S, \nabla_V S\rangle = \epsilon M_p^2 c^2 \left(\frac{1}{2} k_0^t \widehat{X} + q\grave{A}_\alpha\right) \tag{5.27}$$

5.3.7 The inertial observer

The states of the particles and the fields are linked, so to measure one we have to know the other : to measure a charge one uses a known field, and to measure a field one uses a known particle. This process requires actually two measures, involving the motion of the particle, it is done locally and is represented by the standard gauges : $\mathbf{p}_G(m) = \varphi_G(m, 1), \mathbf{p}_U(m) = \varphi_U(m, 1)$ and the related holonomic bases $\mathbf{e}_i(m) = (\mathbf{p}_G(m), e_i), \mathbf{f}_j(m) = (\mathbf{p}_U(m), f_j)$. The measures are done with respect to the standards (represented by 1), which are arbitrary. For this reason the standard gauges and their bases are not sections, just a specific choice done by the observer at each point, they are not given by any physical law. This is consistent with the principle of locality (the measures are done locally) and the free will of the observer (he is not submitted himself to the laws of the system). However, both for modelling purposes and to give a physical meaning to the concepts, we need to assume some rules about the behavior of these gauges. From this point of view the status of the gauge on the principal bundle P_G is special, because its link with the tetrad.

Tetrad and the gravitational field

There are several ways to define an "inertial observer", which lead to the same formulas. This is a useful exercise to review all them.

i) To measure vectorial quantities one wishes to keep the basis as fixed as possible. The observer cannot do anything about the holonomic basis - by definition it is fixed - but he can hope to keep $(\varepsilon_i)_{i=0}^3$ as stable as possible, and for this he can measure the components of ε_i.with respect to $\partial\xi_\alpha$. However we have seen that, for an orthonormal basis, it is not possible to keep them constant, because the metric varies from one point to another. But the connection provides a mean to compare the bases at two different points, by the transport of vectors. A basis whose each vector is transported along a path can be seen as the best approximation for a constant basis. We still deal with a common observer : he has his own trajectory, from which he deduces the folliation, the vector field $\mathbf{O}$, the standard chart, with the constraint $\partial\xi_0 = \mathbf{O}$. The only specificity of the **inertial observer** is that the spatial vectors $(\varepsilon_i)_{i=1}^3$ of his orthonormal basis (over which he has control) are transported by the connection $\mathbf{G}$ along his trajectory. All observers follow trajectories given by $p_o(t) = \varphi_o(t, x_o) = \Phi_{\varepsilon_0}(t, x_o)$ so the condition applies equally at any point along the path given by the integral curves of the vector field $\mathbf{O}$.

The computation is more illuminating with the affine connection $\widehat{\Gamma}$ applied to the vectors $\varepsilon_i = \sum_{\alpha=0}^{3} P_i'^{\alpha}$ expressed in the holonomic basis. The vectors are transported by the connection **G** along his trajectory iff :

i=1,..3 :

$\widehat{\nabla}_{\varepsilon_0}\varepsilon_i = \sum_{\alpha,\beta=0}^{3} P_0^{\alpha}\left(\partial_\alpha P_i^{\beta} + \sum_{\gamma=0}^{3}\widehat{\Gamma}_{\beta\gamma}^{\beta} P_i^{\gamma}\right)\partial\xi_\beta = 0$

$\sum_{\alpha,\beta=0}^{3}\left(\frac{dP_i^{\beta}}{dt} + \sum_{\gamma=0}^{3}\widehat{\Gamma}_{\alpha\gamma}^{\beta} P_i^{\gamma} P_0^{\alpha}\right)\partial\xi_\beta = 0$

with : $\left[\widehat{\Gamma}_\alpha\right]_\gamma^\beta = \left([P]\left([\partial_\alpha P'] + [\Gamma_{M\alpha}][P']\right)\right)_\gamma^\beta = \sum_{j=0}^{3} P_j^{\beta}\left(\partial_\alpha P_\gamma'^{j} + \sum_{a=1}^{6}\sum_{k=0}^{3} G_\alpha^a [\kappa_a]_k^j P_\gamma'^{k}\right)$

$\frac{dP_i^{\beta}}{dt} + \sum_{\alpha,\gamma=0}^{3}\left(\sum_{j=0}^{3} P_j^{\beta}\left(\partial_\alpha P_\gamma'^{j} + \sum_{a=1}^{6}\sum_{k=0}^{3} G_\alpha^a [\kappa_a]_k^j P_\gamma'^{k}\right)\right) P_i^{\gamma} P_0^{\alpha} = 0$

$\frac{dP_i^{\beta}}{dt} + \sum_{\alpha\gamma=0}^{3}\sum_{j=0}^{3} P_j^{\beta} P_i^{\gamma}\frac{dP_\gamma'^{j}}{dt} + \sum_{a=1}^{6} G_\alpha^a [\kappa_a]_i^j P_j^{\beta} P_0^{\alpha} = 0$

$\left[\frac{dP}{dt}\right]_i^\beta + \left([P]\left[\frac{dP'}{dt}\right][P]\right)_i^\beta + \sum_{a=1}^{6}\sum_{\alpha=0}^{3} P_0^{\alpha} G_\alpha^a \left([P][\kappa_a]\right)_i^\beta = 0$

$\left[\frac{dP'}{dt}\right] = \frac{d}{dt}[P]^{-1} = -[P]^{-1}\left[\frac{dP}{dt}\right][P]^{-1}$

$\sum_{a=1}^{6}\sum_{\alpha=0}^{3} P_0^{\alpha} G_\alpha^a \left([P][\kappa_a]\right) = 0$

$[P]\left[\widehat{G}\right] = 0$

$$\widehat{G} = \sum_{a=1}^{6}\sum_{\alpha=0}^{3} P_0^{\alpha} G_\alpha^a \overrightarrow{\kappa}_a = 0 \tag{5.28}$$

The condition sums up to $\widehat{G} = 0$ along the trajectory (given by **O**) of the observer. This is the usual meaning of an inertial observer : it does not feel inertial or gravitational forces. So far one cannot shut down gravitation and the condition is met only if the trajectory itself is special.

ii) We have another way to see this. In any chart the gauge of an inertial observer is given by a section

$$p_I \in \mathfrak{X}(P_G) : p_I(m) = \varphi_G(m, \sigma_I(m)) : \nabla_U^G p_I = 0 \tag{5.29}$$

where U is the velocity of the observer. The formula for $\nabla_\alpha^G p_I = 0$ has been given previously (Total connection) for $p_I = \sigma_w \cdot \sigma_r$ and the usual notations :

$\nabla_\alpha^G \sigma = \left(\sigma^{-1}\partial_\alpha\sigma + \mathbf{Ad}_{\sigma^{-1}} G_\alpha\right) = \upsilon(X_\alpha, Y_\alpha)$

$X_\alpha = [C(r)]^t\left(\left(\left[\frac{1}{a_r} + \frac{1}{2}j(r) + \frac{1}{4a_r}j(r)j(r)\right]\partial_\alpha r + \frac{1}{2}j(w)\partial_\alpha w + [A(w)]G_{r\alpha} + [B(w)]G_{w\alpha}\right)\right.$

$Y_\alpha = [C(r)]^t\left(\frac{1}{4a_w}[4 - j(w)j(w)]\partial_\alpha w - [B(w)]G_{r\alpha} + [A(w)]G_{w\alpha}\right)$

Thus the condition reads :

$\sum_{\alpha=0}^{3} U^\alpha\left\{\left[\frac{1}{a_r} + \frac{1}{2}j(r) + \frac{1}{4a_r}j(r)j(r)\right]\partial_\alpha r + \frac{1}{2}j(w)\partial_\alpha w + [A(w)]G_{r\alpha} + [B(w)]G_{w\alpha}\right\} = 0$

$\sum_{\alpha=0}^{3} U^\alpha\left\{\frac{1}{4a_w}[4 - j(w)j(w)]\partial_\alpha w - [B(w)]G_{r\alpha} + [A(w)]G_{w\alpha}\right\} = 0$

$\left[\frac{1}{a_r} + \frac{1}{2}j(r) + \frac{1}{4a_r}j(r)j(r)\right]\frac{dr}{d\tau} + \frac{1}{2}j(w)\frac{dw}{d\tau} + [A(w)]\widehat{G}_r + [B(w)]\widehat{G}_w = 0$

$\frac{1}{4a_w}[4 - j(w)j(w)]\frac{dw}{d\tau} - [B(w)]\widehat{G}_r + [A(w)]\widehat{G}_w = 0$

$\frac{dw}{d\tau} = \left[a_w I + \frac{1}{4a_w}j(w)j(w)\right]\left([B(w)]\widehat{G}_r - [A(w)]\widehat{G}_w\right)$

$\frac{dw}{d\tau} = \left[a_w I + \frac{1}{4a_w}j(w)j(w)\right]\left([a_w j(w)]\widehat{G}_r - \left[1 - \frac{1}{2}j(w)j(w)\right]\widehat{G}_w\right)$

$= j(w)\widehat{G}_r + \left(a_w I - \frac{1}{4a_w}j(w)j(w)\right)\widehat{G}_w$

$\frac{dr}{dt}$

$= \left[\frac{1}{a_r} + \frac{1}{2} j(r) + \frac{1}{4a_r} j(r) j(r)\right]^{-1} \left(-\frac{1}{2} j(w) \frac{dw}{d\tau} - [A(w)] \widehat{G}_r - [B(w)] \widehat{G}_w\right)$

$= \left[\frac{1}{a_r} + \frac{1}{2} j(r) + \frac{1}{4a_r} j(r) j(r)\right]^{-1} \left(-\widehat{G}_r + \left(\frac{1}{2a_w} - 2a_w\right) j(w) \widehat{G}_w\right)$

The first condition is the same as for a geodesic, as can be expected : the trajectory must be a geodesic. But there is another condition pertaining to the spatial basis. So an inertial observer must follow very precise conditions to adjust his trajectory and his basis to the changes in the gravitational field. It will require a constant acceleration, contrary to the common understanding of the inertial observer[4]. In the absence of field one retrieves : $\sigma = Ct, g = Ct$ and the standard gauge. This is the usual meaning of inertial observers. This fine tuning makes unrealistic the idea of a genuine inertial observer.

iii) One can define an "inertial path" as a path such that a particle, submitted only to gravitation, does not feel any force : the inertial forces balance the gravitational force. This is in accordance with the principle of equivalence : inertial forces, due to the motion, are balanced by the gravitational field.

The sum of the gravitational and inertial forces on a spinor S along its world line u is given by $\nabla_u S$, so an "inertial path" is such that :

$\sigma : \mathbb{R} \to P_G :: \sigma(t) = \sigma_w(t) \cdot \sigma_r(t)$:

$S(\tau) = \gamma C(\sigma(t)) S_0$

$V = \frac{dp}{dt}$

$\nabla_V S = 0$

The condition is :

$\nabla_V S = [\gamma C(\sigma)] \left[\gamma C\left(\sigma^{-1} \cdot \frac{d\sigma}{d\tau} + \mathbf{Ad}_{\sigma^{-1}} \widehat{G}\right)\right] [S_0] = 0$

$\Leftrightarrow \sigma^{-1} \cdot \frac{d\sigma}{d\tau} + \mathbf{Ad}_{\sigma^{-1}} \widehat{G} = 0$

$\widehat{G} = -\mathbf{Ad}_\sigma \left(\sigma^{-1} \cdot \frac{d\sigma}{d\tau}\right) = -\sigma \cdot \sigma^{-1} \cdot \frac{d\sigma}{d\tau} \cdot \sigma^{-1} = -\frac{d\sigma}{d\tau} \cdot \sigma^{-1}$

$$\nabla_V S = 0 \Leftrightarrow \widehat{G} = -\frac{d\sigma}{d\tau} \cdot \sigma^{-1} \tag{5.30}$$

The path does not depend on S_0 : it is the same for any particle (but material bodies with a spatial extension are submitted to tidal forces). We come back to the previous conditions.

Even if it seems difficult for an observer to meet these requirements, the existence of such trajectories is an experimental fact. The gravitational field varies slowly with the location, and there is almost always a dominant source in the surroundings so that there is locally a privileged direction in the universe. This can be expressed as follows.

Take any field U of geodesics, and any observer. Then there is a preferred gauge $p_I \in \mathfrak{X}(P_G)$ such that :

$\nabla_U^G p_I = 0$

If we do the change of gauge :

$\mathbf{p}(m) = \varphi_G(m, 1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m) \cdot \chi(m)^{-1} = p_I(m) = \varphi_G(m, \sigma_I(m))$

$\nabla_U^G \widetilde{p}_I = \sum_{\alpha=0}^{3} \widetilde{G}_\alpha^a U^\alpha = 0$

with the new value of the potential $\widetilde{G}_\alpha^a$.

So, in a theoretical model, to check or compute some properties, one can always see what would happen if the observer was inertial, which could be convenient.

However we do not need to assume that the observer is inertial. For any observer, to get a standard gauge, any euclidean spatial basis will do. This feature can be checked at any point.

[4] An observer in the International Space Station can be considered as inertial, but obviously he is submitted to an acceleration which balances Earth gravity.

And the matrix $[Q]$ is the only quantity which has a physical meaning because it can be subjected to measures. It can be seen as a variable in the representation of the system, as the others. If there is a known motion, such a rotation of the observer with respect to the chart, impacting $[Q]$, it must be accounted for, but other way the evolution of $[Q]$ is seen as determined by the evolution of the system as a whole. In this meaning the variation of $[Q]$ over Ω is assumed to be determinist : it precludes only an unpredictable intervention of the observer, only motivated by the exercise of his free will. And it initial value $[Q(\varphi_o(0,x))]$ is truly arbitrary, and depends only on the choice of the observer.

Spinors

Spinors use a vector space E with basis $(e_i)_{i=1}^4$ and a gauge $\mathbf{e}_i(m) = (p(m), e_i) \in P_G[E, \gamma C]$. There is no physical meaning for $(e_i)_{i=1}^4$ and the choice of the gauge at each point depends on the free will of the observer : there is no constraint attached. The most important measurable quantity issued from the spinor, the energy, does not depend on the choice of the gauge. The spin is directly related to the choice of the gauge (it is expressed in E) but, actually only the variation of the spin, in an instrument such as a Stern-Gerlach analyzer, matters, and the choice of the basis in which the measures are done is indifferent. What is assumed is that the observer keeps his gauge constant, this can always be done and if not, of course the value of the measures should account for the changes.

Other fields

The same remark applies to the vector space F and the basis $(f_j)_{j=1}^n$, which is used only to differentiate the charges of the particles. So actually this is the particles themselves which provide a standard.

A last comment about a procedure common in electromagnetism or linearized gravity. When facing a complicated mathematical relation it is tempting to reduce it to a simpler form by calling to what is called gauge freedom. Actually this procedure uses the fact that the same quantity is expressed in different forms according to the gauge (in the fiber bundle definition). So one can replace one by another, which is equivalent, and better looking. To have any meaning this procedure shall follow the requirements of the change of gauge, clearly stated in the fiber bundle formalism. But we have to keep in mind that a change of gauge has a physical meaning, and an implication on the observer who does the measures. A change of gauge can be physically unacceptable by the constraints which would be imposed to the observer, and any experimental proof which would ignore these requirements in its protocols would be non valid.

5.4 THE PROPAGATION OF FIELDS

The physical phenomenon of propagation of fields is more subtle than it seems and, indeed, it was at the origin of the Relativity. In geometry it is not easy to quit the familiar framework of orthogonal frames with fixed origin, and similarly we are easily confused by the usual representation of a field emanating from a source, propagating at a certain speed, and decreasing with the distance. In this picture a "source" is a point, "speed" is related to the transmission of a signal, and "distance" is the euclidean distance with respect to the source. In a 4 dimensional universe, and notably when there is no source in the area which is studied, these words have no obvious meaning. The field that we perceive comes from sources which are beyond the scope of any bounded system, but we cannot discard their existence (after all we study the spectrum of stars, so their field is a physical entity). In experiments one can create fields which convey a signal, but this is limited to the electromagnetic field, and a signal means a specific variation in time, that is along one of the coordinates, which is specific to each observer. And the speed as well as the range are related to the euclidean distance between points in a given hypersurface. So to study the propagation of fields we will proceed as for geometry, avoiding to go straight to the usual representations, we will look carefully at the concepts, what they mean, and how we can find a pertinent mathematical representation.

The concept of a force field existing everywhere is one of the direct consequence of the Principle of Locality which prohibits action at a distance. From the beginning Faraday and Maxwell came naturally to the conclusion that the fields must be represented by variables whose value is determined locally. They should satisfy a set of local partial differential equations, of which the Maxwell equations are the paradigm. A field manifests its existence, and changes by interacting with particles, but it interacts also with itself and this is at the root of the phenomenon of propagation in the vacuum, where there is no particle. This self-interaction can be modelled with a lagrangian and leads to differential equations as expected as we will see in the next chapters.

The variable which represents the interaction of fields with particles is the connection, through its potential. If it is involved in differential equations we need a derivative. This is the strength of the field $\mathcal{F}$, similar to the electric and magnetic field, which is the key variable to represent the self-interaction of the field. In a dense medium where many interactions with particles occur $\mathcal{F}$ is replaced, in electromagnetism, by similar but different variables which account for these interactions. Here propagation will be seen only as the propagation in the vacuum.

Fields exist even in the vacuum and their value changes, from one point to another, in space and time, through their self interaction. In a relativist context, the distinction between past and future depends on the observer, so there is an issue. The answer depends on the philosophical point of view.

The value of the field is measured through its action on a known particle, so in a strict interpretation of classic QM, one could not say anything about a field before an interaction has occurred. In QTF particles are not localized, there is only a wave function associated to each particle, and at each location all virtual particles combined in a Fock space $\mathcal{H}$. An observable is an operator $P \in \mathcal{L}(\mathcal{H};\mathcal{H})$ acting in this space $\mathcal{H}$. Force fields appear as modifying the state of particles, and this modification is measured through an observable, thus force fields act on the operators representing the observables. It is conceivable to define a system by the algebra $\mathcal{A} \subset \mathcal{L}(\mathcal{H};\mathcal{H})$ of its observables, and force fields are similarly represented as operators acting on $\mathcal{A}$. A complication occurs because the action of the fields depends on the types of particles, so actually force fields are maps over M, valued on a space of distributions acting on spaces of test functions, which represent the waves functions of different kinds of particles : force fields are maps on M valued in the space dual of $\mathcal{H}$, as particles are maps on M valued in $\mathcal{H}$. Another complication comes from the causal structure of the universe. A field is assumed to propagate at the speed of light, and because fields are maps defined over M, the area where they can be active is restricted. This is dealt with through the support (the domain of M where they are not null) of either the wave functions or the operators. This picture has been formalized in the Wightman axioms (see Haag) with variants, which in some way constitute the extension of the "Axioms" of Quantum Mechanics to QTF. The issue of the extension of the force fields is solved (both particles and fields are maps defined all over M, and called "fields") but the concept of propagation vanishes. Actually everything happens at each point, through interactions of identified virtual particles and fields (fermions and bosons), in a picture which is similar to the traditional action at a distance. Most of the studies have been focused on finding solutions to the very complicated computations involved and recurring mathematical inconsistencies. QTF provides methods to represent the phenomena at the atomic and subatomic scale, but, restricted to the SR geometry and, almost by construct, it cannot deal efficiently with the Physics at another scale (whatever the effort).

In a realist point of view, the field is a physical entity. In GR locations are absolute. Particles have a distinct location, they travel along their world line and their proper time identifies uniquely their position. Physical fields are spread over M, they have a definite value at any point, however because of the propagation this value is restricted to an area in which it has already propagated. This area is delimited in the 4 dimensional universe by a 3 dimensional hypersurface, which is the front of the propagation, which moves as the field propagates.

For a given observer the distinction between future and past is clear : the field has a definite value, which can be measured, on any hypersurface $\Omega_3(t)$. In a system the initial value of $\mathcal{F}$ on $\Omega_3(0)$ "comes from the past", interactions with particles and propagation occur pointwise at each time and define the value of $\mathcal{F}$ on the hypersurfaces $\Omega_3(t)$. In a continuous model the field can be represented by a variable $\mathcal{F}(m)$ which has a physical meaning, and a precise space-time extension. Classic methods of the Theory of Fields provide the usual results.

But there is still the issue coming from the causal structure of the universe. This structure is common to all observers, and this fact is one of the motivation to introduce the Lorentz metric. If a physical field has a 4 dimensional extension in M, the area in which it can be perceived does not depend on the observer. We have already met a similar problem with solids. A solid can be defined as a set of material points which travel along world lines given by a common vector field, and such that there is some common proper time τ: at τ all the material points belong to the same hypersurface $\omega_3(\tau)$. So the solid can be considered as a physical entity, with a common structure and even a global spinor, it has a definite support in M, but observers do not see the same solid. For a solid the key is the existence of a common vector field, based on the assumption that the world lines of the particles constituting the solid do not cross. And because the particles have a velocity with Lorentz norm equal to $-c^2$ it is possible to relate the proper time τ of the solid with the time t of an observer. Similarly the field in the vacuum can be characterized by a proper time, the phase s ,which leads to introduce a function $F(m)$ similar to the function f_o of the observers, which defines a folliation of M by three dimensional hypersurfaces $F(m) = s$ and a vector field V along which the field propagates. Meanwhile for a particle there is a relation between τ and the time of an observer, given by the intersection of the world line with $\Omega_3(t)$, there is no such relation for the fields, the intersections are 2 dimensional surfaces, the front waves.

In a beam the field comes from an identified source which is dominant. But in more general systems the physical field is the sum of fields emanating from all the sources, even far away, in accordance with the superposition Huygen's principle, which is mathematically consistent because $\mathcal{F}$ is a vectorial quantity. Each interaction with a particle introduces a perturbation of the field, and the clean laws of propagation are not necessarily preserved. If the interactions can be represented continuously, either because they are small, or because, at the scale of the study one can consider a continuous distribution of particles, then propagation is a continuous process in which the field is represented by smooth variables. But one cannot expect that the perturbations introduced by the interactions with particles can be smeared out quickly in the immediate neighborhood of the particle. The concept of a continuous field is an approximation, which holds only at some scale, and at the atomic scale one cannot ignore the discontinuities which manifest as bosons This will be the topic of the last chapter with discontinuous processes.

The concept of a physical field existing in a definite region of the universe has another consequence. Practically the propagation of fields, specially in the vacuum, is a keystone in the measures of lengths and time, that is in the definition of charts of the universe (for instance we use the light emitted by a distant star to define a direction). The propagation of the field is related to the geometry of the universe : they must adapt to each other. On one hand the physical field provides a grid upon which we build our charts, assumed to be fixed, and on the other hand the tetrad cannot be constant in a fixed chart. So the deformation of the tetrad can be seen the other way around, as resulting from the necessary adjustment to a distorted grid, changing in space and time, provided by a physical field. And this explains the mechanism by which the geometry of the universe is impacted by its physical content.

5.4.1 The strength of the connection

The strength of the connection is a variable $\mathcal{F}$ which is a kind of derivative of the connection. It is related to the curvature, another mathematical object which is commonly used. We give its definition with some details, because they will be useful in the following. We will take $U, P_U, \grave{A}$ as example.

The principles

The tangent space of P_U is given by vectors :

$v_p = \varphi'_{Gm}(m,g)\, v_m + \varphi'_{G\varkappa}(m,g)\, v_g = \sum_{\alpha=0}^{3} v_m^\alpha \partial m_\alpha + \zeta(\theta)(p)$ with $\theta = L'_{g^{-1}g}(v_g)$

where the vertical space $VP_U = \ker \pi'_U = \left\{ \varphi'_{Ug}(m,g)\, v_g, v_g \in T_\chi U \right\}$, is isomorphic to the Lie algebra, does not depend on the trivialization, and is generated by fundamental vectors :

$\zeta : T_1 U \to VP :: \zeta(\theta)(\varphi_U(m,g)) = \varphi'_{Ug}(m,g)\, L'_{g^{-1}g}(\theta)$

with the property :

$\zeta(\theta)(\rho(p,g)) = \rho'_p(p,g)\, \zeta(Ad_g\theta)(p)$

A projectable vector field on TP_U is a vector field $W \in \mathfrak{X}(TP_U)$ such that :

$T\pi_U(W) = (\pi_U(p), \pi'_U(p)(W(p))) = (m, V(m)), V \in \mathfrak{X}(TM)$.

$W(p) = \varphi'_{Gm}(m,g)\, W_m(p) + \varphi'_{G\varkappa}(m,g)\, W_g(p)$

and W is projectable iff $W_m(p)$ does not depend on g : $V(m) = W_m(p)$.

There are holonomic bases of TP_U such that any vector $v_p \in T_pP$ can be uniquely written :

$v_p = \sum_{\alpha=0}^{3} v_m^\alpha \partial m_\alpha + \sum_{a=1}^{m} v_g^a \partial g_a$

where $\partial m_\alpha = \varphi'_{Um}(m,g)\, \partial\xi_\alpha$ with $(\partial\xi_\alpha)_{\alpha=0}^{3}$ a holonomic basis of TM. So that ∂m_α are projectable vector fields :

$\pi'_U(p)\, \partial m_\alpha = \partial\xi_\alpha$

and $\sum_{\alpha=0}^{3} v_m^\alpha \partial m_\alpha$ is a projectable vector field iff v_m^α does not depend on g.

The key object in the representation of the interactions fields / particles is the connection. This is a tensor, a one form $\grave{\mathbf{A}} \in \mathbf{\Lambda}_1(TP_U; VP_U)$ on TP valued in VP. For a principal connection its value depend on the potential $\grave{A}$:

$\grave{\mathbf{A}}(\mathbf{p}(m))(\varphi'_m(m,1)v_m+\zeta(\theta)(\mathbf{p}(m)))=\zeta\left(\theta+\sum_\alpha \grave{A}_\alpha(m)v^\alpha_m\right)(\mathbf{p}(m))$

$\grave{A}\in\Lambda_1(M;T_1U):TM\to T_1U::\grave{A}(m)=\sum_{\alpha=0}^3\sum_{a=1}^m \grave{A}^a_\alpha(m)\overrightarrow{\theta}_a\otimes d\xi^\alpha$

The connection form $\widehat{\grave{A}}$ of $\grave{\mathbf{A}}$ is :

$\widehat{\grave{A}}(p):T_pP_U\to T_1U::\grave{A}(p)(v_p)=\zeta\left(\widehat{\grave{A}}(p)(v_p)\right)(p)$

$\ker\grave{\mathbf{A}}$ is a subvector bundle HP_U of TP_U, its vectors are said to be horizontal, and :

$TP_U=HP_U\oplus VP_U$

$\pi'_U(VP_U)=0$

$\grave{\mathbf{A}}(HP_U)=0$

$\dim VP_U=\dim T_1U=m$

$\dim HP_U=\dim TP_U-\dim VP_U=\dim M=4$

$HP_U=\left\{\sum_{\alpha=0}^3 v^\alpha_m\partial m_\alpha+\zeta(\theta)(p):\theta+Ad_{g^{-1}}\sum_{\alpha=0}^3\grave{A}_\alpha(m)v^\alpha_m=0\right\}$

A r form $\lambda\in\Lambda_r(TP_U;F)$ on TP_U valued in a fixed vector space F is said to be horizontal if it is null for any vertical vector :

$\forall u_p\in VP_U:i_{u_p}\lambda=0$

The definition is independent of the existence of a connection.

It is expressed by :

$\lambda=\sum_{a=1}^m\sum_{\{\alpha_1,...\alpha_r\}=0}^3\lambda^a_{\alpha_1...\alpha_r}dm^{\alpha_1}\wedge....\wedge dm^{\alpha_r}\otimes\overrightarrow{\theta}_a$

The pull back of λ on TM is :

$\pi^*::\lambda\in TP_U^*\to\pi^*\lambda\in TM^*::\pi_U^*\lambda(m)(u_m)=\lambda(\pi(p))\pi'_U(p)u_p\Leftrightarrow\pi^*\lambda=\lambda(T\pi_U)$

$u_p\in VP_U\Leftrightarrow\pi'_U(p)u_p=0\Rightarrow\pi_U^*\lambda(m)(u_m)=0$

A connection can be equivalently defined by the horizontal form :

$\chi(p)::T_pP_U\to H_pP_U::\chi(p)(v_p)=v_p-\grave{\mathbf{A}}(p)(v_p)$

$\chi(p)\left(\sum_\alpha v^\alpha_m\partial m_\alpha+\zeta(\theta)(p)\right)=\sum_\alpha v^\alpha_m\partial m_\alpha-\zeta\left(Ad_{g^{-1}}\grave{A}(m)v_m\right)(p)$

χ is horizontal : $v_p\in V_pP_U:\grave{\mathbf{A}}(p)(v_p)=v_p\Rightarrow\chi(p)(v_p)=0$

If $u_p\in HP_U:\chi(p)(v_p)=v_p$

It enables to define two useful operations.

The horizontal lift of a vector field $V\in\mathfrak{X}(TM)$:

$\chi_L:\mathfrak{X}(TM)\to\mathfrak{X}(HP_U)::$

$\chi_L(p)(V)=\varphi'_{Gm}(m,g)V(m)-\zeta\left(Ad_{g^{-1}}\grave{A}(m)V(m)\right)(p)\in H_pP$

and $\chi_L(p)(V)$ is a vector field projectable on V : $\pi'_U(p)(\chi_L(p)(V))=V(\pi_U(p))$ as well as the commutator $\pi'_U(p)[\chi_L(p)(V_1),\chi_L(p)(V_2)]_{TP_U}=[V_1,V_2]_{TM}$

The horizontalization of any r form ω on TP_U valued in a fixed vector space F is the pull back of ω by χ :

$\chi^*(p):\Lambda_r(T_pP;V)\to\Lambda_r(H_pP;V)::\chi^*(p)\omega(p)(v_1,...,v_r)=\omega(p)(\chi(p)v_1,...,\chi(p)v_r)$

and the result is expressed by a form which depends only on dm^α :

$\chi^*(p)\omega(p)=\sum\mu_{\alpha_1...\alpha_r}(p)dm^{\alpha_1}\wedge...\wedge dm^{\alpha_r}$: it is null whenever a vector v_k is vertical, so that $\chi^*\widehat{\grave{A}}=0$.

We look for a derivative of the one form $\grave{\mathbf{A}}$, acting on TP_U. The method is similar to the one used in differential geometry to define the covariant derivative of forms valued in the tangent bundle. $\grave{\mathbf{A}}$ is expressed with components in dm^α,dg^a, but we are interested in its variation with m and not in its variation with a change of gauge (that we know anyway for a principal connection). One way to neutralize the component in dg^a is the horizontalization. Then the exterior covariant derivative associated to the connection is the map :

$\nabla_e : \Lambda_r (TP; F) \to \Lambda_{r+1} (TP; F) :: \nabla_e \omega = \chi^* (d\omega)$

where $d\omega$ is the exterior differential on TM (the components along dg^a have vanished).

It has the properties that we can expect from a derivative acting on forms :

$\forall \mu \in \Lambda_r (TP_U; F), \omega \in \Lambda_s (TM; F) : \nabla_e \mu \wedge \omega = \nabla_e \mu \wedge \omega + (-1)^r \mu \wedge \nabla_e \omega$

$\nabla_e \circ \pi_U^* = d \circ \pi_U^* = \pi_U^* \circ d$

The form $\widehat{\grave{A}}$ is valued in the fixed vector space T_1U, and the **strength of the field** is :

$$\mathcal{F}_A (m) = -\mathbf{p}^* (m) \nabla_e \widehat{\grave{A}} = -\mathbf{p}^* (m) \chi^* d\widehat{\grave{A}} \tag{5.31}$$

with the standard gauge $\mathbf{p} (m) = \varphi_U (m, 1)$.

$\mathbf{p}^* (m) \nabla_e \widehat{\grave{A}} \in \Lambda_2 (TM; T_1U) :: \mathbf{p}^* (m) \nabla_e \widehat{\grave{A}} (u_m, v_m) = \nabla_e \widehat{\grave{A}} (\varphi_U (m, 1)) (\varphi'_{Um} (m, 1) u_m, \varphi'_{Um} (m, 1) v_m)$

$\nabla_e \widehat{\grave{A}} = -\mathbf{p}_* \mathcal{F}_A$

It has the following expression :

$$\mathcal{F}_A = \sum_{a=1}^m \left(d \left(\sum_{\alpha=0}^3 \grave{A}_\alpha^a d\xi^\alpha \right) + \sum_{\alpha\beta} \left[\grave{A}_\alpha, \grave{A}_\beta \right] d\xi^\alpha \wedge d\xi^\beta \right) \otimes \overrightarrow{\theta}_a$$

where d is the exterior differential on TM and [] is the bracket in T_1U.

Equivalently with ordered indices :

$$\mathcal{F}_A = \sum_{a=1}^m \sum_{\{\alpha,\beta\}} \left(\mathcal{F}_{A\alpha\beta}^a d\xi^\alpha \wedge d\xi^\beta \right) \otimes \overrightarrow{\theta}_a \in \Lambda_2 (M; T_1U) \tag{5.32}$$

and in components :

$$\mathcal{F}_{A\alpha\beta}^a = \partial_\alpha \grave{A}_\beta^a - \partial_\beta \grave{A}_\alpha^a + 2 \left[\grave{A}_\alpha, \grave{A}_\beta \right]^a \tag{5.33}$$

Notice that the indices α, β are ordered, that it involves only the principal bundle, and not the associated vector bundles, and is valued in a fixed vector space. In this representation (with the basis $\left(\overrightarrow{\theta}_a \right)_{a=1}^6$) the group U acts through the map Ad (Maths.23.1.6).

In a change of gauge $\mathcal{F}_A$ changes as :

$$\left[\begin{array}{l} \mathbf{p}_U (m) = \varphi_{P_U} (m, 1) \to \widetilde{\mathbf{p}}_U (m) = \mathbf{p}_U (m) \cdot \varkappa (m)^{-1} : \\ \mathcal{F}_{A\alpha\beta} \to \widetilde{\mathcal{F}}_{A\alpha\beta} (m) = Ad_{\varkappa(m)} \mathcal{F}_{A\alpha\beta} \end{array} \right] \tag{5.34}$$

so that $\mathcal{F}_A$ can be seen as a 2 form on TM valued in the **adjoint bundle** $P_U [T_1U, Ad]$ (Maths.2161). This gives a more geometrical meaning to the concept, and we will see that these relations are crucial in the definition of the lagrangian.

Curvature

There is another introduction of the same concept, through the curvature, which is more usual but less immediate (we will not use it).

The curvature of the connection is the 2 form on P_U :

$\Omega \in \Lambda_2 (TP_U; VP_U) :: \Omega (p) (X, Y) = \zeta \left(\widehat{\Omega} (p) (X, Y)\right) (p) = \grave{\mathbf{A}} (p) \left([\chi (p) X, \chi (p) Y]_{TP_U}\right)$

where the bracket is the commutator of the vector fields $X, Y \in \mathfrak{X} (TP)$

The curvature form is the map such that : $\Omega (p) = \zeta \left(\widehat{\Omega} (p)\right) (p)$

$\widehat{\Omega} \in \Lambda_2 (TP_U; T_1U) : \widehat{\Omega} (p) = -Ad_{g^{-1}} \left(\sum_{a=1}^m \sum_{\alpha,\beta=0}^3 \left(\partial_\alpha \grave{A}^a_\beta + \left[\grave{A}_\alpha, \grave{A}_\beta\right]\right)\right) dm^\alpha \wedge dm^\beta \otimes \overrightarrow{\theta}_a$

where the bracket $\left[\grave{A}_\alpha, \grave{A}_\beta\right]$ is the bracket in the Lie algebra T_1U.

$\widehat{\Omega} = \nabla_e \widehat{\grave{A}}$

$\mathcal{F}_A$ can also be expressed as : $\mathcal{F}_A = -\mathbf{p}^* \widehat{\Omega}$ and because $\nabla_e \widehat{\grave{A}} = \widehat{\Omega} \Rightarrow \mathcal{F}_A = -\mathbf{p}^* \nabla_e \widehat{\grave{A}}$

$\mathcal{F}_A$ acts on TM and $\widehat{\Omega}$ on TP_U, but they are essentially the same 2 form, valued in the Lie algebra. We have the Bianchi identity : $\nabla_e \widehat{\Omega} = 0$.

For any vector fields $U, V \in \mathfrak{X} (TM)$

$\Omega (p) (\chi_L (U), \chi_L (V)) = [\chi_L (p) (U), \chi_L (p) (V)]_{TP} - \chi_L (p) ([U, V]_{TM})$

$= \rho'_p (\mathbf{p},g) \grave{\mathbf{A}} (\mathbf{p}) (v_m + \zeta (\theta) (\mathbf{p} (m)))$

$\nabla_U \circ \nabla_V - \nabla_V \circ \nabla_U = \nabla_{[U,V]} + \pounds_{\Omega(\chi_L(U),\chi_L(V))}$

Electromagnetic field

The strength of the electromagnetic field is a 2 form valued in $\mathbb{R} : \mathcal{F}_A \in \Lambda_2 (M; \mathbb{R})$.

Because the Lie algebra is abelian the bracket is null and : $\mathcal{F}_A = d\grave{A}$ which gives the first Maxwell's law : $d\mathcal{F}_A = 0$.

In a change of gauge : $\mathcal{F}_{A\alpha\beta} \to \widetilde{\mathcal{F}}_{A\alpha\beta} (m) = Ad_{\varkappa(m)} \mathcal{F}_{A\alpha\beta} = \mathcal{F}_{A\alpha\beta}$. The strength of the EM field is invariant in a change of gauge.

Gravitational field

We have the same quantities on $P_G (M, Spin (3, 1), \pi)$.

The strength of the connection is a two form on M valued in the Lie algebra $T_1Spin(3, 1)$ which reads with the basis $(\overrightarrow{\kappa}_a)_{a=1}^6$:

$$\left[\begin{array}{l} \mathcal{F}_G = \sum_{a=1}^6 \left(dG^a + \sum_{\alpha\beta=0}^3 [G_\alpha, G_\beta]^a d\xi^\alpha \wedge d\xi^\beta\right) \otimes \overrightarrow{\kappa}_a \\ \mathcal{F}_G = \sum_{a=1}^6 \sum_{\{\alpha,\beta\}=0}^3 \mathcal{F}^a_{G\alpha\beta} d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\kappa}_a \\ \mathcal{F}_G = \sum_{a=1}^6 \sum_{\{\alpha,\beta\}=0}^3 \left(\partial_\alpha G^a_\beta - \partial_\beta G^a_\alpha + 2 [G_\alpha, G_\beta]^a\right) d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\kappa}_a \end{array}\right] \tag{5.35}$$

where d is the exterior differential on TM and [] is the bracket in $T_1Spin(3, 1)$.[5]

Notice that :

i) in the last 2 formulas the indices α, β are ordered : $\mathcal{F}^a_{G\alpha\beta} = -\mathcal{F}^a_{G\beta\alpha}$

ii) it involves only the principal bundle, and not the associated vector bundles,

[5] The notations and conventions for r forms vary according to the authors and if the indices are ordered or not. On this see Maths.1525,1529.

iii) it is valued in a fixed vector space.

We can distinguish the two parts, $\mathcal{F}_r, \mathcal{F}_w$:
$\mathcal{F}_G = \sum_{\{\alpha,\beta\}=0}^{3} \upsilon \left(\mathcal{F}_{r\alpha\beta}, \mathcal{F}_{w\alpha\beta}\right) d\xi^\alpha \wedge d\xi^\beta$
$\mathcal{F}_G = d\left(\sum_{\alpha=0}^{3} \upsilon \left(G_{r\alpha}, G_{w\alpha}\right) d\xi^\alpha\right) + 2\sum_{\{\alpha\beta\}=0}^{3} \left[\upsilon \left(G_{r\alpha}, G_{w\alpha}\right), \upsilon \left(G_{r\beta}, G_{w\beta}\right)\right] d\xi^\alpha \wedge d\xi^\beta$
and we have :
a = 1,2,3 : $\mathcal{F}_{G\alpha\beta}^a = \mathcal{F}_{r\alpha\beta}^a$
a = 4,5,6 : $\mathcal{F}_{G\alpha\beta}^a = \mathcal{F}_{w\alpha\beta}^a$
with the signature (3,1) :

$$\left[\begin{array}{l} \mathcal{F}_G = \sum_{\{\alpha,\beta\}=0}^{3} \upsilon \left(\mathcal{F}_{r\alpha\beta}, \mathcal{F}_{w\alpha\beta}\right) d\xi^\alpha \wedge d\xi^\beta \\ \mathcal{F}_{r\alpha\beta} = \upsilon \left(\partial_\alpha G_{r\beta} - \partial_\beta G_{r\alpha} + 2\left(j\left(G_{r\alpha}\right) G_{r\beta} - j\left(G_{w\alpha}\right) G_{w\beta}\right), 0\right) \\ \mathcal{F}_{w\alpha\beta} = \upsilon \left(0, \partial_\alpha G_{w\beta} - \partial_\beta G_{w\alpha} + 2\left(j\left(G_{w\alpha}\right) G_{r\beta} + j\left(G_{r\alpha}\right) G_{w\beta}\right)\right) \end{array}\right] \tag{5.36}$$

With the signature (1,3):
$\mathcal{F}_{r\alpha\beta} = -\upsilon \left(\partial_\alpha G_{r\beta} - \partial_\beta G_{r\alpha} + 2\left(j\left(G_{r\alpha}\right) G_{r\beta} - j\left(G_{w\alpha}\right) G_{w\beta}\right), 0\right)$
$\mathcal{F}_{w\alpha\beta} = -\upsilon \left(0, \partial_\alpha G_{w\beta} - \partial_\beta G_{w\alpha} + 2\left(j\left(G_{w\alpha}\right) G_{r\beta} + j\left(G_{r\alpha}\right) G_{w\beta}\right)\right)$

In this representation (with the basis $(\overrightarrow{\kappa}_a)_{a=1}^6$) the group $Spin\left(3,1\right)$ acts through the map **Ad**, and the action is given by 6×6 matrices seen previously. In a change of gauge on the principal bundle the strength changes as :

$$\left[\begin{array}{l} \mathbf{p}_G\left(m\right) = \varphi_G\left(m,1\right) \to \widetilde{\mathbf{p}}_G\left(m\right) = \mathbf{p}_G\left(m\right) \cdot s\left(m\right)^{-1} : \\ \mathcal{F}_{G\alpha\beta} \to \widetilde{\mathcal{F}}_{G\alpha\beta}\left(m\right) = \mathbf{Ad}_{s(m)} \mathcal{F}_{G\alpha\beta} \\ \upsilon\left(\widetilde{\mathcal{F}}_{r\alpha\beta}, \widetilde{\mathcal{F}}_{w\alpha\beta}\right) = \mathbf{Ad}_{s(m)} \upsilon\left(\mathcal{F}_{r\alpha\beta}, \mathcal{F}_{w\alpha\beta}\right) \end{array}\right] \tag{5.37}$$

$$\begin{bmatrix} \widetilde{\mathcal{F}}_{r\alpha\beta} \\ \widetilde{\mathcal{F}}_{w\alpha\beta} \end{bmatrix} = \left[\mathbf{Ad}_{s(m)}\right] \begin{bmatrix} \mathcal{F}_{r\alpha\beta} \\ \mathcal{F}_{w\alpha\beta} \end{bmatrix} = \begin{bmatrix} A & -B \\ B & A \end{bmatrix} \begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix} \begin{bmatrix} \mathcal{F}_{r\alpha\beta} \\ \mathcal{F}_{w\alpha\beta} \end{bmatrix}$$

This feature allows to consider the strength as sections of the adjoint bundle $P_G\left[T_1 Spin(3,1), \mathbf{Ad}\right]$ using the representation of the groups on their Lie algebra through the adjoint map (Ad on $T_1 Spin(3,1)$ is identical to **Ad**) (Maths.2161).

5.4.2 Scalar curvature

In GR another definition of curvature is commonly used, and it is necessary to see how these concepts are related. For this it is useful to use the matrix representation of $\mathcal{F}_G$.

Matrix representation

$\mathcal{F}_G$ can be written in matrix form using the standard representation of $T_1 Spin\left(3,1\right)$ on its Lie algebra (Maths.24.1.3). Then the matrices $\left[\mathcal{F}_G\right]$ representing $\mathcal{F}_G$ belong to $so(3,1)$, and using $\overrightarrow{\kappa}_a \to \left[\kappa_a\right]$:

$$[\mathcal{F}_{G\alpha\beta}] = \sum_{a=1}^{6} \mathcal{F}^a_{G\alpha\beta} [\kappa_a] = [K(\mathcal{F}_{w\alpha\beta})] + [J(\mathcal{F}_{r\alpha\beta})] \tag{5.38}$$

$$[\mathcal{F}_{\alpha\beta}] = \begin{bmatrix} 0 & \mathcal{F}^1_{w\alpha\beta} & \mathcal{F}^2_{w\alpha\beta} & \mathcal{F}^3_{w\alpha\beta} \\ \mathcal{F}^1_{w\alpha\beta} & 0 & -\mathcal{F}^3_{r\alpha\beta} & \mathcal{F}^2_{r\alpha\beta} \\ \mathcal{F}^2_{w\alpha\beta} & \mathcal{F}^3_{r\alpha\beta} & 0 & -\mathcal{F}^1_{\alpha\beta} \\ \mathcal{F}^3_{w\alpha\beta} & -\mathcal{F}^2_{r\alpha\beta} & \mathcal{F}^1_{r\alpha\beta} & 0 \end{bmatrix} = \begin{bmatrix} 0 & \mathcal{F}_{w\alpha\beta} \\ \mathcal{F}_{w\alpha\beta} & j(\mathcal{F}_{r\alpha\beta}) \end{bmatrix}$$

which underlines the rotational feature of the component $\mathcal{F}_r$, and the transversal aspect of the component $\mathcal{F}_w$.

With :

$$\Gamma_{M\alpha} = [K(G_{w\alpha})] + [J(G_{r\alpha})] = \begin{bmatrix} 0 & G^1_{w\alpha} & G^2_{w\alpha} & G^3_{w\alpha} \\ G^1_{w\alpha} & 0 & -G^3_{r\alpha} & G^2_{r\alpha} \\ G^2_{w\alpha} & G^3_{r\alpha} & 0 & -G^1_{\alpha} \\ G^3_{w\alpha} & -G^2_{r\alpha} & G^1_{r\alpha} & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & G_{w\alpha} \\ G_{w\alpha} & j(G_{r\alpha}) \end{bmatrix}$$

$$[\mathcal{F}_{\alpha\beta}] = [\partial_\alpha \Gamma_{M\beta}] - [\partial_\beta \Gamma_{M\alpha}] + [\Gamma_{M\alpha}][\Gamma_{M\beta}] - [\Gamma_{M\beta}][\Gamma_{M\alpha}] \tag{5.39}$$

In this representation the group $Spin(3,1)$ acts through the conjugation of matrices :

$Conj : Spin(3,1) \times so(3,1) \to so(3,1) : Conj_{[g]}[X] = [g][X][g]^{-1}$

Riemann curvature of a principal connection

A curvature on a principal bundle has for image on any *associated vector bundle* a quantity, called **Riemann curvature** (Maths.2203), which is a two-form on M, valued in the endomorphisms on the vector space. In the case of $P_G[\mathbb{R}^4, \mathbf{Ad}]$ it is expressed by the 4 tensor (with ordered indices α, β) :

$R = \sum_{\{\alpha\beta\}ij} R^i_{\alpha\beta j} d\xi^\alpha \wedge d\xi^\beta \otimes \varepsilon_i(m) \otimes \varepsilon^j(m)$ with the dual basis $\varepsilon^j(m)$

$R^i_{\alpha\beta j} = \partial_\alpha \Gamma^i_{M\beta j} - \partial_\beta \Gamma^i_{M\alpha j} + \sum_{k=0}^3 \left(\Gamma^i_{M\alpha k} \Gamma^k_{M\beta j} - \Gamma^i_{M\beta k} \Gamma^k_{M\alpha j} \right)$

In matrix form :

$[R_{\alpha\beta}]^i_j = \left([\partial_\alpha \Gamma_{M\beta}] - [\partial_\beta \Gamma_{M\alpha}] + [\Gamma_{M\alpha}][\Gamma_{M\beta}] - [\Gamma_{M\beta}][\Gamma_{M\alpha}] \right)^i_j = [\mathcal{F}_{\alpha\beta}]^i_j$

$R = \sum_{\{\alpha\beta\}ij} [\mathcal{F}_{\alpha\beta}]^i_j d\xi^\alpha \wedge d\xi^\beta \otimes \varepsilon_i(m) \otimes \varepsilon^j(m)$

The Riemann curvature is the image of the strength of the field on $P_G[\mathbb{R}^4, \mathbf{Ad}]$. This is the same quantity, but in the representation of $T_1 Spin(3,1)$ in the matrix algebra $so(3,1)$.

By construct this quantity is covariant (in a change of chart on M) and equivariant (in a change of gauge on P_G) :

In a change of gauge :

$\mathbf{p}_G(m) = \varphi_G(m,1) \to \widetilde{\mathbf{p}}_G(m) = \mathbf{p}_G(m) \cdot s(m)^{-1}$:

$[\mathcal{F}_{\alpha\beta}]^i_j \to \frac{1}{2} \sum_{kl=0}^3 [h(s)]^i_k [\mathcal{F}_{\alpha\beta}]^k_l \left[h\left(s^{-1}\right)\right]^l_j$

with $[h(s)]$ the matrix of $SO(3,1)$ associated to $s \in Spin(3,1)$

$R \to \widetilde{R} = \sum_{\{\alpha\beta\}ij} \left[\widetilde{\mathcal{F}}_{\alpha\beta}\right]^i_j d\xi^\alpha \wedge d\xi^\beta \otimes \widetilde{\varepsilon}_i(m) \otimes \widetilde{\varepsilon}^j(m)$

$= \sum_{\{\alpha\beta\}ij} [h(s)]^i_k [\mathcal{F}_{\alpha\beta}]^k_l \left[h\left(s^{-1}\right)\right]^l_j d\xi^\alpha \wedge d\xi^\beta \otimes \left[h\left(s^{-1}\right)\right]^p_i \varepsilon_p(m) \otimes [h(s)]^j_q \varepsilon^q(m)$

$= \sum_{\{\alpha\beta\}ij} [\mathcal{F}_{\alpha\beta}]^p_q d\xi^\alpha \wedge d\xi^\beta \otimes \varepsilon_p(m) \otimes \varepsilon^q(m)$

so $\widetilde{R} = R$

It can be expressed in the holonomic basis of any chart on M using the tetrad :

$\varepsilon_i(m) = \sum_{\gamma=0}^{3} P_i^{\gamma} \partial\xi_\gamma$
$\varepsilon^j(m) = \sum_{\eta=0}^{3} P_\eta'^j d\xi^\eta$
$R = \sum_{\{\alpha\beta\}\gamma\eta} ([P][\mathcal{F}_{G\alpha\beta}][P'])_\eta^\gamma d\xi^\alpha \wedge d\xi^\beta \otimes \partial\xi_\gamma \otimes d\xi^\eta$
So we have the steps :
Principal connection $\mathbf{G}$ → Riemann curvature R on $P_G\left[\mathbb{R}^4, \mathbf{Ad}\right]$ → Riemann curvature R on TM in any chart
and the Riemann curvature R on TM is the same object as the strength of the connection $\mathcal{F}$, but expressed in matrix form in any holonomic basis of a chart.

Riemann tensor of an affine connection

With a common affine connection $\widehat{\Gamma}_\alpha$ on TM one can also define similarly a **Riemann tensor** (Maths.1543) :
$\widehat{R} = \sum_{\{\alpha\beta\}} \sum_{\gamma\eta} \widehat{R}_{\alpha\beta\eta}^{\gamma} d\xi^\alpha \wedge d\xi^\beta \otimes \partial\xi_\gamma \otimes d\xi^\eta$
which, expressed in matrix form with : $\widehat{R}_{\alpha\beta\eta}^{\gamma} = \left[\widehat{R}_{\alpha\beta}\right]_\eta^\gamma$, reads :
$\left[\widehat{R}_{\alpha\beta}\right] = \left[\partial_\alpha \widehat{\Gamma}_\beta\right] - \left[\partial_\beta \widehat{\Gamma}_\alpha\right] + \left[\widehat{\Gamma}_\alpha\right]\left[\widehat{\Gamma}_\beta\right] - \left[\widehat{\Gamma}_\beta\right]\left[\widehat{\Gamma}_\alpha\right]$
$\widehat{R} = \sum_{\{\alpha\beta\}} \sum_{\gamma\eta} \widehat{R}_{\alpha\beta\eta}^{\gamma} d\xi^\alpha \wedge d\xi^\beta \otimes \partial\xi_\gamma \otimes d\xi^\eta$
When we take as affine connection the one which is deduced from $\mathbf{G}$:
$\widehat{\Gamma}_{\alpha\beta}^{\gamma} = \left[\widehat{\Gamma}_\alpha\right]_\beta^\gamma = ([P]([\partial_\alpha P'] + [\Gamma_{M\alpha}][P']))_\beta^\gamma$
we get the same result :

$$\left[\widehat{R}_{\alpha\beta}\right] = [R_{\alpha\beta}] = [P][\mathcal{F}_{G\alpha\beta}][P'] \Leftrightarrow [\mathcal{F}_{G\alpha\beta}] = [P'][R_{\alpha\beta}][P] \tag{5.40}$$

Proof. $\left[\widehat{R}_{\alpha\beta}\right]$
$= [\partial_\alpha P][\partial_\beta P'] + [\partial_\alpha P][\Gamma_{M\beta}][P'] + [P]\left[\partial_{\beta\alpha}^2 P'\right] + [P][\partial_\alpha \Gamma_{M\beta}][P']$
$+ [P][\Gamma_{M\beta}][\partial_\alpha P'] - [\partial_\beta P][\partial_\alpha P'] - [\partial_\beta P][\Gamma_{M\alpha}][P'] - [P]\left[\partial_{\alpha\beta}^2 P'\right]$
$- [P][\partial_\beta \Gamma_{M\alpha}][P'] - [P][\Gamma_{M\alpha}][\partial_\beta P'] + [P][\partial_\alpha P'][P][\partial_\beta P']$
$+ [P][\Gamma_{M\alpha}][P'][P][\partial_\beta P'] + [P][\partial_\alpha P'][P][\Gamma_{M\beta}][P']$
$+ [P][\Gamma_{M\alpha}][P'][P][\Gamma_{M\beta}][P'] - [P][\partial_\beta P'][P][\partial_\alpha P']$
$- [P][\Gamma_{M\beta}][P'][P][\partial_\alpha P'] - [P][\partial_\beta P'][P][\Gamma_{M\alpha}][P']$
$- [P][\Gamma_{M\beta}][P'][P][\Gamma_{M\alpha}][P']$
$= +[P]([\partial_\alpha \Gamma_{M\beta}] - [\partial_\beta \Gamma_{M\alpha}] + [\Gamma_{G\alpha}][\Gamma_{M\beta}] - [\Gamma_{M\beta}][\Gamma_{M\alpha}])[P']$
$+ [\partial_\alpha P][\partial_\beta P'] - [\partial_\beta P][\partial_\alpha P'] + [P][\partial_\alpha P'][P][\partial_\beta P']$
$- [P][\partial_\beta P'][P][\partial_\alpha P'] + [\partial_\alpha P][\Gamma_{M\beta}][P'] - [\partial_\beta P][\Gamma_{M\alpha}][P']$
$+ [P][\Gamma_{M\beta}][\partial_\alpha P'] - [P][\Gamma_{M\alpha}][\partial_\beta P'] + [P][\Gamma_{M\alpha}][\partial_\beta P']$
$- [P][\Gamma_{M\beta}][\partial_\alpha P'] + [P][\partial_\alpha P'][P][\Gamma_{M\beta}][P'] - [P][\partial_\beta P'][P][\Gamma_{M\alpha}][P']$
$= [P][\mathcal{F}_{G\alpha\beta}][P'] + [\partial_\alpha P][\partial_\beta P'] - [\partial_\beta P][\partial_\alpha P']$
$- [\partial_\alpha P][P'][P][\partial_\beta P'] + [\partial_\beta P][P'][P][\partial_\alpha P'] + [\partial_\alpha P][\Gamma_{M\beta}][P']$
$- [\partial_\beta P][\Gamma_{M\alpha}][P'] + [P][\Gamma_{M\beta}][\partial_\alpha P'] - [P][\Gamma_{M\alpha}][\partial_\beta P'] + [P][\Gamma_{M\alpha}][\partial_\beta P']$
$- [P][\Gamma_{M\beta}][\partial_\alpha P'] - [\partial_\alpha P][P'][P][\Gamma_{M\beta}][P'] + [\partial_\beta P][P'][P][\Gamma_{M\alpha}][P']$
$= [P][\mathcal{F}_{G\alpha\beta}][P']$
with $[P][\partial_\alpha P'] + [\partial_\alpha P][P'] = 0$ ■

So the Riemann tensor is the Riemann curvature of the principal connection, expressed in the holonomic basis of a chart, and it is the same object as the strength of the connection :
$R = \sum_{\{\alpha\beta\}ij} \sum_{a=1}^{6} \mathcal{F}_{G\alpha\beta}^{a} [\kappa_a]_j^i d\xi^\alpha \wedge d\xi^\beta \otimes \varepsilon_i(m) \otimes \varepsilon^j(m)$

$= \sum_{\{\alpha\beta\}ij} \sum_{a=1}^{6} \mathcal{F}^a_{G\alpha\beta} \left([P] [\kappa_a] [P'] \right)^\gamma_\eta d\xi^\alpha \wedge d\xi^\beta \otimes \partial\xi_\gamma \otimes d\xi^\eta$

The Riemann tensor can be computed with any affine connection, as well as with any principal connection. In the usual RG formalism the Riemann tensor is computed with a special connection : the Levy-Civita connection.

The Riemann tensor is antisymmetric, in the meaning :

$R_{\alpha\beta\gamma\eta} = -R_{\alpha\beta\eta\gamma}$ with $R_{\alpha\beta\gamma\eta} = \sum_\lambda R^\lambda_{\alpha\beta\gamma} g_{\lambda\eta}$

$[\mathcal{F}_{G\alpha\beta}] \in so(3,1)$ so $[\eta] [\mathcal{F}_{G\alpha\beta}] + [\mathcal{F}_{G\alpha\beta}]^t [\eta] = 0$ and

$R_{\alpha\beta\gamma\eta} = \sum_\lambda R^\lambda_{\alpha\beta\gamma} g_{\lambda\eta} = \sum_\lambda \left([P] [\mathcal{F}_{G\alpha\beta}] [P'] \right)^\lambda_\gamma g_{\lambda\eta} = \left([P']^t [\eta] [\mathcal{F}_{G\alpha\beta}] [P'] \right)^\eta_\gamma$

$= \left(\left([P']^t [\eta] [\mathcal{F}_{G\alpha\beta}] [P'] \right)^t \right)^\gamma_\eta = \left([P']^t [\mathcal{F}_{G\alpha\beta}]^t [\eta] [P'] \right)^\gamma_\eta$

$= -\left([P']^t [\eta] [\mathcal{F}_{G\alpha\beta}] [P'] \right)^\gamma_\eta = -R_{\alpha\beta\eta\gamma}$

Thus this symmetry is not specific to the Lévi-Civita connection as it is usually assumed (Wald p.39).

Ricci tensor and scalar curvature

The Riemann tensor R, coming from any connection, is a 2 form but can be expressed as an antisymmetric tensor with non ordered indices with $d\xi^\alpha \wedge d\xi^\beta = d\xi^\alpha \otimes d\xi^\beta - d\xi^\beta \otimes d\xi^\alpha$

$R = \sum_{\alpha\beta\gamma\eta} [R_{\alpha\beta}]^\gamma_\eta d\xi^\alpha \otimes d\xi^\beta \otimes \partial\xi_\gamma \otimes d\xi^\eta$

and we can contract the covariant index α, β or η with the contravariant index γ . The result does not depend on a basis : it is covariant (Maths.385). The different solutions give :

$\alpha : \sum_{\beta\eta} \left(\sum_\alpha [R_{\alpha\beta}]^\alpha_\eta \right) d\xi^\beta \otimes d\xi^\eta$

$\beta : \sum_{\alpha\eta} \left(\sum_\beta [R_{\alpha\beta}]^\beta_\eta \right) d\xi^\alpha \otimes d\xi^\eta$

$\eta : \sum_{\alpha\eta} \left(\sum_\gamma [R_{\alpha\beta}]^\gamma_\gamma \right) d\xi^\alpha \otimes d\xi^\beta$

The last solution has no interest because :

$Tr\left([P] [\mathcal{F}_{G\alpha\beta}] [P'] \right) = Tr\left([\mathcal{F}_{G\alpha\beta}] [P'] [P] \right) = Tr\left([\mathcal{F}_{G\alpha\beta}] \right) = 0$

The first two read :

$\sum_{\beta\gamma} [P]^\alpha_k [\mathcal{F}_{G\alpha\beta}]^k_l [P']^l_\eta [P]^\beta_i \varepsilon^i \otimes [P]^\eta_j \varepsilon^j = \sum_{\beta\gamma} [P]^\alpha_k [\mathcal{F}_{G\alpha\beta}]^k_j [P]^\beta_i \varepsilon^i \otimes \varepsilon^j = \sum_{\alpha\beta j} \left([P] [\mathcal{F}_{G\alpha\beta}] \right)^\alpha_j d\xi^\beta \otimes \varepsilon^j$

$\sum_{\alpha\gamma} [P]^\beta_k [\mathcal{F}_{G\alpha\beta}]^k_l [P']^l_\eta [P]^\alpha_i \varepsilon^i \otimes [P]^\eta_j \varepsilon^j = \sum_{\alpha\gamma} [P]^\beta_k [\mathcal{F}_{G\alpha\beta}]^k_j [P]^\alpha_i \varepsilon^i \otimes \varepsilon^j = \sum_{\beta\gamma} \left([P] [\mathcal{F}_{G\alpha\beta}] \right)^\beta_j d\xi^\alpha \otimes \varepsilon^j$

The **Ricci tensor** is the contraction on the two indices γ, β of R :

$Ric = \sum_{\alpha\eta} Ric_{\alpha\eta} d\xi^\alpha \otimes d\xi^\eta = \sum_{\alpha\eta} \left(\sum_\beta [R_{\alpha\beta}]^\beta_\eta \right) d\xi^\alpha \otimes d\xi^\eta$

This is a tensor, from which one can compute another tensor by lowering the last index:

$\sum_\lambda g^{\eta\lambda} Ric_{\alpha\eta} d\xi^\alpha \otimes d\xi^\eta = \sum_{\alpha\lambda} Ric^\lambda_\alpha d\xi^\alpha \otimes \partial\xi_\lambda$

whose contraction (called the trace of this tensor) provides the **scalar curvature** :

$\mathbf{R} = \sum_\alpha Ric^\alpha_\alpha = \sum_{\alpha\beta\eta} g^{\alpha\eta} [R_{\alpha\beta}]^\beta_\eta$

The same procedure applied to the contraction on the two indices γ, α of R gives the opposite scalar :

$\mathbf{R} = \sum_{\alpha\beta\eta} g^{\beta\eta} [R_{\alpha\beta}]^\alpha_\eta = -\sum_{\alpha\beta\eta} g^{\alpha\eta} [R_{\beta\alpha}]^\beta_\eta = -\sum_{\alpha\beta\eta} g^{\alpha\eta} [R_{\alpha\beta}]^\beta_\eta$

This manipulation is mathematically valid, and provides a unique scalar, which does not depend on a chart, and can be used in a lagrangian. However its physical justification (see Wald) is weak.

In the usual GR formalism the scalar curvature is computed with the Riemann tensor $\widehat{R}$ deduced from the Levy-Civita connection but, as we can see, it can be computed in the tetrad with any principal connection.

Starting from $[R_{\alpha\beta}] = [P][\mathcal{F}_{G\alpha\beta}][P']$ one gets the Ricci tensor :
$Ric = \sum_{\alpha\beta} Ric_{\alpha\beta} d\xi^\alpha \otimes d\xi^\beta = \sum_{\alpha\beta}\sum_\gamma ([P][\mathcal{F}_{G\alpha\gamma}][P'])^\gamma_\beta d\xi^\alpha \otimes d\xi^\beta$

$$Ric = \sum_{\alpha\beta\gamma} ([P][\mathcal{F}_{G\alpha\gamma}][P'])^\gamma_\beta d\xi^\alpha \otimes d\xi^\beta$$

and the scalar curvature :
$\mathbf{R} = \sum_{\alpha\beta\gamma} g^{\alpha\gamma}[R_{\alpha\beta}]^\beta_\gamma = \sum_{\alpha\beta\gamma} g^{\alpha\gamma}([P][\mathcal{F}_{G\alpha\beta}][P'])^\beta_\gamma$ and with $[g]^{-1} = [P][\eta][P]^t$
$\mathbf{R} = \sum_{\alpha\beta\gamma}\left([P][\eta][P]^t\right)^\gamma_\alpha ([P][\mathcal{F}_{G\alpha\beta}][P'])^\beta_\gamma = \sum_{\alpha\beta}\left([P][\mathcal{F}_{G\alpha\beta}][\eta][P]^t\right)^\beta_\alpha$

$$\mathbf{R} = \sum_{\alpha\beta}\sum_{a=1}^3 \mathcal{F}^a_{r\alpha\beta}\left([P][\kappa_a][\eta][P]^t\right)^\beta_\alpha + \mathcal{F}^a_{w\alpha\beta}\left([P][\kappa_{a+3}][\eta][P]^t\right)^\beta_\alpha \tag{5.41}$$

This expression has two important features :

- the scalar curvature is linear with respect to the strength of the field. In the implementation of the Principle of Least Action it provides equations which are linear with respect to $\mathcal{F}_G$, which is a big improvement from the usual computations.

- it shows that *the scalar curvature has a transversal component and a rotational component.* This happens for any scalar curvature, but is just masked in the usual expression through the metric. This feature is not without significance, as it is related to the distinction between the space and the time Universe.

To sum up, with the fiber bundle and connections formalism it is possible to compute, more easily, a scalar curvature which has the usual meaning. And by imposing symmetry to the affine connection we get exactly the same quantity. However, as we have seen before, the symmetry of the connection has no obvious physical meaning, and similarly for the scalar curvature.

5.4.3 Energy

The field interacts with itself, during its propagation, and in this process the value of $\mathcal{F}$ changes locally, so it is rational to look for a quantity, similar to the "energy of the particles", to represent this process. It should involve only $\mathcal{F}$, the tetrad and be independent of the choice of a chart or a gauge, and have as a simple expression as possible. For the gravitational field the scalar curvature can be used for this purpose, and this is the usual solution, however it has no equivalent for the other fields, and it has other drawbacks (it does not represent the totality of the field as we will see). So we will look for a general solution, encompassing all fields, which leads to a scalar product $\langle \mathcal{F}, \mathcal{F} \rangle$, as $\mathcal{F}$ is a vectorial quantity.

Notations

A two form $\mathcal{F} \in \Lambda_2(M;\mathbb{R})$ can be written : $\mathcal{F} = \frac{1}{2}\sum_{\alpha,\beta=0}^3 \mathcal{F}_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta$ with non ordered indices or $\mathcal{F} = \sum_{\{\alpha,\beta\}=0}^3 \mathcal{F}_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta$ with ordered indices. The first is more usual in Differential Geometry, the second in Algebra. It will be convenient in the following to use a precise order of the indices. One can always write :

$\mathcal{F} = \mathcal{F}^r + \mathcal{F}^w$
with
$\mathcal{F}^r = \mathcal{F}_{32} d\xi^3 \wedge d\xi^2 + \mathcal{F}_{13} d\xi^1 \wedge d\xi^3 + \mathcal{F}_{21} d\xi^2 \wedge d\xi^1$
$\mathcal{F}^w = \mathcal{F}_{01} d\xi^0 \wedge d\xi^1 + \mathcal{F}_{02} d\xi^0 \wedge d\xi^2 + \mathcal{F}_{03} d\xi^0 \wedge d\xi^3$
and we will denote the 1×3 row matrices :

$$[\mathcal{F}^r] = \begin{bmatrix} \mathcal{F}_{32} & \mathcal{F}_{13} & \mathcal{F}_{21} \end{bmatrix}; [\mathcal{F}^w] = \begin{bmatrix} \mathcal{F}_{01} & \mathcal{F}_{02} & \mathcal{F}_{03} \end{bmatrix} \quad (5.42)$$

With this notation :
$\mathcal{F} = \mathcal{F}^r + \mathcal{F}^w; K = K^r + K^w$
$\mathcal{F} \wedge K = -\left([\mathcal{F}^r] [K^w]^t + [\mathcal{F}^w] [K^r]^t \right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$
$\mathcal{F}(V, W) = [\mathcal{F}^w] \left(W^0 [v] - V^0 [w] \right) + [w]^t j (\mathcal{F}^r) [v]$ with $[V] = \begin{bmatrix} V^0 \\ [v] \end{bmatrix}$

The strength of the field is valued in the Lie algebra. And we will denote similarly :
$[\mathcal{F}]_{6\times 6} = \begin{bmatrix} \mathcal{F}_r^r & \mathcal{F}_r^w \\ \mathcal{F}_w^r & \mathcal{F}_w^w \end{bmatrix} = \left[\mathcal{F}^a_{G\alpha\beta} \right]$
with the 3×3 matrices :
$[\mathcal{F}_r^r]_{3\times 3} = \begin{bmatrix} \mathcal{F}^1_{G32} & \mathcal{F}^1_{G13} & \mathcal{F}^1_{G21} \\ \mathcal{F}^2_{G32} & \mathcal{F}^2_{G13} & \mathcal{F}^2_{G21} \\ \mathcal{F}^3_{G32} & \mathcal{F}^3_{G13} & \mathcal{F}^3_{G21} \end{bmatrix}$
$[\mathcal{F}_r^w]_{3\times 3} = \begin{bmatrix} \mathcal{F}^1_{G01} & \mathcal{F}^1_{G02} & \mathcal{F}^1_{G03} \\ \mathcal{F}^2_{G01} & \mathcal{F}^2_{G02} & \mathcal{F}^2_{G02} \\ \mathcal{F}^3_{G01} & \mathcal{F}^3_{G02} & \mathcal{F}^3_{G03} \end{bmatrix}$
$[\mathcal{F}_w^r]_{3\times 3} = \begin{bmatrix} \mathcal{F}^4_{G32} & \mathcal{F}^4_{G13} & \mathcal{F}^4_{G21} \\ \mathcal{F}^5_{G32} & \mathcal{F}^5_{G13} & \mathcal{F}^5_{G21} \\ \mathcal{F}^6_{G32} & \mathcal{F}^6_{G13} & \mathcal{F}^6_{G21} \end{bmatrix}$
$[\mathcal{F}_w^w]_{3\times 3} = \begin{bmatrix} \mathcal{F}^4_{G01} & \mathcal{F}^4_{G02} & \mathcal{F}^4_{G03} \\ \mathcal{F}^5_{G01} & \mathcal{F}^5_{G02} & \mathcal{F}^5_{G02} \\ \mathcal{F}^6_{G01} & \mathcal{F}^6_{G02} & \mathcal{F}^6_{G03} \end{bmatrix}$
And for the other fields :
$[\mathcal{F}]_{m\times 6} = \begin{bmatrix} \mathcal{F}_A^r & \mathcal{F}_A^w \end{bmatrix} = \left[\mathcal{F}^a_{A\alpha\beta} \right]$
with $m \times 3$ matrices :
$[\mathcal{F}_A^r]_{m\times 3} = \begin{bmatrix} \mathcal{F}^1_{A32} & \mathcal{F}^1_{A13} & \mathcal{F}^1_{A21} \\ ... & ... & ... \\ \mathcal{F}^m_{A32} & \mathcal{F}^m_{G13} & \mathcal{F}^m_{G21} \end{bmatrix}$
$[\mathcal{F}_A^w]_{m\times 3} = \begin{bmatrix} \mathcal{F}^1_{A01} & \mathcal{F}^1_{A02} & \mathcal{F}^1_{A03} \\ ... & ... & ... \\ \mathcal{F}^m_{A01} & \mathcal{F}^m_{A02} & \mathcal{F}^m_{A03} \end{bmatrix}$

Impact of a change of chart

The split in the two parts $\mathcal{F}^r, \mathcal{F}^w$ does not change in a change of space basis (the vectors $(\partial\xi_\alpha)_{\alpha=1}^3$), that is for a given observer, but changes for another observer who has not the same **O**. It is useful to see how it changes in a change of chart.

If $\partial\xi_\alpha \to \partial\eta_\alpha$ a 2 form changes as :
$\mathcal{F}^a_{\alpha\beta} \to \widetilde{\mathcal{F}}^a_{\alpha\beta} = \sum_{\{\gamma\eta\}=0}^3 \mathcal{F}^a_{\gamma\eta} \det [K]^{\{\gamma\eta\}}_{\{\alpha\beta\}}$

where $[K]$ is the inverse of the jacobian $[K]=[J]^{-1}, [J]=\left[\frac{\partial\eta^{\alpha}}{\partial\xi^{\beta}}\right]$

$\widetilde{\mathcal{F}}^{a}_{\alpha\beta}=\sum_{\{\gamma\eta\}=0}^{3}\mathcal{F}^{a}_{\gamma\eta}\det\begin{bmatrix}K^{\gamma}_{\alpha} & K^{\gamma}_{\beta}\\ K^{\eta}_{\alpha} & K^{\eta}_{\beta}\end{bmatrix}=\sum_{\{\gamma\eta\}=0}^{3}\mathcal{F}^{a}_{G\gamma\eta}\left(K^{\gamma}_{\alpha}K^{\eta}_{\beta}-K^{\gamma}_{\beta}K^{\eta}_{\alpha}\right)$

Let us denote :

$[K]=\begin{bmatrix}K^{0}_{0} & \left[K^{0}\right]\\ [K_0] & [k]_{3\times3}\end{bmatrix}$

and more generally : $[M]_{\lambda}$ the column $\lambda, [M]^{\mu}$ the row μ of the matrix $[M]$.

$\alpha,\beta=1,2,3$:

$\widetilde{\mathcal{F}}_{0\beta}=\left[\widetilde{\mathcal{F}}^{w}\right]_{\beta}=\left([K]^{0}_{0}\right)[\mathcal{F}^{w}][k]_{\beta}-\left[K^{0}\right]_{\beta}[\mathcal{F}^{w}][K_0]-[\mathcal{F}^{r}]\,j\left([K_0]\right)[k]_{\beta}$

$$\left[\widetilde{\mathcal{F}}^{w}\right]=[\mathcal{F}^{w}]\left(\left([K]^{0}_{0}\right)[k]-[K_0]\left[K^{0}\right]\right)-[\mathcal{F}^{r}]\,j\left([K_0]\right)[k] \qquad (5.43)$$

$\widetilde{\mathcal{F}}_{\alpha\beta}=[\mathcal{F}^{w}]\left(\left[K^{0}\right]_{\alpha}[k]_{\beta}-\left[K^{0}\right]_{\beta}[k]_{\alpha}\right)-[\mathcal{F}^{r}]\,j\left([k]_{\alpha}\right)[k]_{\beta}$

For *any matrix* :

$[M]=\begin{bmatrix}M_{11} & M_{12} & M_{13}\\ M_{21} & M_{22} & M_{23}\\ M_{31} & M_{32} & M_{33}\end{bmatrix}$

$[M]^{t}\begin{bmatrix}j\left([M]_2\right)[M]_3 & j\left([M]_3\right)[M]_1 & j\left([M]_1\right)[M]_2\end{bmatrix}=(\det M)\,I_3$

$\begin{bmatrix}j\left([M]_2\right)[M]_3 & j\left([M]_3\right)[M]_1 & j\left([M]_1\right)[M]_2\end{bmatrix}=(\det M)\left[M^{-1}\right]^{t}$

and :

$-[\mathcal{F}^{r}]\,j\left([k]_{\alpha}\right)[k]_{\beta}\rightarrow[\mathcal{F}^{r}](\det k)\left[k^{-1}\right]^{t}$

The other item reads :

$[\mathcal{F}^{w}]\left(\left[K^{0}\right]_{\alpha}[k]_{\beta}-\left[K^{0}\right]_{\beta}[k]_{\alpha}\right)$

$\rightarrow[\mathcal{F}^{w}]\begin{bmatrix}\left[K^{0}\right]_{3}[k]_{2}-\left[K^{0}\right]_{2}[k]_{3} & \left[K^{0}\right]_{1}[k]_{3}-\left[K^{0}\right]_{3}[k]_{1} & \left[K^{0}\right]_{2}[k]_{1}-\left[K^{0}\right]_{1}[k]_{2}\end{bmatrix}$

$=-\begin{matrix}\left[K^{0}\right]_{1} & \left[K^{0}\right]_{2} & \left[K^{0}\right]_{3}\end{matrix}\;j\begin{pmatrix}[\mathcal{F}^{w}][k]_{1}\\ [\mathcal{F}^{w}][k]_{2}\\ [\mathcal{F}^{w}][k]_{3}\end{pmatrix}$

$=-\left[K^{0}\right]j\left(\left([\mathcal{F}^{w}][k]\right)^{t}\right)=[\mathcal{F}^{w}][k]\,j\left(\left[K^{0}\right]\right)$

$$\left[\widetilde{\mathcal{F}}^{r}\right]=[\mathcal{F}^{r}](\det k)\left[k^{-1}\right]^{t}+[\mathcal{F}^{w}][k]\,j\left(\left[K^{0}\right]\right) \qquad (5.44)$$

Which can be summed up in the matrix, which holds also for each component $\mathcal{F}^{a}$:

$$\begin{bmatrix}\left[\widetilde{\mathcal{F}}^{r}\right] & \left[\widetilde{\mathcal{F}}^{w}\right]\end{bmatrix}=\begin{bmatrix}[\mathcal{F}^{r}] & [\mathcal{F}^{w}]\end{bmatrix}\begin{bmatrix}(\det k)\left[k^{-1}\right]^{t} & -j\left([K_0]\right)[k]\\ [k]\,j\left(\left[K^{0}\right]\right) & \left([K]^{0}_{0}\right)[k]-[K_0]\left[K^{0}\right]\end{bmatrix} \qquad (5.45)$$

For a change of spatial chart, with the same time axis, the value of each component $\mathcal{F}^{r},\mathcal{F}^{w}$ changes, but the split holds :

$\begin{bmatrix}\left[\widetilde{\mathcal{F}}^{r}\right] & \left[\widetilde{\mathcal{F}}^{w}\right]\end{bmatrix}=\begin{bmatrix}[\mathcal{F}^{r}] & [\mathcal{F}^{w}]\end{bmatrix}\begin{bmatrix}(\det k)\left[k^{-1}\right]^{t} & 0\\ 0 & [k]\end{bmatrix}$

Computation of the scalar curvature

With this notation it is easy to give a simple formula for the scalar curvature, with any connection (so also for the usual Levy-Civita connection) :

$\mathbf{R}=\sum_{a=1}^{3}\sum_{\alpha\beta}\mathcal{F}_{r\alpha\beta}^{a}\left([P][\kappa_a][\eta][P]^t\right)_{\alpha}^{\beta}+\mathcal{F}_{w\alpha\beta}^{a}\left([P][\kappa_{a+3}][\eta][P]^t\right)_{\alpha}^{\beta}$

$=\sum_{a=1}^{3}$

$\mathcal{F}_{r21}^{a}\left(\left([P][\kappa_a][\eta][P]^t\right)_2^1-\left([P][\kappa_a][\eta][P]^t\right)_1^2\right)+\mathcal{F}_{w21}^{a}\left(\left([P][\kappa_{a+3}][\eta][P]^t\right)_2^1-\left([P][\kappa_{a+3}][\eta][P]^t\right)_1^2\right)$

$+\mathcal{F}_{r13}^{a}\left(\left([P][\kappa_a][\eta][P]^t\right)_1^3-\left([P][\kappa_a][\eta][P]^t\right)_3^1\right)+\mathcal{F}_{w13}^{a}\left(\left([P][\kappa_{a+3}][\eta][P]^t\right)_1^3-\left([P][\kappa_{a+3}][\eta][P]^t\right)_3^1\right)$

$+\mathcal{F}_{r32}^{a}\left(\left([P][\kappa_a][\eta][P]^t\right)_3^2-\left([P][\kappa_a][\eta][P]^t\right)_2^3\right)+\mathcal{F}_{w32}^{a}\left(\left([P][\kappa_{a+3}][\eta][P]^t\right)_3^2-\left([P][\kappa_{a+3}][\eta][P]^t\right)_2^3\right)$

Using :

$\left([P][\kappa_a][\eta][P]^t\right)-\left([P][\kappa_a][\eta][P]^t\right)^t=[P][\kappa_a][\eta][P]^t-[P][\eta][\kappa_a]^t[P]^t=[P]\left([\kappa_a][\eta]-[\eta][\kappa_a]^t\right)[P]^t$

$a=1,2,3:[\kappa_a][\eta]-[\eta][\kappa_a]^t=2[\kappa_a]$

$a=4,5,6:[\kappa_a][\eta]-[\eta][\kappa_a]^t=2[\kappa_a][\eta]$

$\mathbf{R}=2\sum_{a=1}^{3}\mathcal{F}_{r21}^{a}\left([P][\kappa_a][P]^t\right)_2^1+\mathcal{F}_{w21}^{a}\left([P][\kappa_{a+3}][\eta][P]^t\right)_2^1$

$+\mathcal{F}_{r13}^{a}\left([P][\kappa_a][P]^t\right)_1^3+\mathcal{F}_{w13}^{a}\left([P][\kappa_{a+3}][\eta][P]^t\right)_1^3$

$+\mathcal{F}_{r32}^{a}\left([P][\kappa_a][P]^t\right)_3^2+\mathcal{F}_{w32}^{a}\left([P][\kappa_{a+3}][\eta][P]^t\right)_3^2$

$\mathbf{R}=2\{[\mathcal{F}_r^r]_3^1[j([Q]_3)[Q]_2]^3+[\mathcal{F}_r^r]_3^2[j([Q]_1)[Q]_3]^3+[\mathcal{F}_r^r]_3^3[j([Q]_2)[Q]_1]^3$

$+[\mathcal{F}_r^r]_2^1[j([Q]_3)[Q]_2]^2+[\mathcal{F}_r^r]_2^2[j([Q]_1)[Q]_3]^2+[\mathcal{F}_r^r]_2^3[j([Q]_2)[Q]_1]^2$

$+[\mathcal{F}_r^r]_1^1[j([Q]_3)[Q]_2]^1+[\mathcal{F}_r^r]_1^2[j([Q]_1)[Q]_3]^1+[\mathcal{F}_r^r]_1^3[j([Q]_2)[Q]_1]^1$

$+[\mathcal{F}_w^r]_3^1\left([P]_0^1[P]_1^2-[P]_1^1[P]_0^2\right)+[\mathcal{F}_w^r]_3^2\left([P]_0^1[P]_2^2-[P]_0^2[P]_2^1\right)+[\mathcal{F}_w^r]_2^3\left([P]_0^1[P]_3^2-[P]_0^2[P]_3^1\right)$

$-[\mathcal{F}_w^r]_2^1\left([P]_0^1[P]_1^3-[P]_1^1[P]_0^3\right)-[\mathcal{F}_w^r]_2^2\left([P]_0^1[P]_2^3-[P]_0^3[P]_2^1\right)-[\mathcal{F}_w^r]_2^3\left([P]_0^1[P]_3^3-[P]_0^3[P]_3^1\right)$

$+[\mathcal{F}_w^r]_1^1\left([P]_1^3[P]_0^2-[P]_0^3[P]_1^2\right)+[\mathcal{F}_w^r]_1^2\left([P]_0^2[P]_2^3-[P]_0^3[P]_2^2\right)+[\mathcal{F}_w^r]_1^3\left([P]_0^2[P]_3^3-[P]_0^3[P]_3^2\right)\}$

Let us denote the 3×3 matrix :

$[J]=\left[\begin{array}{ccc} j([Q]_3)[Q]_2 & j([Q]_1)[Q]_3 & j([Q]_2)[Q]_1\end{array}\right]=-(\det Q)\left[Q^{-1}\right]^t$

$[\mathcal{F}_r^r]_3^1[j([Q]_3)[Q]_2]^3+[\mathcal{F}_r^r]_3^2[j([Q]_1)[Q]_3]^3+[\mathcal{F}_r^r]_3^3[j([Q]_2)[Q]_1]^3$

$+[\mathcal{F}_r^r]_2^1[j([Q]_3)[Q]_2]^2+[\mathcal{F}_r^r]_2^2[j([Q]_1)[Q]_3]^2+[\mathcal{F}_r^r]_2^3[j([Q]_2)[Q]_1]^2$

$+[\mathcal{F}_r^r]_1^1[j([Q]_3)[Q]_2]^1+[\mathcal{F}_r^r]_1^2[j([Q]_1)[Q]_3]^1+[\mathcal{F}_r^r]_1^3[j([Q]_2)[Q]_1]^1$

$=[\mathcal{F}_r^r]_3^1[J]_1^3+[\mathcal{F}_r^r]_3^2[J]_2^3+[\mathcal{F}_r^r]_3^3[J]_3^3+[\mathcal{F}_r^r]_2^1[J]_1^2+[\mathcal{F}_r^r]_2^2[J]_2^2+[\mathcal{F}_r^r]_2^3[J]_3^2+[\mathcal{F}_r^r]_1^1[J]_1^1+[\mathcal{F}_r^r]_1^2[J]_2^1+[\mathcal{F}_r^r]_1^3[J]_3^1$

$=Tr([\mathcal{F}_r^r][J])=-(\det Q)\,Tr\left([\mathcal{F}_r^r]\left[Q^{-1}\right]^t\right)$

$[\mathcal{F}_w^r]_3^1\left([P]_0^1[P]_1^2-[P]_1^1[P]_0^2\right)+[\mathcal{F}_w^r]_3^2\left([P]_0^1[P]_2^2-[P]_0^2[P]_2^1\right)+[\mathcal{F}_w^r]_2^3\left([P]_0^1[P]_3^2-[P]_0^2[P]_3^1\right)$

$-[\mathcal{F}_w^r]_2^1\left([P]_0^1[P]_1^3-[P]_1^1[P]_0^3\right)-[\mathcal{F}_w^r]_2^2\left([P]_0^1[P]_2^3-[P]_0^3[P]_2^1\right)-[\mathcal{F}_w^r]_2^3\left([P]_0^1[P]_3^3-[P]_0^3[P]_3^1\right)$

$+[\mathcal{F}_w^r]_1^1\left([P]_1^3[P]_0^2-[P]_0^3[P]_1^2\right)+[\mathcal{F}_w^r]_1^2\left([P]_0^2[P]_2^3-[P]_0^3[P]_2^2\right)+[\mathcal{F}_w^r]_1^3\left([P]_0^2[P]_3^3-[P]_0^3[P]_3^2\right)$

$=[P]_0^1([Q][\mathcal{F}_w^r])_3^2-[P]_0^2([Q][\mathcal{F}_w^r])_3^1-[P]_0^1([Q][\mathcal{F}_w^r])_2^3+[P]_0^3([P][\mathcal{F}_w^r])_2^1+[P]_0^2([Q][\mathcal{F}_w^r])_1^3-[P]_0^3([P][\mathcal{F}_w^r])_1^2$

$=[j([P_0])([Q][\mathcal{F}_w^r])_3]^3+[j([P_0])([Q][\mathcal{F}_w^r])_2]^2+[j([P_0])([Q][\mathcal{F}_w^r])_1]^1$

$=Tr(j([P_0])[Q][\mathcal{F}_w^r])=Tr([\mathcal{F}_w^r]j([P_0])[Q])$

$$\mathbf{R} = 2Tr\left(-\left(\det Q\right)\left[\mathcal{F}_r^r\right]\left[Q'\right]^t + \left[\mathcal{F}_w^r\right] j\left(\left[P_0\right]\right)\left[Q\right]\right) \tag{5.46}$$

We see that, $\left[\mathcal{F}_r^w\right], \left[\mathcal{F}_w^w\right]$ are not involved and with a standard basis :

$$\mathbf{R} = -2\left(\det Q\right) Tr\left(\left[\mathcal{F}_r^r\right]\left[Q'\right]^t\right) \tag{5.47}$$

only $\left[\mathcal{F}_r^r\right]$ is involved, which reduces significantly the interest of the scalar curvature to account for $\mathcal{F}_G$. Anyway the computation is far easier with the tetrad. One can do the same computation for the Ricci tensor. Then with the Einstein equation $\mathcal{F}_G$ is, with any connection, solution of a set of linear equations with parameter $[P]$.

But we give now the full justification of this notation.

Scalar product of forms

On any n dimensional manifold endowed with a non degenerate metric g there is a scalar product, denoted G_r for r-forms $\lambda \in \Lambda_r(M;\mathbb{R})$ (Maths.1611). G_r is a bilinear symmetric form, which does not depend on a chart, is non degenerate and definite positive if g is Riemannian.

$G_r(\lambda,\mu) = \sum_{\{\alpha_1...\alpha_r\}\{\beta_1...\beta_r\}} \lambda_{\alpha_1...\alpha_r}\mu_{\beta_1...\beta_r} \det[g]^{\{\alpha_1...\alpha_r\}}_{\{\beta_1...\beta_r\}}$

Hodge dual

G_r defines an isomorphism between r and n-r forms. The **Hodge dual** $*\lambda$ of a r form λ is a $n-r$ form (Maths.1613) such that :

$\forall \mu \in \Lambda_{n-r}(M) : *\lambda \wedge \mu = G_r(\lambda,\mu)\varpi_n$ where ϖ_n is the volume form deduced from the metric. For 2 forms on M :

$$\forall \lambda,\mu \in \Lambda_2(M;\mathbb{R}) : *\lambda \wedge \mu = G_2(\lambda,\mu)\varpi_4 \tag{5.48}$$

The Hodge dual $*\mathcal{F}$ of a scalar 2-form $\mathcal{F} \in \Lambda_2(M,\mathbb{R})$ is a 2 form whose expression, with the Lorentz metric, is simple when a specific ordering is used.

$\mathcal{F} = \mathcal{F}^r + \mathcal{F}^w; K = K^r + K^w$

$K \wedge \mathcal{F} \det P' = -\left(\left[K^r\right]\left[\mathcal{F}^w\right]^t + \left[K^w\right]\left[\mathcal{F}^r\right]^t\right)\varpi_4$

Writing $\mathcal{F} = \mathcal{F}^r + \mathcal{F}^w$ then :

$*\mathcal{F} = *\mathcal{F}^r + *\mathcal{F}^w$

$$\left[\begin{array}{l} *\mathcal{F}^r = -\left(\mathcal{F}^{01}d\xi^3\wedge d\xi^2 + \mathcal{F}^{02}d\xi^1\wedge d\xi^3 + \mathcal{F}^{03}d\xi^2\wedge d\xi^1\right)\det P' \\ *\mathcal{F}^w = -\left(\mathcal{F}^{32}d\xi^0\wedge d\xi^1 + \mathcal{F}^{13}d\xi^0\wedge d\xi^2 + \mathcal{F}^{21}d\xi^0\wedge d\xi^3\right)\det P' \\ \mathcal{F}^{\alpha\beta} = \sum_{\lambda\mu=0}^3 g^{\alpha\lambda}g^{\beta\mu}\mathcal{F}_{\lambda\mu} \end{array}\right] \tag{5.49}$$

The components of the parts are exchanged and the indices are raised with the metric g .Notice that the Hodge dual is a 2 form : even if the notation uses raised indexes, they refer to the basis $d\xi^\alpha \wedge d\xi^\beta$.

The computation of $\mathcal{F}^{\alpha\beta}$ is then easier with the notations :

$[*\mathcal{F}^r] = -\left[\begin{array}{ccc} \mathcal{F}^{01} & \mathcal{F}^{02} & \mathcal{F}^{03} \end{array}\right](\det P');$

$[*\mathcal{F}^w] = -\left[\begin{array}{ccc} \mathcal{F}^{32} & \mathcal{F}^{13} & \mathcal{F}^{21} \end{array}\right](\det P')$

$[g] = \begin{bmatrix} g_{00} & [g^0] \\ [g_0] & [g_3] \end{bmatrix}; [g^{-1}] = \begin{bmatrix} g^{00} & [g^{-1}]^0 \\ [g^{-1}]_0 & [g_3^{-1}] \end{bmatrix}$

$[g^{-1}]_0 = \begin{bmatrix} g^{10} \\ g^{20} \\ g^{30} \end{bmatrix} = \begin{bmatrix} -P_{00}P_{10} + P_{01}P_{11} + P_{02}P_{12} + P_{03}P_{13} \\ -P_{00}P_{20} + P_{01}P_{21} + P_{02}P_{22} + P_{03}P_{23} \\ -P_{00}P_{30} + P_{01}P_{31} + P_{02}P_{32} + P_{03}P_{33} \end{bmatrix}$

Using :

$\begin{bmatrix} j([g_3^{-1}]_2)[g_3^{-1}]_3 & j([g_3^{-1}]_3)[g_3^{-1}]_1 & j([g_3^{-1}]_1)[g_3^{-1}]_2 \end{bmatrix}$

$= (\det [g_3^{-1}])[g_3]^t = (\det [g_3^{-1}])[g_3] = (\det Q)^2 [g_3]$

$$[*\mathcal{F}^r] = \left([\mathcal{F}^w]\left(-g^{00}[g_3^{-1}] + [g_3^{-1}]_0 [g_3^{-1}]^0\right) + [\mathcal{F}^r] j([g_3^{-1}]_0)[g_3^{-1}]\right) \det P' \quad (5.50)$$

$$[*\mathcal{F}^w] = -\left([\mathcal{F}^w][g_3^{-1}] j([g^{-1}]_0) + (\det Q)^2 [\mathcal{F}^r][g_3]\right) \det P' \quad (5.51)$$

For each component $(\mathcal{F}_G^a)_{a=1..6}, (\mathcal{F}_A^a)_{a=1..m}$ we can compute the Hodge dual (with respect to g) and we get :

$[*\mathcal{F}_r^r] = \left([\mathcal{F}_r^w]\left(-g^{00}[g_3^{-1}] + [g^{-1}]_0 [g^{-1}]^0\right) + [\mathcal{F}_r^r] j([g^{-1}]_0)[g_3^{-1}]\right) \det P'$

$[*\mathcal{F}_w^r] = \left([\mathcal{F}_w^w]\left(-g^{00}[g_3^{-1}] + [g^{-1}]_0 [g^{-1}]^0\right) + [\mathcal{F}_w^r] j([g^{-1}]_0)[g_3^{-1}]\right) \det P'$

$[*\mathcal{F}_r^w] = -\left([\mathcal{F}_r^w][g_3^{-1}] j([g^{-1}]_0) + (\det Q)^2 [\mathcal{F}_r^r][g_3]\right) \det P'$

$[*\mathcal{F}_w^w] = -\left([\mathcal{F}_w^w][g_3^{-1}] j([g^{-1}]_0) + (\det Q)^2 [\mathcal{F}_w^r][g_3]\right) \det P'$

$[*\mathcal{F}_A^r] = \left([\mathcal{F}_A^w]\left(-g^{00}[g_3^{-1}] + [g^{-1}]_0 [g^{-1}]^0\right) + [\mathcal{F}_a^r] j([g^{-1}]_0)[g_3^{-1}]\right) \det P'$

$[*\mathcal{F}_A^w] = -\left([\mathcal{F}^w][g_3^{-1}] j([g^{-1}]_0) + (\det Q)^2 [\mathcal{F}^r][g_3]\right) \det P'$

It is quite simpler with a standard chart : $[h] = 0, g^{00} = -1, \det P' = \det Q'$

$$\begin{bmatrix} [*\mathcal{F}_r^r] = [\mathcal{F}_r^w][g_3^{-1}] \det Q' & [*\mathcal{F}_w^r] = [\mathcal{F}_w^w][g_3^{-1}] \det Q' \\ [*\mathcal{F}_r^w] = -[\mathcal{F}_r^r][g_3] \det Q & [*\mathcal{F}_w^w] = -[\mathcal{F}_w^r][g_3] \det Q \\ [*\mathcal{F}_A^r] = [\mathcal{F}_A^w][g_3^{-1}] \det Q' & [*\mathcal{F}_A^w] = -[\mathcal{F}^r][g_3] \det Q \end{bmatrix} \quad (5.52)$$

Notice that :

$\begin{bmatrix} \mathcal{F}^{32} & \mathcal{F}^{13} & \mathcal{F}^{21} \end{bmatrix} = -[*\mathcal{F}^w](\det P)$

$\begin{bmatrix} \mathcal{F}^{01} & \mathcal{F}^{02} & \mathcal{F}^{03} \end{bmatrix} = -[*\mathcal{F}^r](\det P)$

The converse equations are easily computed with :

$**\lambda_r = -(-1)^{r(n-r)}\lambda \Rightarrow **\lambda_2 = -\lambda_2$

Electromagnetic field

The strength of the electromagnetic field is a 2 form valued in $\mathbb{R} : \mathcal{F}_{EM} \in \Lambda_2 (M; \mathbb{R})$.

In electrodynamics the electric field $\overrightarrow{E}$ and the magnetic field $\overrightarrow{H}$ are represented by vectors $[E], [H]$ of $\mathbb{R}^3$. .

To $\overrightarrow{E}, \overrightarrow{H}$ one can associate one forms in the tangent space to $\Omega(t)$:

$E^* = \sum_{\alpha,\beta=1}^3 g_{\alpha\beta} E^\beta d\xi^\alpha$

$H^* = \sum_{\alpha,\beta=1}^3 g_{\alpha\beta} H^\beta d\xi^\alpha$

$d\xi^0 \wedge E^* = \sum_{\alpha,\beta=1}^3 g_{\alpha\beta} E^\beta d\xi^0 \wedge d\xi^\alpha = \mathcal{F}_{EM}^w$

$[\mathcal{F}_{EM}^w] = \left[\sum_{\beta=1}^3 g_{\alpha\beta} E^\beta\right] = [E]^t [g_3]$

In $T\Omega(t)$ one can compute the Hodge dual of H^* which is a 3-1=2 form :

$*H^* = (\det Q') \sum_{\alpha,\beta=1}^3 (-1)^{\alpha+1} g^{\alpha\beta} (H^*)_\beta d\xi^1 \wedge ..\widehat{d\xi^\alpha}..d\xi^3$

$= (\det Q') \sum_{\beta=1}^3 g^{1\beta} (H^*)_\beta d\xi^2 \wedge d\xi^3 - g^{2\beta} (H^*)_\beta d\xi^1 \wedge d\xi^3 + g^{3\beta} (H^*)_\beta d\xi^1 \wedge d\xi^2$

$= (\det Q') \left(H^1 d\xi^2 \wedge d\xi^3 - H^2 d\xi^1 \wedge d\xi^3 + H^3 d\xi^1 \wedge d\xi^2\right)$

$= (\det Q') \sum_{\gamma=1}^3 -H^1 d\xi^3 \wedge d\xi^2 - H^2 d\xi^1 \wedge d\xi^3 - H^3 d\xi^2 \wedge d\xi^1$

$= -\mathcal{F}^r$

$[\mathcal{F}_{EM}^r] = -[H]^t \det Q'$

And :

$$\mathcal{F}_{EM} = d\xi^0 \wedge E^* - *H^* \det Q \tag{5.53}$$

$\mathcal{F}_{EM} = \sum_{\alpha,\beta=1}^3 g_{\alpha\beta} E^\beta d\xi^0 \wedge d\xi^\alpha + H^1 d\xi^3 \wedge d\xi^2 + H^2 d\xi^1 \wedge d\xi^3 + H^3 d\xi^2 \wedge d\xi^1$

$[*\mathcal{F}_{EM}^r] = [\mathcal{F}^w] \left[g_3^{-1}\right] \det Q' = [E]^t [g_3] \left[g_3^{-1}\right] \det Q' = [E]^t \det Q'$

$[*\mathcal{F}_{EM}^w] = -[\mathcal{F}^r] [g_3] \det Q = [H]^t [g_3]$

Scalar product of two forms on M

The scalar product of forms is then easy to compute with the Hodge dual. Take any two scalar 2 forms $\mathcal{F}, K$ and their decomposition as above, a straightforward computation gives :

$*\mathcal{F}^w \wedge K^w = 0$

$*\mathcal{F}^w \wedge K^r = \left(\mathcal{F}^{32} K_{32} + \mathcal{F}^{13} K_{13} + \mathcal{F}^{21} K_{21}\right) \varpi_4$

$\mathcal{F}^w \wedge K^r = \left([\mathcal{F}^r] [K^w]^t\right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$*\mathcal{F}^r \wedge K^w = \left(\mathcal{F}^{01} K_{01} + \mathcal{F}^{02} K_{02} + \mathcal{F}^{03} K_{03}\right) \varpi_4$

$*\mathcal{F}^r \wedge K^r = 0$

$G_2 (\mathcal{F}^w, K^w) = G_2 (\mathcal{F}^r, K^r) = 0$

$G_2 (\mathcal{F}^w, K^r) = \left(\mathcal{F}^{32} K_{32} + \mathcal{F}^{13} K_{13} + \mathcal{F}^{21} K_{21}\right)$

$= \begin{bmatrix} \mathcal{F}^{32} & \mathcal{F}^{13} & \mathcal{F}^{21} \end{bmatrix} \begin{bmatrix} K_{32} \\ K_{13} \\ K_{21} \end{bmatrix}$

$= -\frac{1}{\det P'} [*\mathcal{F}^w] [K^r]^t$

$= \left([\mathcal{F}^w] \left[g_3^{-1}\right] j ([g]_0) + (\det Q)^2 [\mathcal{F}^r] [g_3]\right) [K^r]^t$

$G_2 (\mathcal{F}^r, K^w) = \left(\mathcal{F}^{01} K_{01} + \mathcal{F}^{02} K_{02} + \mathcal{F}^{03} K_{03}\right)$

$= \begin{bmatrix} \mathcal{F}^{01} & \mathcal{F}^{02} & \mathcal{F}^{03} \end{bmatrix} \begin{bmatrix} K_{01} \\ K_{02} \\ K_{03} \end{bmatrix}$

$= -\frac{1}{\det P'} [*\mathcal{F}^r] [K^w]^t$

$= -\left([\mathcal{F}^w] \left(-g^{00} \left[g_3^{-1} \right] + \left[g^{-1} \right]_0 \left[g^{-1} \right]^0 \right) + [\mathcal{F}^r] j \left(\left[g^{-1} \right]_0 \right) \left[g_3^{-1} \right] \right) [K^w]^t$
From there, because G_2 is bilinear :
$G_2 (\mathcal{F}, K)$
$= G_2 (\mathcal{F}^r + \mathcal{F}^w, K^r + K^w)$
$= G_2 (\mathcal{F}^r, K^w) + G_2 (\mathcal{F}^w, K^r)$
$= \left(\mathcal{F}^{32} K_{32} + \mathcal{F}^{13} K_{13} + \mathcal{F}^{21} K_{21} + \mathcal{F}^{01} K_{01} + \mathcal{F}^{02} K_{02} + \mathcal{F}^{03} K_{03} \right)$
$= \sum_{\{\alpha\beta\}} \mathcal{F}^{\alpha\beta} K_{\alpha\beta}$
$= \left([\mathcal{F}^w] \left[g_3^{-1} \right] j \left(\left[g_3^{-1} \right]_0 \right) + (\det Q)^2 [\mathcal{F}^r] [g_3] \right) [K^r]^t$
$- \left([\mathcal{F}^w] \left(-g^{00} \left[g_3^{-1} \right] + \left[g_3^{-1} \right]_0 \left[g_3^{-1} \right]^0 \right) + [\mathcal{F}^r] j \left(\left[g_3^{-1} \right]_0 \right) \left[g_3^{-1} \right] \right) [K^w]^t$
$= (\det Q)^2 [\mathcal{F}^r] [g_3] [K^r]^t - [\mathcal{F}^r] j \left(\left[g_3^{-1} \right]_0 \right) \left[g_3^{-1} \right] [K^w]^t$
$+ [\mathcal{F}^w] \left[g_3^{-1} \right] j ([g]_0) [K^r]^t - [\mathcal{F}^w] \left(-g^{00} \left[g_3^{-1} \right] + \left[g_3^{-1} \right]_0 \left[g_3^{-1} \right]^0 \right) [K^w]^t$
and in the standard chart :
$G_2 (\mathcal{F}, K) = (\det Q)^2 [\mathcal{F}^r] [g_3] [K^r]^t - [\mathcal{F}^w] \left[g_3^{-1} \right] [K^w]^t$

$$G_2 (\mathcal{F}, K) = -\frac{1}{\det P'} \left([*\mathcal{F}^w] [K^r]^t + [*\mathcal{F}^r] [K^w]^t \right) = \sum_{\{\alpha\beta\}} \mathcal{F}^{\alpha\beta} K_{\alpha\beta} \tag{5.54}$$

For the EM field :
$G_2 (\mathcal{F}_{EM}, \mathcal{F}_{EM}) = -\frac{1}{\det P'} \left([H]^t [g_3] \left(- [H]^t \det Q' \right)^t + [E]^t \det Q' \left[[E]^t [g_3] \right]^t \right)$
$= -\frac{1}{\det P'} \left(- [H]^t [g_3] [H] \det Q' + [E]^t [g_3] [E] \det Q' \right) = \|H\|_3^2 - \|E\|_3^2$

Norm on $\Lambda_2 (M; \mathbb{R})$

On one hand the decomposition $\mathcal{F}^r, \mathcal{F}^w$ can be done by choosing any vector $\partial\xi_0$: it does not involve any specific property of $\partial\xi_0$.

For any vector field $X \in \mathfrak{X}(TM)$ and its associate form $X^* \in \mathfrak{X}(TM^*) : X_\alpha = \sum_\beta g_{\alpha\beta} X^\beta$ let us define the maps :
$\lambda (X^*) : \Lambda_r (M; \mathbb{R}) \to \Lambda_{r+1} (M; \mathbb{R}) :: \lambda (X^*) \mu = \lambda (X^*) \wedge \mu$
$i_X : \Lambda_r (M; \mathbb{R}) \to \Lambda_{r-1} (M; \mathbb{R}) :: i_X \mu = \mu (X)$
$\{\mathcal{F}^w\} = \ker \lambda \left(d\xi^0 \right)$
$\{\mathcal{F}^r\} = \ker i_{\partial\xi_0}$
$\{\mathcal{F}^w\} \cap \{\mathcal{F}^r\} = 0$
The choice of $\partial\xi_0$ gives a decomposition : $\Lambda_2 (M; \mathbb{R}) = \{\mathcal{F}^w\} \oplus \{\mathcal{F}^r\}$

On the other hand G_r relies on the existence of the metric g, for which $\partial\xi_0$ is time-like, thus the expression of the scalar product as well as of the Hodge dual assumes that $\partial\xi_0$ is time-like.

The bilinear form G_2 is symmetric, non degenerate, but not definite positive. There is at each point m a partition of $\Lambda_2 (M; \mathbb{R})$ into 3 components :
$C^+ = \{\mathcal{F} \in \Lambda_2 (M; \mathbb{R}), G_2 (\mathcal{F}, \mathcal{F}) > 0\}$
$C^- = \{\mathcal{F} \in \Lambda_2 (M; \mathbb{R}), G_2 (\mathcal{F}, \mathcal{F}) < 0\}$
$C_0 = \{\mathcal{F} \in \Lambda_2 (M; \mathbb{R}), G_2 (\mathcal{F}, \mathcal{F}) = 0\}$
$\Lambda_2 (M; \mathbb{R})$ is a vector space, whose topology is inherited from M : continuous tensor fields are defined by continuous maps through charts. It is infinite dimensional. C^+, C^- are open sets, C_0 is closed in $\Lambda_2 (M; \mathbb{R})$, and they are cone with apex 0.

The value of $G_2(\mathcal{F},\mathcal{F})$ does not depend on the chart, thus we can always choose a standard chart, for which : $[g_3]$ is definite positive, $g^{00}=-1, \left[g_3^{-1}\right]_0=\left[g_3^{-1}\right]^0=0$. And we see that $\mathcal{F}^r\in C^+, \mathcal{F}^w\in C^-$. And conversely, as seen in the standard chart.

$C^+=\{\mathcal{F}\in\Lambda_2(M;\mathbb{R}):\mathcal{F}=\mathcal{F}^r\}$

$C^-=\{\mathcal{F}\in\Lambda_2(M;\mathbb{R}):\mathcal{F}=\mathcal{F}^w\}$

C^+ is the set of 2 forms which, in any chart such that $\partial\xi_0$ is time-like, can be expressed as $\mathcal{F}^r$. And similarly for C^-.

The metric is part of the physical structure of the universe, it implies a specific structure for vectors fields and also for the 2 forms.

In any physical chart there is necessarily one vector field of the holonomic basis which is time-like. It defines a decomposition of any 2 form in $\mathcal{F}^r,\mathcal{F}^w$, and we can compute $G_2(\mathcal{F}^r,\mathcal{F}^r)>0, G_2(\mathcal{F}^w,\mathcal{F}^w)<0$ and :

$\|\mathcal{F}\|^2=G_2(\mathcal{F}^r,\mathcal{F}^r)-G_2(\mathcal{F}^w,\mathcal{F}^w)>0$

$G_2(\mathcal{F}^w,\mathcal{F}^w), G_2(\mathcal{F}^r,\mathcal{F}^r)$ do not depend on the chart, so $\|\mathcal{F}\|^2$ does not depend on the chart, and it is easy to check that it defines a norm on $\Lambda_2(M;\mathbb{R})$.

It can be expressed as :

$\|\mathcal{F}\|^2=[\mathcal{F}^w]\left(-g^{00}\left[g_3^{-1}\right]+\left[g_3^{-1}\right]_0\left[g_3^{-1}\right]^0\right)[\mathcal{F}^w]^t+(\det Q)^2[\mathcal{F}^r][g_3][\mathcal{F}^r]^t>0$

Notice that :

i) The value of the norm for a 2 form depends on the choice of $\partial\xi_0$, but there is always a norm.

ii) This norm is defined point-wise : $\|\mathcal{F}(m)\|^2$. But with the volume form ϖ_4 we have a norm on any relatively compact subset $\Omega\subset M$: $\|\mathcal{F}\|^2=\int_\Omega\|\mathcal{F}(m)\|^2\varpi_4(m)$.

iii) The scalar curvature involves only $[\mathcal{F}_r^r],[\mathcal{F}_w^r]$, and this raises a serious issue because it seems clear that the part of $\mathcal{F}_G$ which lies in the cone C^- should play a specific role.

To extend this scalar product on 2 forms valued in the Lie algebras we need first a scalar product on these vector spaces.

Scalar products on the Lie algebras

The strength can be seen as a section of the associated vector bundles $P_G[T_1Spin(3,1),\mathbf{Ad}]$,

$P_U[T_1U,Ad]$ and then the scalar product must be preserved by the adjoint map Ad. There are not too many possibilities. It can be shown that, for simple groups of matrices, the only scalar products on their Lie algebra which are invariant by the adjoint map are of the kind : $\langle[X],[Y]\rangle=kTr\left([X]^*[Y]\right)$ which sums up, in our case, to use the Killing form. This is a bilinear form (Maths.1609) which is preserved by any automorphism of the Lie algebra (thus in any representation). However it is negative definite if and only if the group is compact and semi-simple (Maths.1847).

Scalar product for the gravitational field

The scalar product on $T_1Spin(3,1)$, induced by the scalar product on the Clifford algebra, is, up to a constant, the Killing form :

$\langle\upsilon(r,w),\upsilon(r',w')\rangle_{Cl(3,1)}=\frac{1}{4}(r^tr'-w^tw')$

a= 1,2,3 : $\mathcal{F}^a_{G\alpha\beta}=\mathcal{F}^a_{r\alpha\beta}$

a = 4,5,6 : $\mathcal{F}^a_{G\alpha\beta}=\mathcal{F}^a_{w\alpha\beta}$

For fixed indices α,β,λ,μ :

$\left\langle\mathcal{F}_{G\alpha\beta}(m),\mathcal{F}'_{G\lambda\mu}(m)\right\rangle_{Cl}=\left\langle\upsilon(\mathcal{F}_{r\alpha\beta},\mathcal{F}_{w\alpha\beta}),\upsilon\left(\mathcal{F}'_{r\lambda\mu},\mathcal{F}'_{w\lambda\mu}\right)\right\rangle_{Cl}$

$= \frac{1}{4}\left(\mathcal{F}^t_{r\alpha\beta}\mathcal{F}'_{r\lambda\mu} - \mathcal{F}^t_{w\alpha\beta}\mathcal{F}'_{w\lambda\mu}\right) = \frac{1}{4}\left(\sum_{a=1}^{3}\mathcal{F}^a_{G\alpha\beta}\mathcal{F}'^a_{G\lambda\mu} - \sum_{a=4}^{6}\mathcal{F}^a_{G\alpha\beta}\mathcal{F}'^a_{G\lambda\mu}\right)$

The result does not depend on the signature. This scalar product is invariant in a change of gauge, non degenerate but not definite positive.

Scalar product for the other fields

The group U is assumed to be compact and connected. If U is semi-simple, its Killing form, which is invariant by the adjoint map, is then definite negative, and we can define a definite positive scalar product, invariant in a change of gauge, on its Lie algebra. This is the case for $SU(2)$ and $SU(3)$ but not for $U(1)$, however the Lie algebra of $U(1)$ is $\mathbb{R}$ and there is an obvious definite positive scalar product. As T_1U is a real vector space the scalar product is a *bilinear symmetric* form.

So we will assume that :

Proposition 104 *There is a definite positive scalar product on the Lie algebra T_1U, defined by a bilinear symmetric form preserved by the adjoint map, that we will denote $\langle\rangle_{T_1U}$. The basis $\left(\overrightarrow{\theta}_a\right)_{a=1}^m$ of T_1U is orthonormal for this scalar product.*

Notice that it is different from the scalar product on F (which defines the charges), which is Hermitian. In the standard model, because several groups are involved, three different constants are used, called the "gauge coupling". Here we consider only one group, and we can take the basis $\left(\overrightarrow{\theta}_a\right)_{a=1}^m$ as orthonormal for the scalar product.

The scalar product between sections $\mathcal{F}_A$ of $\Lambda_2(M;T_1U)$ is then defined, pointwise, as

$$\left\langle \mathcal{F}_{A\alpha\beta}(m), \mathcal{F}'_{A\lambda\mu}(m)\right\rangle_{T_1U} = \sum_{a=1}^{m}\mathcal{F}^a_{A\alpha\beta}(m)\,\mathcal{F}'^a_{A\lambda\mu}(m) \tag{5.55}$$

Scalar product for the strength of the fields

We have to combine both scalar products. They can all be expressed with $\mathcal{F}^{ar}, \mathcal{F}^{aw}$.

For the gravitational field

$\langle \mathcal{F}, K\rangle_G = \left\langle \sum_{a=1}^{6}\sum_{\{\alpha,\beta\}=0}^{3}\mathcal{F}^a_{\alpha\beta}d\xi^\alpha\wedge d\xi^\beta\otimes\overrightarrow{\kappa}_a, \sum_{b=1}^{6}\sum_{\{\alpha,\beta\}=0}^{3}K^b_{\alpha\beta}d\xi^\alpha\wedge d\xi^\beta\otimes\overrightarrow{\kappa}_b\right\rangle$

$= \sum_{a,b=1}^{6}\left\langle\sum_{\{\alpha,\beta\}=0}^{3}\mathcal{F}^a_{\alpha\beta}d\xi^\alpha\wedge d\xi^\beta, \sum_{\{\alpha,\beta\}=0}^{3}K^b_{\alpha\beta}d\xi^\alpha\wedge d\xi^\beta\right\rangle_{TM}\langle\overrightarrow{\kappa}_a,\overrightarrow{\kappa}_b\rangle_{Cl}$

$= \sum_{a=1}^{3}\left\langle\sum_{\{\alpha,\beta\}=0}^{3}\mathcal{F}^a_{r\alpha\beta}d\xi^\alpha\wedge d\xi^\beta, \sum_{\{\alpha,\beta\}=0}^{3}K^a_{r\alpha\beta}d\xi^\alpha\wedge d\xi^\beta\right\rangle_{TM}$

$-\sum_{a=1}^{3}\left\langle\sum_{\{\alpha,\beta\}=0}^{3}\mathcal{F}^a_{w\alpha\beta}d\xi^\alpha\wedge d\xi^\beta, \sum_{a=1}^{6}\sum_{\{\alpha,\beta\}=0}^{3}K^a_{w\alpha\beta}d\xi^\alpha\wedge d\xi^\beta\right\rangle_{TM}$

$= \sum_{a=1}^{3}G_2\left(\mathcal{F}^a_r, K^a_r\right) - G_2\left(\mathcal{F}^a_w, K^a_w\right)$

$= \sum_{a=1}^{3}G_2\left(\mathcal{F}^{ar}_r, K^{aw}_r\right) + G_2\left(\mathcal{F}^{aw}_r, K^{ar}_r\right) - G_2\left(\mathcal{F}^{ar}_w, K^{aw}_w\right) - G_2\left(\mathcal{F}^{aw}_w, K^{ar}_w\right)$

Which can be expressed equivalently :

$$\left[\begin{array}{l} \langle\mathcal{F},K\rangle_G = \sum_{a=1}^{3}\sum_{\{\alpha\beta\}}\mathcal{F}^{a\alpha\beta}_r K^a_{r\alpha\beta} - \mathcal{F}^{a\alpha\beta}_w K^a_{w\alpha\beta} = \frac{1}{2}\sum_{a=1}^{3}\sum_{\alpha\beta=0}^{3}\mathcal{F}^{a\alpha\beta}_r K^a_{r\alpha\beta} - \mathcal{F}^{a\alpha\beta}_w K^a_{w\alpha\beta} \\ \langle\mathcal{F},K\rangle_G = -\frac{1}{\det P'}Tr\left(\left([*\mathcal{F}^w_r][K^r_r]^t + [*\mathcal{F}^r_r][K^w_r]^t\right) - \left([*\mathcal{F}^w_w][K^r_w]^t + [*\mathcal{F}^r_w][K^w_w]^t\right)\right) \end{array}\right] \tag{5.56}$$

Remark : the scalar product can also be computed by
$\langle \mathcal{F}, K \rangle_G$
$= \sum_{\{\alpha,\beta\}=0}^{3} \sum_{\{\lambda,\mu\}=0}^{3} \left\langle d\xi^\alpha \wedge d\xi^\beta, d\xi^\lambda \wedge d\xi^\mu \right\rangle_{TM} \left\langle \sum_{a=1}^{6} \mathcal{F}_{\alpha\beta}^a \overrightarrow{\kappa}_a, \sum_{b=1}^{6} \sum_{\{\lambda,\mu\}=0}^{3} K_{\lambda\mu}^b \overrightarrow{\kappa}_b \right\rangle_{Cl}$
$= \sum_{a=1}^{3} \sum_{\{\alpha,\beta\}=0}^{3} \sum_{\{\lambda,\mu\}=0}^{3} \left\langle d\xi^\alpha \wedge d\xi^\beta, d\xi^\lambda \wedge d\xi^\mu \right\rangle_{TM} \left(\mathcal{F}_{r\alpha\beta}^a K_{r\lambda\mu}^a - \mathcal{F}_{w\alpha\beta}^a K_{w\lambda\mu}^a \right)$
$= \sum_{a=1}^{3} \left\langle \sum_{\{\alpha,\beta\}=0}^{3} \mathcal{F}_{r\alpha\beta}^a d\xi^\alpha \wedge d\xi^\beta, \sum_{\{\lambda,\mu\}=0}^{3} K_{r\lambda\mu}^a d\xi^\lambda \wedge d\xi^\mu \right\rangle_{TM}$
$- \left\langle \sum_{\{\alpha,\beta\}=0}^{3} \mathcal{F}_{w\alpha\beta}^a d\xi^\alpha \wedge d\xi^\beta, \sum_{\{\lambda,\mu\}=0}^{3} K_{w\lambda\mu}^a d\xi^\lambda \wedge d\xi^\mu \right\rangle_{TM}$
$= \sum_{a=1}^{3} G_2\left(\mathcal{F}_r^a, K_r^a\right) - G_2\left(\mathcal{F}_w^a, K_w^a\right)$

For the other fields

And similarly :

$$\left[\begin{array}{l} \langle \mathcal{F}, K \rangle_A = \sum_{a=1}^{m} \sum_{\{\alpha\beta\}} \mathcal{F}^{a\alpha\beta} K_{\alpha\beta}^a = \frac{1}{2} \sum_{a=1}^{m} \sum_{\alpha\beta=0}^{3} \mathcal{F}^{a\alpha\beta} K_{\alpha\beta}^a \\ \langle \mathcal{F}, K \rangle_A = -\frac{1}{\det P'} Tr\left([*\mathcal{F}^w] [K^r]^t + [*\mathcal{F}^r] [K^w]^t \right) \end{array} \right] \tag{5.57}$$

These scalar products are, by construct, invariant in a change of gauge or chart. So we can compute them in any chart, and of course their expression is simpler in a standard chart. They do not depend on the signature of the metric but they are not definite positive. These quantities can be computed in the usual frame work : for gravity it then involves the Riemann tensor.

From the computation above we have :
$\langle \mathcal{F}, K \rangle_G \varpi_4 = \sum_{a=1}^{3} *\mathcal{F}_r^{ar} \wedge K_r^{aw} + *\mathcal{F}_r^{aw} \wedge K_r^{ar} - *\mathcal{F}_w^{ar} \wedge K_w^{aw} - *\mathcal{F}_w^{aw} \wedge K_w^{ar}$
$\langle \mathcal{F}, \mathcal{F} \rangle_G \varpi_4 = \sum_{a=1}^{3} *\mathcal{F}_r^{ar} \wedge \mathcal{F}_r^{aw} + *\mathcal{F}_r^{aw} \wedge \mathcal{F}_r^{ar} - *\mathcal{F}_w^{ar} \wedge \mathcal{F}_w^{aw} - *\mathcal{F}_w^{aw} \wedge \mathcal{F}_w^{ar}$
$= \sum_{a=1}^{3} *\mathcal{F}_r^{ar} \wedge \mathcal{F}_r^{aw} + *\mathcal{F}_r^{ar} \wedge *\mathcal{F}_r^{aw} - *\mathcal{F}_w^{ar} \wedge \mathcal{F}_w^{aw} - *\mathcal{F}_w^{ar} \wedge \mathcal{F}_w^{aw}$
$\langle \mathcal{F}, \mathcal{F} \rangle_G \varpi_4 = 2\left(\sum_{a=1}^{3} *\mathcal{F}_r^{ar} \wedge \mathcal{F}_r^{aw} - *\mathcal{F}_w^{ar} \wedge \mathcal{F}_w^{aw} \right)$
Similarly :
$\langle \mathcal{F}, \mathcal{F} \rangle_A \varpi_4 = 2\left(\sum_{a=1}^{3} *\mathcal{F}_A^{ar} \wedge \mathcal{F}_A^{aw} \right)$

Identity

We have a useful property which is more general, and holds for all the fields:

Theorem 105 *On the Lie algebra T_1U of a Lie group U, endowed with a symmetric scalar product $\langle\rangle_{T_1U}$ which is preserved by the adjoint map :*

$$\forall X, Y, Z \in T_1U : \langle X, [Y, Z] \rangle = \langle [X, Y], Z \rangle \tag{5.58}$$

Proof. $\forall g \in U : \langle Ad_g X, Ad_g Y \rangle = \langle X, Y \rangle$
take the derivative with respect to g at $g = 1$ for $Z \in T_1U$:
$(Ad_g X)'(Z) = ad(Z)(X) = [Z, X]$
$\langle [Z, X], Y \rangle + \langle X, [Z, Y] \rangle = 0 \Leftrightarrow \langle X, [Y, Z] \rangle = \langle [Z, X], Y \rangle$
exchange X, Z:
$\Rightarrow \langle Z, [Y, X] \rangle = \langle [X, Z], Y \rangle = -\langle [Z, X], Y \rangle = -\langle X, [Y, Z] \rangle = -\langle Z, [X, Y] \rangle$ ■

For the gravitational field :

Let be
$X=\sum_{a=1}^{6}\sum_{\{\alpha,\beta\}=0}^{3}X_{\alpha}^{a}d\xi^{\alpha}\wedge d\xi^{\beta}\otimes\overrightarrow{\kappa}_{a}$,
$Y=\sum_{a=1}^{6}Y_{\alpha}^{a}d\xi^{\alpha}\otimes\overrightarrow{\kappa}_{a}$,
$Z=\sum_{a=1}^{6}Z_{\alpha}^{a}d\xi^{\alpha}\otimes\overrightarrow{\kappa}_{a}$
$[Y,Z]=\sum_{a=1}^{6}\sum_{\{\alpha,\beta\}=0}^{3}[Y_{\alpha},Z_{\beta}]^{a}d\xi^{\alpha}\wedge d\xi^{\beta}\otimes\overrightarrow{\kappa}_{a}$
$\langle X,[Y,Z]\rangle_{G}$
$=\sum_{a=1}^{3}\sum_{\{\alpha\beta\}}X_{r}^{a\alpha\beta}[Y_{\alpha},Z_{\beta}]_{r}^{a}-X_{w}^{a\alpha\beta}[Y_{\alpha},Z_{\beta}]_{w}^{a}$
$=\sum_{a=1}^{3}\sum_{\alpha<\beta}X_{r}^{a\alpha\beta}[Y_{\alpha},Z_{\beta}]_{r}^{a}-X_{w}^{a\alpha\beta}[Y_{\alpha},Z_{\beta}]_{w}^{a}$
$=4\sum_{\alpha<\beta}\left\langle X^{\alpha\beta},[Y_{\alpha},Z_{\beta}]\right\rangle_{Cl}=4\sum_{\alpha<\beta}\left\langle\left[X^{\alpha\beta},Y_{\alpha}\right],Z_{\beta}\right\rangle_{Cl}$
$\langle X,[Y,Z]\rangle_{G}$
$=\frac{1}{2}\sum_{a=1}^{3}\sum_{\alpha,\beta=0}^{3}X_{r}^{a\alpha\beta}[Y_{\alpha},Z_{\beta}]_{r}^{a}-X_{w}^{a\alpha\beta}[Y_{\alpha},Z_{\beta}]_{w}^{a}$
$=2\sum_{\alpha,\beta=0}^{3}\left\langle X^{\alpha\beta},[Y_{\alpha},Z_{\beta}]\right\rangle_{Cl}$

$$\langle X,[Y,Z]\rangle_{G}=2\sum_{\alpha,\beta=0}^{3}\left\langle\left[X^{\alpha\beta},Y_{\alpha}\right],Z_{\beta}\right\rangle_{Cl}\tag{5.59}$$

For the other fields :
$X=\sum_{a=1}^{6}\sum_{\{\alpha,\beta\}=0}^{3}X_{\alpha}^{a}d\xi^{\alpha}\wedge d\xi^{\beta}\otimes\overrightarrow{\theta}_{a},Y=\sum_{a=1}^{6}Y_{\alpha}^{a}d\xi^{\alpha}\otimes\overrightarrow{\theta}_{a},Z=\sum_{a=1}^{6}Z_{\alpha}^{a}d\xi^{\alpha}\otimes\overrightarrow{\theta}_{a}$
$\langle X,[Y,Z]\rangle_{A}$
$=\sum_{a=1}^{m}\sum_{\{\alpha\beta\}}X^{a\alpha\beta}[Y_{\alpha},Z_{\beta}]_{r}^{a}$
$=\sum_{\{\alpha\beta\}}\left\langle X^{\alpha\beta},[Y_{\alpha},Z_{\beta}]\right\rangle_{T_1U}$

$$\langle X,[Y,Z]\rangle_{A}=\sum_{\{\alpha\beta\}}\left\langle\left[X^{\alpha\beta},Y_{\alpha}\right],Z_{\beta}\right\rangle_{T_1U}\tag{5.60}$$

Energy

We have a quantity, which involves only the strength of the field, is invariant in a change of gauge or of chart, and can be used as the variable which we were looking for.

In a system over an area Ω the equilibrium between fields and particles is characterized by an exchange of energy, so naturally $\langle\mathcal{F},\mathcal{F}\rangle$ can be seen as the energy of the field in the vacuum : due to the interactions $\mathcal{F}$ must change all over Ω and this change propagates. But, as well as $\frac{1}{i}\langle\psi,\nabla\psi\rangle$ does not represent an absolute energy of the particle, $\langle\mathcal{F},\mathcal{F}\rangle$ does not represent an absolute energy of the field, only the quantity of energy which is exchanged by the field with itself through the propagation. $\mathcal{F}$ can be seen as the "rate of change of the field" (this is a derivative) with respect to a chart and a gauge, and $\langle\mathcal{F},\mathcal{F}\rangle$ as a scalar measure of this rate.

We will call it, to keep it simple, the density of energy of the field. And we state :

Proposition 106 *The* ***energy density of the fields****, with respect to the volume form ϖ_4, is, up to a constant,*

for the gravitational field :
$\langle\mathcal{F}_G,\mathcal{F}_G\rangle_G=\sum_{a=1}^{3}\sum_{\{\alpha\beta\}}\mathcal{F}_{w}^{a\alpha\beta}\mathcal{F}_{w\alpha\beta}^{a}-\mathcal{F}_{r}^{a\alpha\beta}\mathcal{F}_{r\alpha\beta}^{a}$
$=-\frac{1}{\det P'}Tr\left(\left([*\mathcal{F}_{r}^{w}][\mathcal{F}_{r}^{r}]^{t}+[*\mathcal{F}_{r}^{r}][\mathcal{F}_{r}^{w}]^{t}\right)-\left([*\mathcal{F}_{w}^{w}][\mathcal{F}_{w}^{r}]^{t}+[*\mathcal{F}_{w}^{r}][\mathcal{F}_{w}^{w}]^{t}\right)\right)$
for the other fields :
$\langle\mathcal{F}_A,\mathcal{F}_A\rangle=\sum_{a=1}^{m}\sum_{\{\alpha\beta\}}\mathcal{F}_{A}^{a\alpha\beta}\mathcal{F}_{A\alpha\beta}^{a}=-\frac{1}{\det P'}Tr\left([*\mathcal{F}^{w}][\mathcal{F}^{r}]^{t}+[*\mathcal{F}^{r}][\mathcal{F}^{w}]^{t}\right)$

All these quantities are, of course, estimated up to constants depending on the units.When incorporated in a lagrangian, which represents the energy of a system it corresponds to the energy in the interaction of the field with itself, and provides the usual results.

Because it measures a variation, this quantity is not necessarily positive.

Conservation of energy

For a system comprised only of fields and a given observer, the conservation of energy means that

$\mathcal{E}(t) = \int_{\Omega(t)} \langle \mathcal{F}, \mathcal{F} \rangle \varpi_3 = Ct = \int_{\Omega(t)} i_{\varepsilon_0} \left(\langle \mathcal{F}, \mathcal{F} \rangle \varpi_4 \right)$

Consider the manifold $\Omega([t_1, t_2])$ with borders $\Omega(t_1), \Omega(t_2)$:

$\mathcal{E}(t_2) - \mathcal{E}(t_1) = \int_{\partial\Omega([t_1,t_2])} i_{\varepsilon_0} \left(\langle \mathcal{F}, \mathcal{F} \rangle \varpi_4 \right) = \int_{\Omega([t_1,t_2])} d \left(i_{\varepsilon_0} \langle \mathcal{F}, \mathcal{F} \rangle \varpi_4 \right)$

with the Lie derivative $\pounds_{\varepsilon_0} \left(\langle \mathcal{F}, \mathcal{F} \rangle \varpi_4 \right)$:

$d \left(i_{\varepsilon_0} \langle \mathcal{F}, \mathcal{F} \rangle \varpi_4 \right) = \pounds_{\varepsilon_0} \left(\langle \mathcal{F}, \mathcal{F} \rangle \varpi_4 \right) - i_{\varepsilon_0} d \left(\langle \mathcal{F}, \mathcal{F} \rangle \varpi_4 \right)$

$i_{\varepsilon_0} d \left(\langle \mathcal{F}, \mathcal{F} \rangle \varpi_4 \right) = i_{\varepsilon_0} \left(d \langle \mathcal{F}, \mathcal{F} \rangle \wedge \varpi_4 \right) + i_{\varepsilon_0} \langle \mathcal{F}, \mathcal{F} \rangle d\varpi_4 = i_{\varepsilon_0} \left(d \langle \mathcal{F}, \mathcal{F} \rangle \wedge \varpi_4 \right) = 0$

$d \left(i_{\varepsilon_0} \langle \mathcal{F}, \mathcal{F} \rangle \varpi_4 \right) = \pounds_{\varepsilon_0} \left(\langle \mathcal{F}, \mathcal{F} \rangle \varpi_4 \right)$

$= \left(\pounds_{\varepsilon_0} \langle \mathcal{F}, \mathcal{F} \rangle \right) \varpi_4 + \langle \mathcal{F}, \mathcal{F} \rangle \pounds_{\varepsilon_0} \varpi_4$

$= \frac{1}{c} \frac{\partial}{\partial t} \langle \mathcal{F}, \mathcal{F} \rangle \varpi_4 + \langle \mathcal{F}, \mathcal{F} \rangle \left(div \varepsilon_0 \right) \varpi_4$

$div \varepsilon_0 = div P_0 = \sum_{\alpha=0}^{3} \partial_\alpha P_0^\alpha + P_0^\alpha Tr \left([\partial_\alpha P'] [P] \right)$

$d \left(i_{\varepsilon_0} \langle \mathcal{F}, \mathcal{F} \rangle \varpi_4 \right) = \left(\frac{1}{c} \frac{\partial}{\partial t} \langle \mathcal{F}, \mathcal{F} \rangle + \langle \mathcal{F}, \mathcal{F} \rangle \sum_{\alpha=0}^{3} \partial_\alpha P_0^\alpha + P_0^\alpha Tr \left([\partial_\alpha P'] [P] \right) \right) \varpi_4$

The conservation of energy implies for the observer :

$$\frac{1}{c} \frac{\partial}{\partial t} \langle \mathcal{F}, \mathcal{F} \rangle + \langle \mathcal{F}, \mathcal{F} \rangle \sum_{\alpha=0}^{3} \partial_\alpha P_0^\alpha + P_0^\alpha Tr \left([\partial_\alpha P'] [P] \right) = 0 \tag{5.61}$$

For an observer who uses its standard chart it sums up to $\frac{\partial}{\partial t} \langle \mathcal{F}, \mathcal{F} \rangle + \langle \mathcal{F}, \mathcal{F} \rangle Tr \left(\left[\frac{\partial}{\partial t} Q' \right] [Q] \right) = 0$

Norm for the strength of the field

There is a norm on $\Lambda_2(M; \mathbb{R})$.

On T_1U there is a norm, induced by the definite positive scalar product, so the norm on $\Lambda_2(M; T_1U)$ is :

$\|\mathcal{F}_A\|^2 = \sum_{a=1}^{m} \|\mathcal{F}_A^a\|^2$

Similarly on $\Lambda_2(M; T_1 Spin(3,1))$:

$\|\mathcal{F}_G\|^2 = \sum_{a=1}^{3} \|\mathcal{F}_r^{ra}\|^2 + \|\mathcal{F}_r^{wa}\|^2 + \|\mathcal{F}_w^{ra}\|^2 + \|\mathcal{F}_w^{wa}\|^2$

These norms are defined point wise, and

$\int_\omega \|\mathcal{F}_A(m)\|^2 \varpi_4(m), \int_\omega \|\mathcal{F}_G(m)\|^2 \varpi_4(m)$

are norms on any compact $\omega \subset M$.

If Ω is a relatively compact open of M, the spaces :

$L^2(\Omega, T_1 Spin(3,1), \varpi_4) : \int_\omega \|\mathcal{F}_G(m)\|^2 \varpi_4(m) < \infty$

$L^2(\Omega, T_1U, \varpi_4) : \int_\omega \|\mathcal{F}_A(m)\|^2 \varpi_4(m) < \infty$

where ω is any compact of Ω, are Fréchet spaces.

5.4.4 The phenomenon of propagation

Propagation results from the self-interaction of fields. A system comprised only of fields can be modelled using $\mathcal{F}$ and the tetrad, and the conditions at its equilibrium can be computed using the lagrangian formalism. It provides a set of local partial differential equations, which gives the change of $\mathcal{F}$, and $\mathcal{F}$ itself should be computed with respect to initial values. However the propagation of a physical field is essentially linked to the concept of field, as a physical entity existing in an area of the universe which expands towards the future. And this is this phenomenon, which is not related to the existence of an equilibrium or a lagrangian, that we will presently try to represent mathematically. We will take P_U and $\grave{A}$ as examples.

Mathematical representation

A physical field in the vacuum is located in a delimited 4 dimensional area of M, it has a border, which is a 3 dimensional hypersurface which moves toward the future, in a way similar to a particle which travels along a world line. These borders define a folliation of M by hypersurfaces $W(s)$ as the field propagates There are several ways to consider this general picture.

The field is represented by a connection $\grave{\mathbf{A}}$. The first idea is that W would be surfaces on which the connection is constant. Because this is a one form, the only possible interpretation is that $\grave{\mathbf{A}}$ has a constant value on the tangent space to $W(s)$, and, because of the linearity of $\grave{\mathbf{A}}$, this value must be zero : the tangent space to $W(s)$ should belong to the horizontal bundle HP_U. And W would be generated by an integrable distribution of horizontal vectors, that is by a set of linearly independent vectors fields $\{W_k, k = 1...r \in \mathfrak{X}(HP_U)\}$ such that at any point $p \in P_U$ the vector space spanned by $W_k(p)$ is the tangent space to a submanifold $W \subset P_U$. Using the connection form $\widehat{\grave{A}}$ on TP_U valued in the fixed vector space T_1U, a theorem (Maths.1494) tells that the distribution given by W_k is integrable iff : $\forall u_p, v_p \in T_pW : \widehat{\grave{\mathbf{A}}}(p)\, u_p = \widehat{\grave{\mathbf{A}}}(p)\, v_p = 0 \Rightarrow d\widehat{\grave{\mathbf{A}}}(p)(u_p, v_p) = 0$ where $d\widehat{\grave{\mathbf{A}}}(p)$ is the exterior differential of $\widehat{\grave{\mathbf{A}}}$.

If $W_k \in \mathfrak{X}(HP_U) : \widehat{\grave{\mathbf{A}}}(p)\, W_k = 0, \chi(W_k) = W_k$ thus the condition sums up to

$d\widehat{\grave{\mathbf{A}}}(p)(W_j, W_k) = d\widehat{\grave{\mathbf{A}}}(p)(\chi(W_j), \chi(W_k)) = \nabla_e \widehat{\grave{\mathbf{A}}}(W_j, W_k) = 0$

The surface $W \subset P_U$ projects in a surface $S \subset M$ if the vectors W_k are projectable, which can be done by defining W_k from a vector field $V_k \in \mathfrak{X}(TM)$:

$W(p) = \sum_{\alpha=0}^{3} V_m^\alpha \partial m_\alpha - \zeta\left(Ad_{g^{-1}} \sum_{\alpha=0}^{3} \grave{A}_\alpha(m)\, V_m^\alpha\right)(p)$

To have a hypersurface $S \subset M$ we need a set of 3 linearly independent vector fields V_1, V_2, V_3. To respect the causal structure S must be space like (two points in the same S are never related) and the vector fields V_1, V_2, V_3 are space like. We can then define a usual chart with $\partial\xi_\alpha = V_\alpha, \alpha = 1, 2, 3$ and $\partial\xi_0$ time like, future oriented (it will be transversal to S). In the dual basis $\mathcal{F}_A$ reads :

$\mathcal{F}_{A\alpha\beta} = \mathcal{F}_A(\partial\xi_\alpha, \partial\xi_\beta) = 0$ for $\alpha, \beta = 1, 2, 3$

$\mathcal{F}_{A0\beta} = \mathcal{F}_A(\partial\xi_0, \partial\xi_\beta)$

$\mathcal{F}_A = \mathcal{F}_w = \sum_{a=1}^{m} \sum_{\beta=1}^{3} \mathcal{F}_{0\beta}^a d\xi^0 \wedge d\xi^\beta$

If we change the spatial basis, with the same $\partial\xi_0 : \widetilde{\mathcal{F}}^r = 0$.

For the EM field it implies that the magnetic field is null, and for the gravitational field that the scalar curvature is null. So we can conclude that usually these conditions are not met : there are no 3 dimensional surfaces of equal value of $\grave{\mathbf{A}}$.

Remark : this holds also for the EM field. In the general framework used here, the principal bundle is $P_U(M, U(1), \pi_U)$, the potential is a one form on M valued in $\mathbb{R}$, it can have a constant value on any hypersurface only if it is null. What is commonly called potential is the function from which the electric current is the differential.

But there is another picture. The propagation of the field in the vacuum implies a change in the energy, whatever it can be. So the simplest way to represent the propagation of the field is to define the hypersurfaces $W(s) \subset M$ by :

$W(s) = \{m \in M : F(m) = s\}$

where F is related to $\langle \mathcal{F}_A, \mathcal{F}_A \rangle$, which measures the rate at which the field changes. The function $F(s)$ defines a folliation of M if $F'(s) \neq 0$: there is no local extremum, which can be expected if there is no interaction with particles. The vector field $V = grad(F')$ is normal to $W(s)$ and its flow is a diffeomorphism $\Phi_V(s,m)$. To respect the causal structure V must be time like, future oriented, then two points in the same $W(s)$ are never related (V is transversal). The parameter s, the **phase** of the field, can be seen as its "proper time", so s is increasing for outwards oriented vectors such as V. One can expect that $\langle \mathcal{F}_A, \mathcal{F}_A \rangle$ decreases with s increasing. $s = 0$ at the points closest to the sources, with $\langle \mathcal{F}(m), \mathcal{F}(m) \rangle = \mathcal{E}_0$. At the maximum of $s : \langle \mathcal{F}(m), \mathcal{F}(m) \rangle = 0$. So we take : $F(m) = \mathcal{E}_0 - \langle \mathcal{F}, \mathcal{F} \rangle$

Then F tells how the field varies along its propagation, it is linked to the **range** of the field. It should be a characteristic of the field.

There is no source in the areas delimited by $W(s), s = s_0, s_1$ but $W(s_0)$ appears as a collection of point wise sources, with a field originating from each point and travelling along V.

Any observer has a partial and distorted perception of the field. At each time t the field is seen in the intersection $\Omega_3(t) \cap W(s)$. The intersections are usually 2 dimensional surfaces, the **waves** of the field. Propagation occurs along the integral curves of V, in the 4 dimensional universe, however for an observer it appears as occurring between points in the three dimensional space. These points are actually located in different hypersurfaces $\Omega_3(t)$ if the propagation is not instantaneous, as seen on the diagram below :

$$\begin{array}{cccccc} & t = t_0 & & \varepsilon_0 & & t = t_1 \\ & B = \varphi_o(t_0, b) & \rightarrow & \rightarrow & \rightarrow & C = \varphi_o(t_1, b) = \Phi_V(s_1, n) \\ & \uparrow & & & \nearrow & \\ \overrightarrow{v} & \uparrow & & \nearrow & V & \\ & \uparrow & \nearrow & & & \\ & A = \varphi_0(t_0, a) = \Phi_V(s_0, n) & \rightarrow & \rightarrow & \rightarrow & \varphi_o(t_1, a) \end{array}$$

We have along an integral curve of V going through $n \in W(0)$

$\Phi_V(s,n) = \varphi_o(t(s), x(s))$

By derivation with respect to s

$\varphi'_{ot}(t(s), x(s)) \frac{dt}{ds} + \varphi'_{ox}(t(s), x(s)) \frac{dx}{ds} = V(\Phi_V(s,n))$

$\varphi'_{ot}(t(s), x(s)) \frac{dt}{ds} = c \frac{dt}{ds} \varepsilon_0$

$\varphi'_{ox}(t(s), x(s)) \frac{dx}{ds} = \varphi'_{ox}(t(s), x(s)) \frac{dx}{dt} \frac{dt}{ds}$

$\varphi'_{ox}(t(s), x(s)) \frac{dx}{dt} = \overrightarrow{v}$ is the spatial speed of propagation of the field on the hypersurfaces $\Omega_3(t)$.

$$V(\Phi_V(s,n)) = \frac{dt}{ds}(c\varepsilon_0 + \overrightarrow{v}) \tag{5.62}$$

$\langle V, V \rangle = \left(\frac{dt}{ds}\right)^2 \left(-c^2 + \langle \overrightarrow{v}, \overrightarrow{v} \rangle\right)$

The formula above holds for any observer.

The fact that the propagation is not instantaneous is reflected by $\langle V, V \rangle \leq 0$: V has a time-like component. Because $\overrightarrow{v}$ is a space like vector $\langle \overrightarrow{v}, \overrightarrow{v} \rangle \geq 0$.

$0 \leq \langle \overrightarrow{v}, \overrightarrow{v} \rangle = c^2 + \langle V, V \rangle \left(\frac{d\tau}{dt} \right)^2 \leq c^2$

$\langle V, V \rangle$ does not depend on the observer, so if $\langle \overrightarrow{v}, \overrightarrow{v} \rangle$ does not depend on the observer, necessarily $\langle V, V \rangle = 0, \langle \overrightarrow{v}, \overrightarrow{v} \rangle = c^2$. Then the field propagates along null curves. This property does not depend on the parameter s . Null curves can be oriented by continuity.

V is null iff $\sum_{\lambda\mu} g^{\lambda\mu} (\partial_\lambda F)(\partial_\mu F) = 0$

The experimental problem is that, if the point C belongs to the integral curve going through B (they are spatially the same points) one needs to identify the time at which a field originating from A reaches C. This is done by a characteristic change in the field in A. So the experiments about the propagation of fields concern essentially the EM field that can be modulated at will. They show that :

i) The spatial speed $\langle \overrightarrow{v}, \overrightarrow{v} \rangle$ is constant and the same for all observers.

ii) There is no distortion of the field in the vacuum

iii) The intensity (seen as the energy of the field) decreases with the distance (measured as AB)

Similar experiments are difficult to do with the other fields, however it is generally assumed that the gravitational field shares the same features. The range of the weak and strong interactions is very short, so we will not consider their propagation in this section.

So we will state :

Proposition 107 *In a connected area Ω without interaction with particles, for each field force, the function : $F(m) = \mathcal{E}_0 - \langle \mathcal{F}, \mathcal{F} \rangle$ defines a folliation of Ω in space like hypersurfaces $W(s) = \{m \in M : F(m) = s\}$, the front waves, which are the set with the same value of $\langle \mathcal{F}, \mathcal{F} \rangle$. The field propagates along the null vector field $V = gradF$ normal to each $W(s)$, with*

$F'(m) \neq 0, \sum_{\lambda\mu} g^{\lambda\mu} (\partial_\lambda F)(\partial_\mu F) = 0$

As a consequence :

$V(\Phi_V(s,n)) = \sum_{\alpha=0}^{3} V^\alpha \partial\xi_\alpha = \frac{dt}{ds} (c\varepsilon_0 + \overrightarrow{v})$ with $\langle v, v \rangle = c^2, V^0 = c\frac{dt}{ds}$

In the vacuum there are at least 4 force fields present. It is usually assumed that the fields do not interact with each other, if they propagate along null curves the condition $\sum_{\lambda\mu} g^{\lambda\mu} (\partial_\lambda F)(\partial_\mu F) = 0$ applies independently for each of them. So we are lead to conclude that this condition actually is a constraint on g. The tensor g has 10 degrees of freedom, and these constraints would reduce them at 6, which is the dimension of $SO(3,1)$.

By definition (which does not depend on the existence of a metric), the vectors u tangent to $W(s)$ are $F'(m)u = 0$, the outgoing transversal vectors are : $F'(m)u > 0$. By definition of the gradient : $V = gradF : \forall u : \langle V, u \rangle = F'(m)u$. So, in the definition of the tangent space and transversal vectors, the first definition prevails. But of course V is normal to $W(s)$ because $F'(m)u = 0 \Rightarrow \langle V, u \rangle = 0$.

There is no usual volume form on $W(s)$ because there is no unitary normal to $W(s)$. However the interior product $i_V \varpi_4$ is well defined and can be seen as the measure of the flow of the field through $W(s)$. And the flow through $W(s)$ as $i_V \langle \mathcal{F}, \mathcal{F} \rangle \varpi_4$. The conservation of energy between the hypersurfaces $W(s)$ reads :

$\int_{W(s_2)} i_V \langle \mathcal{F}, \mathcal{F} \rangle \varpi_4 - \int_{W(s_1)} i_V \langle \mathcal{F}, \mathcal{F} \rangle \varpi_4 = \int_{W(s_1,s_2)} d(i_V \langle \mathcal{F}, \mathcal{F} \rangle \varpi_4) = 0$

And as seen before :

$\pounds_V \langle \mathcal{F}, \mathcal{F} \rangle \varpi_4 = d(i_V \langle \mathcal{F}, \mathcal{F} \rangle \varpi_4) + i_V d(\langle \mathcal{F}, \mathcal{F} \rangle \varpi_4) = d(i_V \langle \mathcal{F}, \mathcal{F} \rangle \varpi_4)$

$\pounds_V \langle \mathcal{F}, \mathcal{F} \rangle \varpi_4$
$= \langle \mathcal{F}, \mathcal{F} \rangle \pounds_V \varpi_4 + (\pounds_V \langle \mathcal{F}, \mathcal{F} \rangle) \varpi_4$
$= (\mathcal{E}_0 - F)(divV) \varpi_4 - F'(V) \varpi_4 = (\mathcal{E}_0 - F)(divV) \varpi_4 - (\sum_\alpha V^\alpha \partial_\alpha F) \varpi_4$
$= (\mathcal{E}_0 - F)(divV) \varpi_4 - \left(\sum_{\alpha\beta} g^{\alpha\beta} \partial_\alpha F \partial_\beta F\right) \varpi_4 = (\mathcal{E}_0 - F)(divV) \varpi_4$
$d(i_V \langle \mathcal{F}, \mathcal{F} \rangle \varpi_4) = 0 \Rightarrow (\mathcal{E}_0 - F)(divV) = 0 \Rightarrow$

$$divV = 0 \tag{5.63}$$

$divV == \frac{1}{\det P'} \sum_{\alpha=0}^3 \frac{d}{d\xi^\alpha} (V^\alpha \det P')$
$= \sum_{\alpha=0}^3 \partial_\alpha V^\alpha + V^\alpha Tr([\partial_\alpha P'][P])$
$= \sum_{\alpha=0}^3 \partial_\alpha V^\alpha + Tr\left(\left[\frac{d}{ds} P'\right][P]\right) = 0$
This should hold for any force field :
$Tr\left(\left[\frac{d}{ds} P'\right][P]\right) = -\sum_{\alpha=0}^3 \partial_\alpha V^\alpha$
It can be seen as above as a constraint on V, but also as a constraint on the geometry through P'.

Specification of $\mathcal{F}$

Along the integral curves of V the values of $\mathcal{F}$ are related. We can approach the problem from two sides : using QM theorems and Differential Geometry.

Quantization of $\mathcal{F}$

On a connected, relatively compact area Ω without particles, one can assume that
$\mathcal{F}_G \in \mathcal{V}_G \subset L^2(\Omega, T_1 Spin(3,1), \varpi_4)$,
$\mathcal{F}_A \in \mathcal{V}_A \subset L^2(\Omega, T_1 U, \varpi_4)$
which are Fréchet spaces. Moreover they can be seen as sections of the vector bundles $\mathcal{F}_G \in \mathfrak{X}(P_G[T_1 Spin(3,1), \mathbf{Ad}])$, $\mathcal{F}_A \in \mathfrak{X}(P_U[T_1 U, Ad])$, which are representations of the groups, and the value of the scalar product, representing the energy, does not depend on the choice of a chart or gauge.

We can implement the Theorem 24 : the spaces $\mathcal{V}_G, \mathcal{V}_A$ can be embedded in Hilbert spaces, which are unitary representation of $Spin(3,1), U$. To each function F_G, F_A is associated a subspace of the Hilbert space invariant by the action of the groups, that is an irreducible representation.

For the gravitational field the irreducible representations of $Spin(3,1)$ are parametrized by a scalar, which is the total energy in Ω, and a signed integer $z \in \mathbb{Z}$. For a given observer, $\mathcal{V}_G$ is globally invariant by $SO(3)$ which is compact, so it is isomorphic to one finite dimensional representation, parametrized by an integer and $j = z$.

For the EM field each irreducible representation is defined by a single vector $\mathcal{F}$.

For the other fields U is compact, the irreducible representations are finite dimensional, so for each function F_A the 2 form $\mathcal{F}_A$ belongs to a finite dimensional vector space.

The field is quantized : for each map F, that is for each area without particles, $\mathcal{F}$ belongs to a finite dimensional vector space of maps, characterized by integers.

As a consequence the local PDE defining $\mathcal{F}$ should be linear, such that a set of general solutions (adjusted to the initial conditions on the border of Ω) constitute a finite dimensional space, characterized by a finite number of parameters.

We can go further. Let us consider an integral curve of V originating at $n \in W(0)$: $\langle \mathcal{F}(n), \mathcal{F}(n) \rangle = \mathcal{E}_0$.We can expect that the evaluation map :

$\mathcal{E}(s)\mathcal{F} = \mathcal{F}(\Phi_V(s,n))$

is continuous, and the evolution is semi-determinist, because we are in the vacuum. Then we can implement the theorem 26:

- there is a Hilbert space H on the set of values $\{\mathcal{F}(\Phi_V(s,n)), s \in [0,S]\}$

- for each s there is a unitary operator $\Theta(s) \in \mathcal{L}(H;H)$ such that : $\Theta(s)\mathcal{F}(\Phi_V(0,n)) = \mathcal{F}(\Phi_V(s,n))$

Because $\mathcal{V}$ is finite dimensional for a given F, H is finite dimensional, however we do not know the scalar product. We could expect that this is the scalar product defined previously, but it is definite only on cones, and this would imply for the EM field that either the electric or the magnetic field is null.

Moreover Θ depends on n. What we can say is that there is a linear map : $L(s,n)$ such that :

$\mathcal{F}(\Phi_V(s,n)) = L(s,n)\mathcal{F}(n)$

Consider a global change of gauge on $\mathfrak{X}(P_U[T_1U, Ag])$:

$\mathbf{p}_U(m) \to \widetilde{\mathbf{p}}_U(m) = \widetilde{\varphi}_U(m,1) = \mathbf{p}_U(m) \cdot g^{-1}$

$\mathcal{F}(\Phi_V(s,n)) \to Ad_g\mathcal{F}(\Phi_V(s,n))$

$\mathcal{F}(n) \to Ad_g\mathcal{F}(n)$

and we must have :

$Ad_g\mathcal{F}(\Phi_V(s,n)) = Ad_g L(s,n)\mathcal{F}(n) = L(s,n) Ad_g\mathcal{F}(n)$

thus $L(s,n)$ must commute with Ad. The only solution is : $L(s,n) = \lambda(s,n) Ad_{g(s,n)}$ for a scalar function $\lambda \in C_1(M;\mathbb{R})$ and $g \in P_U$ with $\lambda(s_0,n) = 1, g(s_0,n) = 1$:

$\mathcal{F}(\Phi_V(s,n)) = \lambda(s,n) Ad_{g(s,n)}\mathcal{F}(n)$

As we have :

$\langle \mathcal{F}(\Phi_V(s,n)), \mathcal{F}(\Phi_V(s,n)) \rangle = \left\langle \lambda(s,n) Ad_{g(s,n)}\mathcal{F}(n), \lambda(s,n) Ad_{g(s,n)}\mathcal{F}(n) \right\rangle = (\lambda(s,n))^2 \langle \mathcal{F}(n), \mathcal{F}(n) \rangle$

$\langle \mathcal{F}(\Phi_V(s,n)), \mathcal{F}(\Phi_V(s,n)) \rangle = \mathcal{E}_0 - F(\Phi_V(s,n))$

$\langle \mathcal{F}(n), \mathcal{F}(n) \rangle = \mathcal{E}_0$

$\lambda(s,n) = \sqrt{\frac{\mathcal{E}_0 - F(\Phi_V(s,n))}{\mathcal{E}_0}}$ does not depend on n

$$\mathcal{F}(\Phi_V(s,n)) = \sqrt{\frac{\mathcal{E}_0 - F(\Phi_V(s,n))}{\mathcal{E}_0}} Ad_{g(s,n)}\mathcal{F}(n) \tag{5.64}$$

The formula implies that $\mathcal{E}_0 - s > 0$. The phase does not mark an absolute location, but a location with respect to the inner border of Ω. If one assumes that the field propagates at a finite speed and originates in its interactions with particles, it is unavoidable that Ω covers a finite area, beyond which rests the true vacuum.

Differential Geometry

The formula $\mathcal{F}(\Phi_V(s,n)) = \lambda(s,n) Ad_{g(s,n)}\mathcal{F}(n)$ involves 2 forms, which are not defined in the same vector space. It can be made more precise. The basic idea is that $\mathcal{F}$ is transported along the integral curve of V. Such transport can be represented mathematically in the fiber bundle, using the connection and a section of P_U. But if we look at the definition of $\mathcal{F} = -\mathbf{p}^*(m)\chi^* d\widehat{\grave{A}}$ it incorporates already all these ingredients. The other transport is by push-forward, it relies on the images of $\mathcal{F}$ by the derivative of the flow of V.

The pull back of $\mathcal{F}$ by Φ_V is :

$\Phi_V(s,\cdot)^*\mathcal{F}(m) = \mathcal{F}(\Phi_V(s,m))(\Phi_V(s,m))'$

The push forward of $\mathcal{F}$ by Φ_V is :

$\Phi_V(s,\cdot)_* \mathcal{F}(\Phi_V(s,m)) = \mathcal{F}(m) \circ (\Phi_V(-s,m))'$

Pull back and push forward are inverse of each other :

$\Phi_V(s,\cdot)_* = (\Phi_V(s,\cdot)^*)^{-1} = \Phi_V(-s,\cdot)^*$

$\mathcal{F}(\Phi_V(s,m)), \Phi_V(s,\cdot)_* \mathcal{F}(\Phi_V(s,m))$ belong to the same tangent space, and are valued in the same fixed vector space.

Then along the integral curve $m(s) = \Phi_V(s,n)$ one can say that the values of $\mathcal{F}$ are related if, using $\mathcal{F}_A$ as example :

$\exists L \in C_1(\mathbb{R} \times W(0); \mathcal{L}(T_1U; T_1U)), L(W(0)) = Id$

$\mathcal{F}(\Phi_V(s,n)) = L(s,n)(\Phi_V(s,\cdot)_* \mathcal{F}(\Phi_V(s,n)))$

and we retrieve the formula above. With the same reasoning : $L(s,n) = \lambda(s) Ad_{g(s,n)}$.

Then the value of $\mathcal{F}$ at any point $\Phi_V(s,n)$ along the integral curve is deduced from the value of $\mathcal{F}(n)$ by

$\mathcal{F}(\Phi_V(s,n)) = L(s)\mathcal{F}(n) \circ (\Phi_V(-s,n))'$

Let us denote : $m = \Phi_V(s,n) \Leftrightarrow n = \Phi_V(-s,m)$

$\mathcal{F}(\Phi_V(s,\Phi_V(-s,m))) = L(s,n)(\Phi_V(s,\cdot)_* \mathcal{F}(\Phi_V(s,\Phi_V(-s,m))))$

$\mathcal{F}(m) = L(s)(\Phi_V(s,\cdot)_* \mathcal{F}(m))$

$\mathcal{F}(m) = L(-s)(\Phi_V(-s,\cdot)_* \mathcal{F}(m))$

$\forall s : \mathcal{F}(m) = L(-s)\left(\Phi_V(s,\cdot)^* \mathcal{F}(m)\right)$

By taking the derivative with respect to s at $s = 0$:

$-L'(0)\mathcal{F}(m) + \frac{d}{ds}\Phi_V(s,m)^* \mathcal{F}(m)|_{s=0} = 0$

$\frac{d}{ds}\Phi_V(s,m)^* \mathcal{F}(m)|_{s=0}$ is the Lie derivative of $\mathcal{F}$ along V :

$$\pounds_V \mathcal{F}(m) = L'(0)\mathcal{F}(m) \tag{5.65}$$

$L'(0)\mathcal{F}_A(m) = \sum_{\{\alpha\beta\}} \sum_{a,b=1}^m [L'(0)]_b^a \mathcal{F}_{\alpha\beta}^a(m) d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\theta}_a$

The value of the Lie derivative is (see Annex) :

$\pounds_V \mathcal{F}(m) = \sum_{a=1}^m \left(\sum_{\{\alpha\beta\}} \sum_{\gamma=0}^3 V^\gamma \partial_\gamma \mathcal{F}_{\alpha\beta}^a d\xi^\alpha \wedge d\xi^\beta + \mathcal{F}_{\alpha\beta}^a \left(\partial_\gamma V^\alpha d\xi^\gamma \wedge d\xi^\beta + \partial_\gamma V^\beta d\xi^\alpha \wedge d\xi^\gamma\right)\right) \otimes \overrightarrow{\theta}_a$

$[(\pounds_V \mathcal{F}(m))^r] = \sum_{\gamma=0}^3 V^\gamma [\partial_\gamma \mathcal{F}^{ar}] + [\mathcal{F}^{aw}] j\left(\left[\partial V^0\right]\right) + [\mathcal{F}^{ar}]\left(-[\partial v]^t + (div(v)) I_3\right)$

$[(\pounds_V \mathcal{F}(m))^w] = \sum_{\gamma=0}^3 V^\gamma [\partial_\gamma \mathcal{F}^{aw}] + [\mathcal{F}^{aw}]\left(\partial_0 V^0 + [\partial v]\right) - [\mathcal{F}^{ar}] j(\partial_0 V)$

with :

$[\partial_\beta V^\alpha]_{\beta=0...3}^{\alpha=0...3} = \begin{bmatrix} \partial_0 V^0 & [\partial_0 V]_{1\times 3} \\ \left[\partial V^0\right]_{3\times 1} & [\partial V]_{3\times 3} \end{bmatrix}$

$[\partial v] = \begin{bmatrix} \partial_1 V^1 & \partial_2 V^1 & \partial_3 V^1 \\ \partial_1 V^2 & \partial_2 V^2 & \partial_3 V^2 \\ \partial_1 V^3 & \partial_2 V^3 & \partial_3 V^3 \end{bmatrix}$

$div(v) = \partial_1 V^1 + \partial_2 V^2 + \partial_3 V^3 = Tr[\partial v]$

$j(\partial_0 V) = \begin{bmatrix} 0 & -\partial_0 V^3 & \partial_0 V^2 \\ \partial_0 V^3 & 0 & -\partial_0 V^1 \\ -\partial_0 V^2 & \partial_0 V^1 & 0 \end{bmatrix}$

Moreover :

$divV = \sum_{\alpha=0}^3 \partial_\alpha V^\alpha + V^\alpha Tr([\partial_\alpha P'][P]) = 0$

$div(v) = \partial_0 V^0 - Tr\left(\left[\frac{d}{ds}P'\right][P]\right)$

$\langle V, V\rangle = 0 \Rightarrow \left(V^0\right)^2 = \langle v, v\rangle$

$\sum_{\alpha=1}^3 V^\alpha \partial_0 V^\alpha = V^0 \partial_0 V^0 \Leftrightarrow [v]^t [\partial_0 V] = V^0 \partial_0 V^0$

$\sum_\alpha V^\alpha \partial_\beta V^\alpha = V^0 \partial_\beta V^0 \Leftrightarrow [v]^t [\partial v] = V^0 \left[\partial V^0\right]$

So $\mathcal{F}$ is solution of the $6m$ PDE with parameters the fixed matrix $[L'(0)]$:
For $a = 1...m$:
$\sum_{b=1}^{m} [L'(0)]_b^a [\mathcal{F}^{br}] = \sum_{\gamma=0}^{3} V^\gamma [\partial_\gamma \mathcal{F}^{ar}] + [\mathcal{F}^{aw}] j(\partial V^0) + [\mathcal{F}^{ar}] \left(-[\partial v]^t + (div(v)) I_3\right)$
$\sum_{b=1}^{m} [L'(0)]_b^a [\mathcal{F}^{bw}] = \sum_{\gamma=0}^{3} V^\gamma [\partial_\gamma \mathcal{F}^{aw}] + [\mathcal{F}^{aw}] (\partial_0 V^0 + [\partial v]) - [\mathcal{F}^{ar}] j(\partial_0 V)$
Pointwise $\mathcal{F}$ has $6m$ degrees of freedom, and, for a given vector field V (that is a function F) the solutions of the $6m$ PDE constitute a vector space parametrized by the m^2 values of $[L'_s(0,n)]_b^a$. The general solutions of the PDE are then adjusted to the initial conditions, on the border of the area Ω. The quantization told us that this vector space is finite dimensional, so the coefficients of $[L'(0)]$ are linearly related.

A solution to the PDE is found by a kind of separation of variables, by stating : $\mathcal{F} = \widehat{\mathcal{F}} \otimes X$ with $\widehat{\mathcal{F}} \in \Lambda_2(M;\mathbb{R})$, and $X \in C_1(M;T_1U)$. Then :
$\pounds_V \mathcal{F} = \widehat{\mathcal{F}} \otimes \sum_{\gamma=0}^{3} V^\gamma \partial_\gamma X + \pounds_V \widehat{\mathcal{F}} \otimes X$
The PDE reads :
$\widehat{\mathcal{F}} \otimes L'_s(0,n) X = \widehat{\mathcal{F}} \otimes \sum_{\gamma=0}^{3} V^\gamma \partial_\gamma X + \pounds_V \widehat{\mathcal{F}} \otimes X$
If $\widehat{\mathcal{F}}$ is solution of $\pounds_V \widehat{\mathcal{F}} = 0$, then $\mathcal{F}$ is solution of the PDE if X is solution of :
$L'(0) X(m) = \sum_{\gamma=0}^{3} V^\gamma \partial_\gamma X(m)$ with $m = \Phi_V(s,n)$
$X(\Phi_V(s,n)) = \lambda(s) Ad_{g(s,n)} X(n)$
Except in the modelling of deformable solids the gauge is arbitrary. So it makes sense to take as gauge $X(n) = X_0 = Ct$ on the border $\partial\Omega$ of the vacuum. Then :

$$X(\Phi_V(s,n)) = \lambda(s) Ad_{g(s,n)} X_0 \tag{5.66}$$

The field appears as a deformation of the field X_0 on $\partial\Omega$, which is similar to the fundamental state ψ_0 of particles, by a local change of gauge given by a section g.

The condition on $\widehat{\mathcal{F}}$ reads :

$$\left[\begin{array}{l} \pounds_V \widehat{\mathcal{F}} = 0 \\ \left[\frac{d}{ds}\widehat{\mathcal{F}}^r\right] = \left[\widehat{\mathcal{F}}^r\right] \left([\partial v]^t - (div(v)) I_3\right) - \left[\widehat{\mathcal{F}}^w\right] j([\partial V^0]) \\ \left[\frac{d}{ds}\widehat{\mathcal{F}}^w\right] = \left[\widehat{\mathcal{F}}^r\right] j(\partial_0 V) - \left[\widehat{\mathcal{F}}^w\right] (\partial_0 V^0 + [\partial v]) \end{array}\right] \tag{5.67}$$

$\Leftrightarrow \widehat{\mathcal{F}}(\Phi_V(s,m)) = \widehat{\mathcal{F}}(m) \circ (\Phi_V(-s,m))'$

For the EM field $T_1U = \mathbb{R}$ and X is limited to $\lambda(s,n)$. $\mathcal{F}$ is a closed form so the condition reads :
$\widehat{\mathcal{F}} = d\grave{A} \Rightarrow \pounds_V \widehat{\mathcal{F}} = di_V \widehat{\mathcal{F}} = d\left(i_V d\grave{A}\right) = 0$

The solution of $\pounds_V \widehat{\mathcal{F}} = 0$ can be easily found, at least along the integral curves of V. It does not depend on the chart. Take as chart $\Phi_V(s,n)$ with $n \in \partial\Omega$ the border of the area considered. The jacobian is : $\partial\eta_\alpha = \sum_\beta J_\alpha^\beta \partial\xi_\beta$, the inverse $[K] = [J]^{-1}$, the components of V :
$V \to \widetilde{V} = [K] V = \begin{bmatrix} K_0^0 & [K^0] \\ [K_0] & [k]_{3\times 3} \end{bmatrix} \begin{bmatrix} V^0 \\ v \end{bmatrix} = \begin{bmatrix} c \\ 0 \end{bmatrix} = \begin{bmatrix} K_0^0 V^0 + [K^0] v \\ [K_0] V^0 + [k] v \end{bmatrix}$
$K_0^0 V^0 + [K^0] v = c$
$[K_0] V^0 + [k] v = 0$
Take

$\left[K^0\right] = -\frac{1}{c}\left[v\right]^t \Rightarrow \left[K^0\right] v = -\frac{1}{c}\left(V^0\right)^2$
$K_0^0 V^0 - \frac{1}{c}\left(V^0\right)^2 = c$
$K_0^0 = \frac{c^2+\left(V^0\right)^2}{cV^0}$
$\left[k\right] = \frac{1}{c}\left(j\left(v\right) - V^0 I_3\right), \left[K_0\right] = \frac{1}{c}v,$
$\left[K_0\right]V^0 + \left[k\right]v = \frac{1}{c}vV^0 + \frac{1}{c}\left(j\left(v\right) - V^0 I_3\right)v = 0$
$\det\left[k\right] = -\frac{1}{c^3}V^0\left(\left(V^0\right)^2 + v^t v\right) = -2\left(\frac{V^0}{c}\right)^3 \neq 0$
$\left[K\right] = \left[\begin{array}{cc} \frac{c^2+\left(V^0\right)^2}{cV^0} & -\left[\frac{v}{c}\right]^t \\ \frac{v}{c} & \frac{1}{c}\left(j\left(v\right) - V^0 I_3\right) \end{array}\right]$
$\partial\xi_0 = \left(\frac{c^2+\left(V^0\right)^2}{cV^0}\right)\partial\eta_0 - \sum_{\alpha=1}^3 \frac{v^\alpha}{c}\partial\eta_\alpha$
$\alpha = 1,2,3 : \partial\xi_\alpha = \frac{v^\alpha}{c}\partial\eta_0 + \frac{1}{c}\sum_{\beta=1}^3\left(j\left(v\right)^\beta - V^0\right)\partial\eta_\beta$
$\det\left[K\right] = -2\left(\frac{V^0}{c}\right)^2 \neq 0$

In the new chart the equations $\pounds_V \widehat{\mathcal{F}} = 0$ sum up to $\frac{d}{ds}\widetilde{\left[\widehat{\mathcal{F}}^r\right]} = 0, \frac{d}{ds}\widetilde{\left[\widehat{\mathcal{F}}^w\right]} = 0$. The components of $\widehat{\mathcal{F}}$ in the holonomic basis linked to V are constant (but $\widehat{\mathcal{F}}$ itself is not constant).

$\left[\widehat{\mathcal{F}}^r\left(\widetilde{\Phi_V\left(s,n\right)}\right)\right] = \left[\widehat{\mathcal{F}}^r\left(\widetilde{\Phi_V\left(0,n\right)}\right)\right]$
$\left[\widehat{\mathcal{F}}^w\left(\widetilde{\Phi_V\left(s,n\right)}\right)\right] = \left[\widehat{\mathcal{F}}^w\left(\widetilde{\Phi_V\left(0,n\right)}\right)\right]$
To come back to the original chart :
$\left[\begin{array}{cc} \widetilde{\left[\widehat{\mathcal{F}}^r\left(m\right)\right]} & \widetilde{\left[\widehat{\mathcal{F}}^w\left(m\right)\right]} \end{array}\right] = \left[\begin{array}{cc} \left[\widehat{\mathcal{F}}^r\left(m\right)\right] & \left[\widehat{\mathcal{F}}^w\left(m\right)\right] \end{array}\right]\left[M\right] = \left[\begin{array}{cc} \widetilde{\left[\widehat{\mathcal{F}}^r\left(n\right)\right]} & \widetilde{\left[\widehat{\mathcal{F}}^w\left(n\right)\right]} \end{array}\right]$
$\left[M\right] = \frac{1}{c^2}\left[\begin{array}{cc} j\left(v\right)j\left(v\right) - V^0 j\left(v\right) + 2\left(V^0\right)^2 & -j\left(v\right)j\left(v\right) + V^0 j\left(v\right) \\ -j\left(v\right)j\left(v\right) + V^0 j\left(v\right) & \left(\frac{c^2+\left(V^0\right)^2}{V^0}\right)j\left(v\right) - c^2 \end{array}\right]$
$\left[\begin{array}{cc} \left[\widehat{\mathcal{F}}^r\left(m\right)\right] & \left[\widehat{\mathcal{F}}^w\left(m\right)\right] \end{array}\right]$
$= c^2\left[\begin{array}{cc} \left[\widehat{\mathcal{F}}^r\left(n\right)\right] & \left[\widehat{\mathcal{F}}^w\left(n\right)\right] \end{array}\right]\left[\begin{array}{cc} j\left(v\right)j\left(v\right) - V^0 j\left(v\right) + 2\left(V^0\right)^2 & -j\left(v\right)j\left(v\right) + V^0 j\left(v\right) \\ -j\left(v\right)j\left(v\right) + V^0 j\left(v\right) & \left(\frac{c^2+\left(V^0\right)^2}{V^0}\right)j\left(v\right) - c^2 \end{array}\right]^{-1}$

Moreover, according to our assumptions, on each integral curve, for some $n \in W\left(0\right)$:
$\mathcal{F}\left(\Phi_V\left(s,n\right)\right) = \sqrt{\frac{\mathcal{E}_0 - s}{\mathcal{E}_0}}Ad_{g(s,n)}\mathcal{F}\left(n\right)$
$\left\langle \mathcal{F}\left(\Phi_V\left(s,n\right)\right), \mathcal{F}\left(\Phi_V\left(s,n\right)\right)\right\rangle = s - \mathcal{E}_0$
$= \left\langle \widehat{\mathcal{F}}\left(\Phi_V\left(s,n\right)\right) \otimes X\left(\Phi_V\left(s,n\right)\right), \widehat{\mathcal{F}}\left(\Phi_V\left(s,n\right)\right) \otimes X\left(\Phi_V\left(s,n\right)\right)\right\rangle$
$= \left\langle \widehat{\mathcal{F}}\left(\Phi_V\left(s,n\right)\right) \otimes \lambda\left(s\right)Ad_{g(s,n)}X\left(n\right), \widehat{\mathcal{F}}\left(\Phi_V\left(s,m\right)\right) \otimes \lambda\left(s\right)Ad_{g(s,n)}X\left(n\right)\right\rangle$
$= G_2\left(\widehat{\mathcal{F}}\left(\Phi_V\left(s,n\right)\right), \widehat{\mathcal{F}}\left(\Phi_V\left(s,n\right)\right)\right)\left\langle \lambda\left(s\right)Ad_{g(s,n)}X\left(n\right), \lambda\left(s\right)Ad_{g(s,n)}X\left(n\right)\right\rangle_{T_1U}$
$= \left(\lambda\left(s\right)\right)^2 G_2\left(\widehat{\mathcal{F}}\left(\Phi_V\left(s,n\right)\right), \widehat{\mathcal{F}}\left(\Phi_V\left(s,n\right)\right)\right)\left\langle X\left(n\right), X\left(n\right)\right\rangle_{T_1U}$
$= \left(\frac{\mathcal{E}_0 - s}{\mathcal{E}_0}\right)G_2\left(\widehat{\mathcal{F}}\left(\Phi_V\left(s,n\right)\right), \widehat{\mathcal{F}}\left(\Phi_V\left(s,n\right)\right)\right)\left\langle X\left(n\right), X\left(n\right)\right\rangle_{T_1U}$
$= s - \mathcal{E}_0$
$G_2\left(\widehat{\mathcal{F}}\left(\Phi_V\left(s,n\right)\right), \widehat{\mathcal{F}}\left(\Phi_V\left(s,n\right)\right)\right)\left\langle X\left(n\right), X\left(n\right)\right\rangle_{T_1U} = -\mathcal{E}_0 = G_2\left(\widehat{\mathcal{F}}\left(n\right), \widehat{\mathcal{F}}\left(n\right)\right)\left\langle X\left(n\right), X\left(n\right)\right\rangle_{T_1U}$
$\left\langle X\left(n\right), X\left(n\right)\right\rangle_{T_1U} \neq 0$

$$G_2\left(\widehat{\mathcal{F}}\left(\Phi_V\left(s,n\right)\right),\widehat{\mathcal{F}}\left(\Phi_V\left(s,n\right)\right)\right)=G_2\left(\widehat{\mathcal{F}}\left(n\right),\widehat{\mathcal{F}}\left(n\right)\right) \tag{5.68}$$

And we state :

Proposition 108 *In a connected area Ω without interaction with particles, there are, for each field :*
i) A scalar 2 form $\widehat{\mathcal{F}}\in\Lambda_2\left(M;\mathbb{R}\right)$, *such that* $\pounds_V\widehat{\mathcal{F}}=0\iff\widehat{\mathcal{F}}\left(\Phi_V\left(s,m\right)\right)=\widehat{\mathcal{F}}\left(m\right)\circ\left(\Phi_V\left(-s,m\right)\right)'$
ii) Sections
$X_G\in\mathfrak{X}\left(P_G\left[T_1Spin\left(3,1\right),\mathbf{Ad}\right]\right),X_A\in\mathfrak{X}\left(P_U\left[T_1U,Ad\right]\right),$
iii) maps :
$g_A:\mathbb{R}\times W\left(0\right)\to U,g_G:\mathbb{R}\times W\left(0\right)\to Spin\left(3,1\right)$
such that
$\mathcal{F}_G\left(m\right)=\widehat{\mathcal{F}}_G\left(m\right)\otimes X_G\left(m\right);\mathcal{F}_A\left(m\right)=\widehat{\mathcal{F}}_A\left(m\right)\otimes X_A\left(m\right)$
and, along the integral curves of V going through $n\in W\left(0\right)$:
$X_G\left(\Phi_V\left(s,n\right)\right)=\sqrt{\frac{\mathcal{E}_{0G}-s}{\mathcal{E}_{0G}}}\mathbf{Ad}_{g_G(s,n)}X_G\left(n\right)$
$X_A\left(\Phi_V\left(s,n\right)\right)=\sqrt{\frac{\mathcal{E}_{0A}-s}{\mathcal{E}_{0A}}}Ad_{g_A(s,n)}X_A\left(n\right)$

Then :
$[\mathcal{F}_A^r]=[X_A]_{m\times1}\left[\widehat{\mathcal{F}}_A^r\right]_{1\times3};[\mathcal{F}_A^w]=[X_A]_{m\times1}\left[\widehat{\mathcal{F}}_A^w\right]_{1\times3}$
$[\mathcal{F}_r^r]=[X_r]_{3\times1}\left[\widehat{\mathcal{F}}_G^r\right]_{1\times3};[\mathcal{F}_r^w]=[X_r]_{m\times1}\left[\widehat{\mathcal{F}}_G^w\right]_{1\times3}$
$[\mathcal{F}_w^r]=[X_w]_{3\times1}\left[\widehat{\mathcal{F}}_G^r\right]_{1\times3};[\mathcal{F}_w^w]=[X_w]_{3\times1}\left[\widehat{\mathcal{F}}_G^w\right]_{1\times3}$
$[*\mathcal{F}_r^r]=[X_r]\left[*\widehat{\mathcal{F}}^r\right];[*\mathcal{F}_w^r]=[X_w]\left[*\widehat{\mathcal{F}}^r\right]$
$[*\mathcal{F}_r^w]=[X_r]\left[*\widehat{\mathcal{F}}^w\right];[*\mathcal{F}_w^w]=[X_w]\left[*\widehat{\mathcal{F}}^w\right]$
$[*\mathcal{F}_A^r]=[X_A]\left[*\widehat{\mathcal{F}}_A^r\right];[\mathcal{F}_A^w]=[X_A]\left[*\widehat{\mathcal{F}}_A^w\right]$

$\left\langle\mathcal{F}_A,\mathcal{F}_A\right\rangle_A$
$=\frac{1}{\det P'}Tr\left([X_A]\left[*\widehat{\mathcal{F}}_A^w\right]\left[\widehat{\mathcal{F}}_A^r\right]^t[X_A]^t+[X_A]\left[*\widehat{\mathcal{F}}_A^r\right]\left[\widehat{\mathcal{F}}_A^w\right]^t[X_A]^t\right)$

$$\left\langle\mathcal{F}_A,\mathcal{F}_A\right\rangle_A=\frac{1}{\det P'}\left([X_A]^t[X_A]\right)\left(\left(\left[*\widehat{\mathcal{F}}_A^w\right]\left[\widehat{\mathcal{F}}_A^r\right]^t\right)+\left(\left[*\widehat{\mathcal{F}}_A^r\right]\left[\widehat{\mathcal{F}}_A^w\right]^t\right)\right) \tag{5.69}$$

$\left\langle\mathcal{F},\mathcal{F}\right\rangle_G=\frac{1}{\det P'}Tr\{\left([X_r]\left[*\widehat{\mathcal{F}}_G^w\right]\left[\widehat{\mathcal{F}}_G^r\right]^t[X_r]^t+[X_r]\left[*\widehat{\mathcal{F}}_G^r\right]\left[\widehat{\mathcal{F}}^w\right]^t[X_r]^t\right)$
$-\left([X_w]\left[*\widehat{\mathcal{F}}^w\right]\left[\widehat{\mathcal{F}}_G^r\right]^t[X_w]^t+[X_w]\left[*\widehat{\mathcal{F}}^r\right]\left[\widehat{\mathcal{F}}_G^w\right]^t[X_w]^t\right)\}$
$=\frac{1}{\det P'}\{\left([X_r]^t[X_r]\right)\left(\left[*\widehat{\mathcal{F}}^w\right]\left[\widehat{\mathcal{F}}_G^r\right]^t+\left[*\widehat{\mathcal{F}}^r\right]\left[\widehat{\mathcal{F}}^w\right]^t\right)-\left([X_w]^t[X_w]\right)\left(\left[*\widehat{\mathcal{F}}^w\right]\left[\widehat{\mathcal{F}}_G^r\right]^t+\left[*\widehat{\mathcal{F}}^r\right]\left[\widehat{\mathcal{F}}_G^w\right]^t\right)\}$

$$\left\langle\mathcal{F},\mathcal{F}\right\rangle_G=\frac{1}{\det P'}\left(\left[*\widehat{\mathcal{F}}_G^w\right]\left[\widehat{\mathcal{F}}_G^r\right]^t+\left[*\widehat{\mathcal{F}}_G^r\right]\left[\widehat{\mathcal{F}}^w\right]^t\right)\left([X_r]^t[X_r]-[X_w]^t[X_w]\right) \tag{5.70}$$

The cosmological constant

So we have a sensible model for the propagation of the field. It relies on the existence of a scalar function F, characteristic of the field and which defines the vector V.

We will see in the following chapters continuous models, based on the application of the Principle of Least Action and a lagrangian. Their purpose is to find an equilibrium. When they involve particles and fields the lagrangian expresses the fact that there is a balance between all the physical components of the system, measured by the exchange of energy through $\frac{1}{i}\langle\psi,\nabla_V\psi\rangle$ and $\langle\mathcal{F},\mathcal{F}\rangle$ (or similar expressions). They can be used "in the vacuum", that is where there is no particle involved, and should give the conditions of the propagation. However their solutions are not compatible with the results presented in this section : put together they lead to $\langle\mathcal{F},\mathcal{F}\rangle=0$. There is a reason for that : the application of the principle of least action to a system comprised only of fields, with a lagrangian limited to $\langle\mathcal{F},\mathcal{F}\rangle$ leads naturally to the solution $\langle\mathcal{F},\mathcal{F}\rangle=0:\mathcal{F}$ measures the rate of change of the field, and an equilibrium is reached when this rate is null. But such a continuous model is somewhat unphysical. One should account for the initial values, which are not null because there is a particle somewhere, and more fundamentally because propagation is not an equilibrium. There is always an area, in the future of Ω which should be accounted for, eventually a genuine vacuum is Ω is large enough, which is filled by the field.

This issue has been raised, in another way, in the cosmological models. General Relativity enables to conceive models based on the Einstein Equation $Ric_{\alpha\beta}-\frac{1}{2}g_{\alpha\beta}\mathbf{R}=0$, which can be proven by the application of the principle of least action to a system comprised only of the gravitational field, and the Hilbert Action $\mathbf{R}\varpi_4$. They lead to "static universe", without singularity (no Big Bang). To give more flexibility to the model a cosmological constant Λ, a fixed scalar, has been added, ex-model, to the Einstein equation :$Ric_{\alpha\beta}-\frac{1}{2}g_{\alpha\beta}(R+\Lambda)=0$. The cosmological constant Λ acts as a pressure, positive or negative, to impact the expansion of the Universe. Its existence and value have been a hot topic, but it is nowadays generally acknowledged that, at least for cosmological models, it should be non null. Even if the purpose and the formulation are different, this is the same issue as above. An equilibrium means that there is no evolution of the field, and this is contradictory with the concept of a field which propagates, all the more so if the system is the whole universe.

The point of view of the observer

An observer does not follow the integral curves of V, he crosses them. For a given observer, following an integral curve of ε_0,his world line is transversal to the hypersurfaces $W(s)$ at each time t, because $\langle V(m),\varepsilon_0(m)\rangle=c>0$, so there is a bijective correspondence between s and t, but it holds only at his own location : this is the converse of the particle case.

PARTICLE CASE		FIELDS CASE	
Particle :		**Field :**	
world line :	$\Phi_u(\tau,a)$	area :	$W(s)=\{F(m)=s\}$
proper time :	τ	phase :	s
Observer :		**Observer :**	
present :	$\Omega(t)=\{f_o(m)=t\}$	world line :	$\Phi_{\varepsilon_0}(t,x)$
Intersection :	$\tau(t)$	Intersection :	$s(t)$

The location of particles are precisely defined, but with respect to a proper time which is specific to each of them. An observer can identify and locate particles in his present time, but there is no way to tell where are all the particles. Similarly a physical field has a definite location in the universe, with respect to its phase, and a given observer can measure the part of the physical field which is in his present, but there is no way to tell where the whole physical field lies.

Because there is no bijective correspondence between s, t the last results cannot be fully used. However it sticks that :

i) $\mathcal{F}$ is a decomposable tensor : $\mathcal{F}_G(m) = \widehat{\mathcal{F}}_G(m) \otimes X_G(m) ; \mathcal{F}_A(m) = \widehat{\mathcal{F}}_A(m) \otimes X_A(m)$

ii) $\pounds_V \widehat{\mathcal{F}} = 0$ and the previous solutions holds on the integral curves of V

One aspect of propagation which is of interest is the spatial curve along which the field seems to propagate. Let us consider the integral curve of V passing through $n = \varphi_o(0, x)$, with x fixed. V is a vector transversal to $\Omega_3(t)$ for any t because :

$\langle V(\Phi_V(s,n)), \varepsilon_0(\Phi_V(s,n))\rangle = c > 0$

Thus it crosses $\Omega_3(t)$ at a unique point $m(t)$ such that : $\Phi_V(s,n) = m(t)$ and there is a unique point $p(t) = \Phi_{\varepsilon_0}(-t, \Phi_V(s,n)) \in \Omega_3(0)$.

By continuity the set $\{p(t), t \geq 0\}$ is a continuous curve C in $\Omega_3(0)$ originating in n. Its tangent $\overrightarrow{v}(p(t))$ is given by :

$V(m(t)) = c\varepsilon_0(m(t)) + \sum_{\alpha=1}^{3} \overrightarrow{v}^{\alpha}(m(t))\, \varepsilon_\alpha(m(t))$

$\Phi'_{\varepsilon_0}(-t, \cdot)^* \overrightarrow{v}(p(t)) = \Phi'_{\varepsilon_0}(-t, m(t))\, \overrightarrow{v}(m(t))$

C is the spatial line along which the field seems to propagate. The distance between n and $p(t)$ is $\int_0^t \langle \overrightarrow{v}(p(\tau)), \overrightarrow{v}(p(\tau))\rangle d\tau$ and it increases with t because $\overrightarrow{v}(p(\tau))$ is space like. If not a straight line, according to the First Law of Optics, one could expect that C would be a geodesic in $\Omega_3(0)$. But, with the assumptions that we have made so far, this cannot be proven. Actually one cannot even exclude the possibility that C crosses itself : if $s \neq s' \Rightarrow m(t) \neq m(t') \Rightarrow t \neq t'$, $m(t) = \varphi_o(t, p(t)), m(t') = \varphi_o(t', p(t'))$, but one can have $p(t) = p(t')$.

5.4.5 Chern-Weil theory

The strength of the field is a somewhat complicated derivation of the potential, so one can expect that $\mathcal{F}$ meets some identities related to its definition. This is the case but, what is more significant, is that these properties do not depend on the connection, but on the principal bundle structure itself, which gives a specific, physical meaning on the fiber bundles structures P_G, P_U, Q. This is the topic of the Chern-Weil theory, which is quite abstract but as practical consequences (see Maths.27.4.5 and Kobayashi II p.298). It is a purely mathematical theory, which does not rely on any physical assumption.

Chern-Weil theorem

Let (V, ρ) be the representation of a Lie group G., and $I_n(V, \rho, G)$ the set of scalar n linear symmetric form $\varphi \in \mathcal{L}_{ns}(V; \mathbb{R})$ which are invariant by G :

$\forall X_1...X_n, Y_1, ..Y_n \in V, k_1, ...k_n \in \mathbb{R}, g \in G, \sigma \in \mathfrak{S}(n)$:

multilinear :

$\varphi(k_1X_1, ..., k_nX_n) = k_1...k_n\varphi(X_1, ..., X_n)$

$\varphi(X_1, , .., X_i + Y_i, ., X_n) = \varphi(X_1, , .., X_i, ., X_n) + \varphi(X_1, , .., Y_i, ., X_n)$

symmetric : $\varphi(X_1, .., X_n) = \varphi\left(X_{\sigma(1)}, .., X_{\sigma(n)}\right)$

invariant by G : $\varphi(\rho(g)X_1, .., \rho(g)X_n) = \varphi(X_1, .., X_n)$

φ reads in any basis of V as :

$\varphi(X_1, ..., X_n) = \sum_{i_1...i_n=1}^{\dim V} \varphi_{i_1...i_n} X_1^{i_1}...X_n^{i_n}$ where the coefficients $\varphi_{i_1...i_n}$ are symmetric by permutation of the indices.

$I_n(V,\rho,G)$ is a vector space, as well as $I(V,\rho,G) = \oplus_{n=0}^{\infty} I_n(V,\rho,G)$ with $I_0(V,\rho,G) = \mathbb{R}$ and can be endowed with a product with which it has the structure of a real algebra.

Any group has the representation (T_1G, Ad) on its Lie algebra thus one can consider (T_1G, Ad) and $I_n(T_1G, Ad, G)$.

For any principal bundle $P(M,G,\pi)$ the space of sections $\mathfrak{X}(P[T_1G, Ad], Ad)$ of the adjoint bundle is a representation of G. For any connection on $P(M,G,\pi)$ the strength $\mathcal{F}$ of the connection is a map $\mathcal{F} : M \to \Lambda_2(M; T_1G)$. So from $\mathcal{F}$, for any form $\varphi_n \in I_n(T_1G, Ad, G)$, one can define the 2n form $\widehat{\varphi}_n(\mathcal{F}) \in \Lambda_{2n}(M;\mathbb{R})$ by symmetrization :

$\forall u_1, ..., u_{2n} \in \mathfrak{X}(TM) :: \widehat{\varphi}_n(\mathcal{F})(u_1, ..., u_{2n})$

$= \frac{1}{(2n)!} \sum_{\sigma \in \mathfrak{S}(2n)} \epsilon(\sigma) \varphi\left(\mathcal{F}\left(u_{\sigma(1)}, u_{\sigma(2)}\right), ...\mathcal{F}\left(u_{\sigma(2n-1)}, u_{\sigma(2n)}\right)\right)$

$\mathcal{F}(u_p, u_q) = \sum_{a=1}^{\dim T_1G} \sum_{\alpha,\beta=1}^{\dim M} \mathcal{F}_{\alpha\beta}^a u_p^\alpha u_q^\beta \overrightarrow{\kappa}_a$

$\varphi_n(\kappa_1, ..., \kappa_n) = \sum_{a_1..a_n=1}^{\dim T_1G} \varphi_{i_1...i_n} \kappa_1^{a_1} ... \kappa_n^{a_n}$

$\widehat{\varphi}_n(\mathcal{F})$

$= \sum_{\beta_1...\beta_{2n}=1}^{\dim M} \left(\frac{1}{(2n)!} \sum_{a_1..a_n=1}^{\dim T_1G} \varphi_{a_1...a_n} \sum_{\sigma \in \mathfrak{S}(2n)} \epsilon(\sigma) \mathcal{F}_{\beta_{\sigma(1)}\beta_{\sigma(2)}}^{a_1} ... \mathcal{F}_{\beta_{\sigma(2n-3)}\beta_{\sigma(2n)}}^{a_n} \right) d\xi^{\beta_1} \wedge d\xi^{\beta_2} ... \wedge d\xi^{\beta_{2n}}$

For $n = 1$:

$\varphi_n(\kappa) = \sum_{a=1}^{\dim T_1G} \varphi_a \kappa^a$

$\widehat{\varphi}_1(\mathcal{F}) = \sum_{\alpha,\beta=1}^{\dim M} \sum_{a=1}^{\dim T_1G} \varphi_a \mathcal{F}_{\alpha\beta}^a d\xi^\alpha \wedge d\xi^\beta$

For $n = 2$:

$\widehat{\varphi}_2(\mathcal{F})$

$= \frac{1}{24} \sum_{\sigma \in \mathfrak{S}(4)} \sum_{a,b=1}^{\dim T_1G} \varphi_{ab} \sum_{\alpha_1\alpha_2\alpha_3\alpha_4=1}^{\dim M} \epsilon(\sigma) \mathcal{F}_{\alpha_{\sigma(1)}\sigma(2)}^a \mathcal{F}_{\alpha_{\sigma(3)}\alpha_{\sigma(4)}}^b d\xi^{\alpha_1} \wedge d\xi^{\alpha_2} \wedge d\xi^{\alpha_3} \wedge d\xi^{\alpha_4}$

Of course $\widehat{\varphi}_n(\mathcal{F}) \equiv 0$ whenever $2n > \dim M$.

The set of closed forms $\lambda \in \Lambda_n(M;\mathbb{R})$ on a manifold is an algebra with the exterior product, by taking the quotient space one gets a vector space $H^n(M)$ (the n cohomology class of M) and $H^*(M) = \oplus_{n=0}^{\dim M} H^n(M)$ is an algebra. Any form λ of $H^n(M)$ can be defined, up to a closed form, by a representative $c_n \in \Lambda_n(M;\mathbb{R})$ of $H^n(M) : d(\lambda - c_n) = 0$.

The Chern-Weil theorem tells that :

i) For any given map $\varphi_n \in I(T_1G, Ad, G)$ and any connection with strength $\mathcal{F}$ the exterior differential $d\widehat{\varphi}_n(\mathcal{F}) = 0$.

ii) For two principal connections with strengths $\mathcal{F}_1, \mathcal{F}_2$ there is some form $\lambda \in \Lambda_{2n-1}(M;\mathbb{R})$ such that $\widehat{\varphi}_n(\mathcal{F}_1 - \mathcal{F}_2) = d\lambda_n$.

iii) The map : $\chi : I(T_1G, Ad, G) \to H^*(M) :: \chi(\varphi) = [d\lambda_n]$ is a morphism of algebras.

So, whatever the connection, the 2n scalar forms $\widehat{\varphi}_n(\mathcal{F})$ are equivalent, up to a closed form. The class of cohomology to which belongs $\widehat{\varphi}_n(\mathcal{F})$, called the characteristic class of (P, φ_n), depends not on the connection on P, but on φ_n, and is specific to the structure of principal bundle P. In particular if P is trivial (it can be defined without patching open subsets of M) then the characteristic class is null : $H^0(M) \simeq \mathbb{R}^p$ where p is the number of connected components of M.

From a principal bundle one can define any vector bundle, but the converse is true : given a vector bundle one can define a principal bundle whose group is the one by which one goes from one holonomic basis to another (for the usual vector bundle on a m dimensional manifold this is just $GL(\mathbb{R}, m)$). So, because they depend only on the principal bundle structure and not on the connection, one can associate characteristic classes to any vector bundle E, which are called Chern classes, and each characteristic class of $H^{2n}(E)$ is defined by a 2n-form $c_n(E) = \widehat{\varphi}_n(\mathcal{F}) \in \Lambda_{2n}(M;\mathbb{R})$. Then the strength $\mathcal{F}$ is represented, in holonomic basis, by a matrix which is the Riemann tensor. For instance for $P_G\left[\mathbb{R}^4, \mathbf{Ad}\right]$

$R=\sum_{\{\alpha\beta\}ij}\left[\mathcal{F}_{\alpha\beta}\right]_j^i d\xi^\alpha\wedge d\xi^\beta\otimes\varepsilon_i(m)\otimes\varepsilon^j(m)$ and $\left[\mathcal{F}_{\alpha\beta}\right]=\sum_{a=1}^{6}\mathcal{F}_{\alpha\beta}^a\left[\kappa_a\right].$

The issue is then to compute the maps $\varphi\in I\left(T_1G,Ad,G\right).$ Notice that the Chern-Weil theorem assumes their existence, but maps φ_n which meet the properties above are quite special and do not necessarily exist. The usual way to look for them is through symmetric polynomials , the function $f(X)=\det\left(I-t\left[X\right]\right)$ (Kobayashi p.298), and polarization (Kolar p.218) but we will proceed here in a more direct method.

Application to M

The manifold M is 4 dimensional, so we have to consider only n forms for $n=1$ and $n=2$.

For $n=1$ the multilinear maps are just covectors $\varphi_1\in T_1G^*:\varphi_1(\kappa)=\sum_{a=1}^{\dim T_1G}\varphi_a\kappa^a$

The map Ad is represented in T_1G by a matrix, and φ_1 is invariant iff :

$\forall g\in G,X\in T_1G::\varphi_1(\kappa)=\sum_{a=1}^{\dim T_1G}\varphi_a\kappa^a=\varphi_1\left(Ad_g\kappa\right)=\sum_{a,b=1}^{\dim T_1G}\varphi_a\left[Ad_g\right]_b^a\kappa^b$

So there is no solution, except if $T_1G=\mathbb{R}$ because then $Ad_g=Id$ (the conjugation is the identity). This is the case of the EM field. Then the 2 form

$\widehat{\varphi}_1(\mathcal{F})=\sum_{\alpha,\beta=1}^{\dim M}\varphi\mathcal{F}_{\alpha\beta}d\xi^\alpha\wedge d\xi^\beta=\varphi\mathcal{F}_{EM}$

and we know that, indeed, $\mathcal{F}_{EM}=d\grave{A}_{EM}$ is a closed form because the bracket is null in $T_1U(1)$.

For $n=2$ the multilinear maps are bilinear symmetric form on T_1G

$\varphi_2(X,Y)=[X]^t\left[\varphi_2\right][Y]$

with a symmetric matrix $\left[\varphi_2\right]$. So this is a scalar product on the Lie algebra which is preserved by the adjoint map. For any Lie algebra there is such a scalar product, given by the Killing form (other scalar products can be defined by morphisms). So there is always a solution (except for the EM field because the Lie algebra is abelian), and

$\varphi_2(X,Y)=\langle X,Y\rangle_{T_1G}$

$\widehat{\varphi}_2(\mathcal{F})=\frac{1}{24}\left(\sum_{\sigma\in\mathfrak{S}(4)}\epsilon(\sigma)\left\langle\mathcal{F}_{\sigma(0)\sigma(1)},\mathcal{F}_{\sigma(2)\sigma(3)}\right\rangle\right)d\xi^0\wedge d\xi^1\wedge d\xi^2\wedge d\xi^3$

By considering all the permutations one gets :

$\widehat{\varphi}_2(\mathcal{F})=-\frac{1}{3}\left(\langle\mathcal{F}_{01},\mathcal{F}_{32}\rangle+\langle\mathcal{F}_{02},\mathcal{F}_{13}\rangle+\langle\mathcal{F}_{03},\mathcal{F}_{21}\rangle\right)d\xi^0\wedge d\xi^1\wedge d\xi^2\wedge d\xi^3$

the scalar product being : $\langle\mathcal{F}_{01},\mathcal{F}_{32}\rangle=\sum_{a,b=1}^{\dim T_1G}\varphi_{ab}\mathcal{F}_{01}^a\mathcal{F}_{r32}^b$

$\left(\langle\mathcal{F}_{01},\mathcal{F}_{32}\rangle+\langle\mathcal{F}_{02},\mathcal{F}_{13}\rangle+\langle\mathcal{F}_{03},\mathcal{F}_{21}\rangle\right)d\xi^0\wedge d\xi^1\wedge d\xi^2\wedge d\xi^3$

$=\sum_{a,b=1}^{\dim T_1G}\varphi_{ab}\left(\mathcal{F}_{01}^a\mathcal{F}_{r32}^b+\mathcal{F}_{02}^a\mathcal{F}_{r13}^b+\mathcal{F}_{03}^a\mathcal{F}_{r21}^b\right)d\xi^0\wedge d\xi^1\wedge d\xi^2\wedge d\xi^3$

$=\sum_{a,b=1}^{\dim T_1G}\varphi_{ab}\left(\left[\mathcal{F}^{ar}\right]\left[\mathcal{F}^{bw}\right]^t+\left[\mathcal{F}^{aw}\right]\left[\mathcal{F}^{br}\right]^t\right)d\xi^0\wedge d\xi^1\wedge d\xi^2\wedge d\xi^3$

$=-\sum_{a,b=1}^{\dim T_1G}\varphi_{ab}\mathcal{F}^a\wedge\mathcal{F}^b$

and because $\varphi_{ab}=\varphi_{ba}:\widehat{\varphi}_2(\mathcal{F})=0$ which sums up to, for any group and scalar product on the Lie algebra preserved by the adjoint map, the identity :

$$\langle\mathcal{F}_{01},\mathcal{F}_{32}\rangle+\langle\mathcal{F}_{02},\mathcal{F}_{13}\rangle+\langle\mathcal{F}_{03},\mathcal{F}_{21}\rangle=0 \tag{5.71}$$

The identity reads for the gravitational field, with $\langle\mathcal{F}_{G01},\mathcal{F}_{G32}\rangle=\left[\mathcal{F}_{r01}\right]^t\left[\mathcal{F}_{r32}\right]-\left[\mathcal{F}_{w01}\right]^t\left[\mathcal{F}_{w32}\right]$

$\sum_{a=1}^{3}\mathcal{F}_{r01}^a\mathcal{F}_{r32}^a+\mathcal{F}_{r02}^a\mathcal{F}_{r13}^a+\mathcal{F}_{r03}^a\mathcal{F}_{r21}^a-\mathcal{F}_{w01}^a\mathcal{F}_{w32}^a-\mathcal{F}_{w02}^a\mathcal{F}_{w13}^a-\mathcal{F}_{w03}^a\mathcal{F}_{w21}^a=0$

or :

$$Tr\left(\left[\mathcal{F}_r^r\right]^t\left[\mathcal{F}_r^w\right]-\left[\mathcal{F}_w^r\right]^t\left[\mathcal{F}_w^w\right]\right)=0 \tag{5.72}$$

and for the other fields (except EM) :
$\sum_{a=1}^m \mathcal{F}_{01}^a\mathcal{F}_{32}^a+\mathcal{F}_{02}^a\mathcal{F}_{13}^a+\mathcal{F}_{03}^a\mathcal{F}_{21}^a=0$
or :

$$Tr\left(\left[\mathcal{F}_A^r\right]^t\left[\mathcal{F}_A^w\right]\right)=0 \tag{5.73}$$

Remarks :

i) This identity holds for any connection (except the EM field), and without any assumption about M or Ω beyond that M is 4 dimensional, whenever there is a vector bundle.

ii) It is the consequence of the assumption of the existence of a vector bundle, and does not imply anything about the conditions of an equilibrium of a system.

iii) This identity comforts the decomposition of $\mathcal{F}$ in $\mathcal{F}^r,\mathcal{F}^w$.

With decomposable tensor

Using the previous decomposition :

$Tr\left(\left[\mathcal{F}_A^r\right]^t\left[\mathcal{F}_A^w\right]\right)$
$=Tr\left(\left[\widehat{\mathcal{F}}_A^r\right]^t\left[X_A\right]^t\left[X_A\right]\left[\widehat{\mathcal{F}}_A^w\right]\right)$
$=Tr\left(\left[\widehat{\mathcal{F}}_A^w\right]\left[\widehat{\mathcal{F}}_A^r\right]^t\left[X_A\right]^t\left[X_A\right]\right)$
$=\left(\left[\widehat{\mathcal{F}}_A^w\right]\left[\widehat{\mathcal{F}}_A^r\right]^t\right)Tr\left(\left[X_A\right]^t\left[X_A\right]\right)$

$$\left(\left[\widehat{\mathcal{F}}_A^w\right]\left[\widehat{\mathcal{F}}_A^r\right]^t\right)\left(\left[X_A\right]\left[X_A\right]^t\right)=0 \tag{5.74}$$

$Tr\left(\left[\mathcal{F}_r^r\right]^t\left[\mathcal{F}_r^w\right]-\left[\mathcal{F}_w^r\right]^t\left[\mathcal{F}_w^w\right]\right)$
$=Tr\left(\left[\widehat{\mathcal{F}}_G^r\right]^t\left[X_r\right]^t\left[X_r\right]\left[\widehat{\mathcal{F}}_G^w\right]-\left[\widehat{\mathcal{F}}_G^r\right]^t\left[X_w\right]^t\left[X_w\right]\left[\widehat{\mathcal{F}}_G^w\right]\right)$
$=\left(\left[\widehat{\mathcal{F}}_G^w\right]\left[\widehat{\mathcal{F}}_G^r\right]^t\right)Tr\left[X_r\right]^t\left[X_r\right]-\left(\left[\widehat{\mathcal{F}}_G^w\right]\left[\widehat{\mathcal{F}}_G^r\right]^t\right)Tr\left[X_w\right]^t\left[X_w\right]$

$$\left(\left[\widehat{\mathcal{F}}_G^w\right]\left[\widehat{\mathcal{F}}_G^r\right]^t\right)\left(\left[X_r\right]^t\left[X_r\right]\right)-\left(\left[X_w\right]^t\left[X_w\right]\right)=0 \tag{5.75}$$

As we have :
$\left\langle\mathcal{F},\mathcal{F}\right\rangle_G=\frac{1}{\det P'}\left(\left[*\widehat{\mathcal{F}}_G^w\right]\left[\widehat{\mathcal{F}}_G^r\right]^t+\left[*\widehat{\mathcal{F}}_G^r\right]\left[\widehat{\mathcal{F}}_G^w\right]^t\right)\left(\left[X_r\right]^t\left[X_r\right]-\left[X_w\right]^t\left[X_w\right]\right)$
$\left\langle\mathcal{F}_A,\mathcal{F}_A\right\rangle_A=\frac{1}{\det P'}\left(\left[X_A\right]^t\left[X_A\right]\right)\left(\left(\left[*\widehat{\mathcal{F}}_A^w\right]\left[\widehat{\mathcal{F}}_A^r\right]^t\right)+\left(\left[*\widehat{\mathcal{F}}_A^r\right]\left[\widehat{\mathcal{F}}_A^w\right]^t\right)\right)$
necessarily :

$$\left[\widehat{\mathcal{F}}_A^w\right]\left[\widehat{\mathcal{F}}_A^r\right]^t=0;\left[\widehat{\mathcal{F}}_G^w\right]\left[\widehat{\mathcal{F}}_G^r\right]^t=0 \tag{5.76}$$

We can then give a more precise solution of $\pounds_V\widehat{\mathcal{F}}=0$.
Along an integral curve of V :
$\left[\frac{d}{ds}\widehat{\mathcal{F}}^w\right]=\left[\widehat{\mathcal{F}}^r\right]j\left(\partial_0V\right)-\left[\widehat{\mathcal{F}}^w\right]\left(\partial_0V^0+\left[\partial v\right]\right)$
$\left[\frac{d}{ds}\widehat{\mathcal{F}}^r\right]^t=\left(\left[\partial v\right]-\left(div\left(v\right)\right)I_3\right)\left[\widehat{\mathcal{F}}^r\right]^t+j\left(\left[\partial V^0\right]\right)\left[\widehat{\mathcal{F}}^w\right]^t$
$\left[\frac{d}{ds}\widehat{\mathcal{F}}^w\right]\left[\widehat{\mathcal{F}}^r\right]^t=\left[\widehat{\mathcal{F}}^r\right]j\left(\partial_0V\right)\left[\widehat{\mathcal{F}}^r\right]^t-\left[\widehat{\mathcal{F}}^w\right]\left(\partial_0V^0+\left[\partial v\right]\right)\left[\widehat{\mathcal{F}}^r\right]^t$
$\left[\widehat{\mathcal{F}}^w\right]\left[\frac{d}{ds}\widehat{\mathcal{F}}^r\right]^t=\left[\widehat{\mathcal{F}}^w\right]\left(\left[\partial v\right]-\left(div\left(v\right)\right)I_3\right)\left[\widehat{\mathcal{F}}^r\right]^t+\left[\widehat{\mathcal{F}}^w\right]j\left(\left[\partial V^0\right]\right)\left[\widehat{\mathcal{F}}^w\right]^t$
Thus :
$\frac{d}{ds}\left(\left[\widehat{\mathcal{F}}^w\right]\left[\widehat{\mathcal{F}}^r\right]^t\right)=0$
$=\left[\widehat{\mathcal{F}}^r\right]j\left(\partial_0V\right)\left[\widehat{\mathcal{F}}^r\right]^t-\left[\widehat{\mathcal{F}}^w\right]\left(\partial_0V^0+\left[\partial v\right]\right)\left[\widehat{\mathcal{F}}^r\right]^t+\left[\widehat{\mathcal{F}}^w\right]\left(\left[\partial v\right]-\left(div\left(v\right)\right)I_3\right)\left[\widehat{\mathcal{F}}^r\right]^t+\left[\widehat{\mathcal{F}}^w\right]j\left(\left[\partial V^0\right]\right)\left[\widehat{\mathcal{F}}^w\right]^t$
$=\left[\widehat{\mathcal{F}}^r\right]j\left(\partial_0V\right)\left[\widehat{\mathcal{F}}^r\right]^t-\left(\partial_0V^0\right)\left[\widehat{\mathcal{F}}^w\right]\left[\widehat{\mathcal{F}}^r\right]^t-\left[\widehat{\mathcal{F}}^w\right]\left[\partial v\right]\left[\widehat{\mathcal{F}}^r\right]^t+\left[\widehat{\mathcal{F}}^w\right]\left[\partial v\right]\left[\widehat{\mathcal{F}}^r\right]^t-\left(div\left(v\right)\right)\left[\widehat{\mathcal{F}}^w\right]\left[\widehat{\mathcal{F}}^r\right]^t+\left[\widehat{\mathcal{F}}^w\right]j\left(\left[\partial V^0\right]\right)\left[\widehat{\mathcal{F}}^w\right]^t$
So, for the solutions, whenever $\left[\widehat{\mathcal{F}}^w\right]\left[\widehat{\mathcal{F}}^r\right]^t=0$
$\Rightarrow$

$$\left[\widehat{\mathcal{F}}^r\right]j\left(\partial_0V\right)\left[\widehat{\mathcal{F}}^r\right]^t+\left[\widehat{\mathcal{F}}^w\right]j\left(\partial V^0\right)\left[\widehat{\mathcal{F}}^w\right]^t=0 \tag{5.77}$$

Chapter 6

THE PRINCIPLE OF LEAST ACTION

In this chapter we will introduce the main tools and review the issues in continuous models, in the more general picture, that is including interactions.

The Principle of Least Action states that for any system there is some quantity (the action) which is stationary when the system is at its equilibrium. It does not tell anything about the physical content of this quantity. However, in almost all its applications, the quantity is some representation of the total energy of the system, or more precisely of the energy which is exchanged between the physical objects in the system. In an equilibrium the total balance should be null.

For a system which is represented by quantitative variables and their derivatives, defined in the context of a vector bundle $E(M, V, \pi)$ the involved quantity can be expressed in a general way as a functional, that is a map :

$\ell : \mathfrak{X}(J^r E) \to \mathbb{R}$

which acts on sections Z of the r-jet prolongation of E, represented by their coordinates $Z = \left(z^i_{\alpha_1 \ldots \alpha_s}, i = 1...n, s = 0, ..., r\right)$. And the stationary is understood as a local extremum : around the equilibrium Z_0 it is impossible to change $\ell(Z_0)$ by an infinitesimal variation δZ of Z. It is natural to assume that there is some function $L : \mathfrak{X}(J^r E) \to \mathbb{C}$ over which ℓ acts and, if this action is linear, then the functional can be expressed as an integral (Maths.2342), in this case:

ℓ

$= \int_\Omega \mathcal{L}(Z(m))\, d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$= \int_\Omega L(Z(m)) \det P' d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$= \int_\Omega L(Z(m))\, \varpi_4$

L is then a real scalar function called the **scalar lagrangian**.

In Mathematics a lagrangian is usually defined as a $\dim M$ form on M, and their theory, as well as the Variational Calculus, is extensive but also rather technical (see Maths.34.2). When only vector bundles are considered, as this is usually the case in Physics, the problems are more simple and one can use an extension of the theory of distributions (or generalized functions) which is similar to the functional derivatives well known by the physicists. We will see its application in the next chapter, and for the time being, we will study the specification of the scalar lagrangian L.

This is a crucial step, because L sums up much of the Physics of the model. The specification of a lagrangian is an art in itself and many variants have been proposed. The Standard Model is built around a complicated lagrangian (see Wikipedia “Standard Model” for its expression) which is the result of many attempts and patches to find a solution which fits the results of experiments. It is useful to remind, at this step, that one of the criteria in the choice and validation of a scientific theory is efficiency. Physicists must be demanding about their basic concepts, upon which everything is built, but, as they proceed to more specific problems, they can relax a bit. There is no Theory or a unique Model of Everything, which would be suited to all problems. The framework that we have exposed provides several tools, which can be selected according to the problems at hand. So we continue in the same spirit, and, fortunately, in the choice of the right lagrangian there are logical rules, coming essentially from the Principle of Relativity : the solution should be equivariant in a change of observer, which entails that the lagrangian itself, which is a scalar function, should be invariant. This condition provides strong guidelines in its specification, that we will see now. The methods that we expose are general, but as we have done so far, they are more easily understood when implemented on an example, and we will use the variables and representations which have been developed in the previous chapters.

6.1 THE SPECIFICATION OF THE LAGRANGIAN

6.1.1 General issues

Which variables ?

We have to decide which are the variables that enter the lagrangian and the order of their derivatives. The lagrangian is a function

$L\left(z^i_{\alpha_1...\alpha_s}, i = 1...n, s = 0, ..., r\right)$ and this is under this form that a solution is found. However usually the variables appear as composite expressions : for instance the derivatives of the potential $\partial_\alpha \grave{A}$ appear not as such but in the strength $\mathcal{F}$. So this is always a two steps process, from the composite variables to the variables and their derivatives.

We will limit ourselves to the variables which have been introduced previously, as they give a comprehensive picture of the problems. The key variables are : ψ for the state of the particle, considering that the particle belongs to some matter field, G the potential of the gravitational field, $\grave{A}$ the potential of the other fields, and the tetrad P which, in the fiber bundle model is a variable as the others (here it is not necessary to distinguish $[Q]$ and $[C_q]$) and defines the metric g. We can add a density μ, then the measure with respect to which the integral is computed is $\mu\varpi_4$.

All these variables are maps defined on a bounded area Ω of M, and valued in various vector bundles, so expressed in components in the relevant holonomic frames. The use of the formalism of fiber bundle enables us to study the most general problem with 4 variables only, and the consideration of matter fields (which are no more than general definition of motion and states) to address the issue of the trajectories and of rotation. Of course, whenever only gravitation is involved, ψ is reduced to the spinor S, but we will keep the most general framework.

The model is based on first order derivatives : the covariant derivative is at its core, and this is a first order operator. The strength $\mathcal{F}$ is of first order with respect to the potentials. So in the lagrangian it is legitimate to stay at : $\partial_\alpha \psi, \partial_\alpha G_\beta, \partial_\alpha \grave{A}_\beta, \partial_\alpha P$.

Equivariance and Covariance

An equilibrium, in the meaning of the Principle of Least Action, is a specific state of the system, which does not refer to a specific observer : an equilibrium for an observer should be also an equilibrium for another observer. So, even if the variables which are used in the model refer to measures taken by a specific observer, the conditions which are met should hold, up to a classic change of variable, for any other observer. So the lagrangian and the solutions should be, not invariant, by equivariant in a change of observer. The equilibrium is not expressed by the same figures, but it is still an equilibrium and one can go from one set of data to another by using mathematical relations deduced from the respective disposition of the observers.

In any model based on manifolds (and I remind that an affine space is a manifold, so this applies also in Galilean Geometry) a lagrangian, as any other mathematical relation, should stay the same in a change of chart. This condition is usually called covariance.

In a model based on fiber bundles there is an additional condition : the expressions must change according to the rules in a change of gauge. This condition is usually called equivariance, but it has the same meaning.

Covariance and equivariance are expressed as conditions that any quantity, and of course the lagrangian, must meet. These conditions are also a way to deal with the uncertainty which comes for the choice of some variables. For instance the orthonormal basis $(\varepsilon_i(m))$ is defined (and the tetrad with it) up to a $SO(3,1)$ matrix. The equivariance relations account for this fact.

Equivariance is usually expressed as Noether's currents (from the Mathematician Emmy Noether) and presented as the consequence of symmetries in the model. Of course if there are additional, physical symmetries, they can be accounted for in the same way. But the Noether's currents are the genuine expression of the freedom of gauge.

Once we have checked that our lagrangian (and more generally any quantity) is compliant with equivariance and covariance, of course we can exercise our freedom of gauge by choosing one specific gauge. This is how Gauge Freedom is usually introduced in Physics (in Electromagnetism we have the Gauss gauge, the Coulomb gauge,...). The goal is to simplify an expression by imposing some relations between variables. This is legitimate but, as noticed before, one must be aware that it has practical implications on the observer himself who must actually use this gauge in the collection of his data.

Time

The Principle of Locality leads naturally to express all quantities related to particles with respect to their proper time. But, whenever the propagation of the fields or several particles are considered, the state of the system must be related to a unique time, which is the time of an observer (who is arbitrary). This is necessary to have a common definition of the area of integration in the action.

The proper time of a particle and the time of the observer are related. The basic relations are (with the notations used previously) :

between the proper time τ of a particle and the time t of an observer :

$\frac{d\tau}{dt} = \sqrt{1 - \frac{\left\|\overrightarrow{v}\right\|^2}{c^2}} = \frac{1}{c}\sqrt{-\langle V, V\rangle} = \frac{c}{u^0} = \frac{1}{2a_w^2 - 1}$

between the velocity u of a particle and the speed V as measured by an observer :

$u = \frac{dp}{d\tau} = V\frac{c}{\sqrt{-\langle V,V\rangle}} = \frac{1}{\sqrt{1 - \frac{\left\|\overrightarrow{v}\right\|^2}{c^2}}}\left(\overrightarrow{v} + c\varepsilon_0\left(m\right)\right)$

Whenever particles are represented as matter fields these relations can be fully expressed with the σ_w component of the spinor.

The distinction between proper time and time of the observer, by breaking the fundamental symmetry, requires to distinguish the relativist spin (represented by $\epsilon = \pm 1$). It is assumed that its value is a continuous variable : a change along the world line is a discontinuous process,

The distinction between proper time and time of the observer is usually ignored in QTF, in spite of its obvious significance. Some attempts have been made to confront this issue, which is linked, in Quantum Physics, to the speed of propagation of the perturbation of a wave function (see Schnaid).

Fundamental state

The assumption of the existence of a fundamental state ψ_0 is at the core of the theory. For elementary particles it is fixed, and given by the type of the particle. For a composite body, the assumption that the state can be represented by a element of $E \otimes F$ implies that there is some way to define a fundamental state : there are internal forces which keep the cohesion of the body, and the definition of the body itself provides a way to compute (in a separate model) some fundamental state.

For a solid body the computation given previously provides :

- for a rigid solid a constant inertial spinor S_0, which is simply the basis for a spinor $S\left(t\right) = \gamma C\left(\sigma\left(t\right)\right) S_0$ along the trajectory

- for a deformable solid an inertial spinor S_B which varies along the world line : $S_B\left(\tau\left(t\right)\right) = N_B\left(\tau\left(t\right)\right)\gamma C\left(\sigma_B\left(\tau\left(t\right)\right)\right) S_0$ so that $S\left(t\right) = \gamma C\left(\sigma\left(t\right)\right) S_B\left(\tau\left(t\right)\right)$.

Moreover the introduction of a density (assimilated to a number of particles by unit of volume) provides some flexibility.

For the fields there is no equivalent, however, because for the fields the vacuum exist almost everywhere, the "normal state" of a field is that which it takes when it propagates. So the conditions of its propagations play an important role.

Partial derivatives and covariant derivatives

To implement the rules of Variational Calculus the partial derivatives $\partial_\alpha \psi$,

$\partial_\alpha \grave{A}, \partial_\alpha G, ...$ are required. However other quantities, and notably the lagrangian, can be expressed in using the covariant derivative ∇ or the strength $\mathcal{F}$, which have a more physical meaning. The question is then : is it legitimate to express $\partial_\alpha G, \partial_\alpha \grave{A}$ only through the strength $\mathcal{F}$, and $\partial_\alpha \psi$ through the covariant derivative ? And we will see that the answer is definitively positive.

However there is an issue with the use of the covariant derivative. The interaction of the force fields with particles goes through $\nabla_\alpha \psi$ but the covariant derivative is a 1 form, and we need to choose a vector along which to take this derivative.

Because particles move along world lines, the most natural choice is to take as vector the velocity u of the particle, or, using the rules above, the speed V. The resulting quantities are proportional :

$\nabla_V \psi = \frac{\sqrt{-\langle V,V \rangle}}{c} \nabla_u \psi$

The difference between the two quantities has a physical meaning : $\nabla_V \psi$ is the covariant derivative as measured by the observer, $\nabla_u \psi$ is an intrinsic quantity, which does not depend on the observer. However we will use the chart of an observer, with its time t, and u is related to σ_w in the standard basis of an observer, thus the choice of $\nabla_V \psi$ is more legitimate.

The value of the potentials along the trajectory have been denoted

$\widehat{G} = \sum_{\alpha=0}^{3} V^\alpha G_\alpha$

$\widehat{\grave{A}} = \sum_{\alpha=0}^{3} V^\alpha \grave{A}_\alpha$

V is a variable of the model, it is related to the motion that we should compute, for a given ψ_0 and for an observer by :

$[\psi] = [\gamma C (\sigma_w \cdot \sigma_r)] [\psi_0] [\varrho(\varkappa)]$

$V = \frac{dp}{dt} = u \frac{\sqrt{-\langle V,V \rangle}}{c} = \overrightarrow{v} + c\varepsilon_0 (m) = c \left(\varepsilon_0 + \epsilon \frac{a_w}{2a_w^2 - 1} \sum_{i=1}^{3} w_i \varepsilon_i \right)$

The choice of the key variables depends on the problem at hand :

i) Whenever the trajectory is known (for instance for bonded particles) V and then u, σ_w are known. But usually σ_r is not fixed and is a variable of the problem.

ii) For particles studied individually the trajectories are given (with respect to ε_0 for an observer) by a map : $\sigma : [0,T] \to Spin(3,1)$ which is a variable. There is a "state equation" which is an ODE in ψ with $\widehat{G}, \widehat{\grave{A}}$ as parameters, and key variables maps $r(t), w(t)$ on one hand, and "fields equations" which relate the value of the strength $\mathcal{F}_G, \mathcal{F}_A$ to the motion of the particles, on the other hand. The combination of these equations and their fitting with the initial conditions provide the trajectories.

iii) For particles represented by matter fields, the trajectories are given by a map : $\sigma : \Omega \to Spin(3,1)$ which is a variable. Usually there is also a density, with the continuity equation. The solutions are computed as in ii).

In QTF the solution which is commonly chosen is different, this is the Dirac's operator, celebrated because it is mathematically clever, but has serious drawbacks.

Dirac operator

The Dirac operator is a differential operator, and no longer a 1-form on M, defined from the covariant derivative, which does not require the choice of a vector : so it "absorbs" the α of the covariant derivative. Actually this is required in the Standard Model because the world lines are not explicit, but the Dirac's operator can be defined in a very large context (Maths.32.28), including GR, and in our formalism its meaning is more obvious. The general Dirac operator D is weakly elliptic, and D^2 is a scalar operator (Maths.2495,2496).

The mechanism is the following :

i) using the isomorphism between TM and the dual bundle TM^* provided by the metric g, to each covector $\omega = \sum_{\alpha=0}^3 \omega_\alpha d\xi^\alpha$ one can associate a vector : $\omega^* = \sum_{\alpha\beta=0}^3 g^{\alpha\beta}\omega_\alpha \partial\xi_\beta$

ii) vectors $v = \sum_{\alpha=0}^3 v^\alpha \partial\xi^\alpha$ of TM can be seen as elements of the Clifford bundle $Cl(M)$ and as such acts on $\mathbf{e}_p(m) \otimes \mathbf{f}_q(m)$ by :

$v = \sum_{\alpha j=0}^3 v^\alpha P_\alpha'^j \varepsilon_j(m)$ in the orthogonal frame

ε_j acts on $\mathbf{e}_p(m) \otimes \mathbf{f}_q(m)$ by γC :

$(\mathbf{q}(m), \mathbf{e}_p(m) \otimes \mathbf{f}_q(m))$

$\rightarrow \left(\mathbf{q}(m), \sum_{\alpha j=0}^3 v^\alpha P_\alpha'^j ([\gamma C(\varepsilon_j)] \mathbf{e}_p(m)) \otimes \mathbf{f}_q(m)\right)$

iii) thus there is an action of TM^* on $\mathbf{e}_p(m) \otimes \mathbf{f}_q(m)$ with $v = \omega^*$

$(\mathbf{q}(m), \gamma C(\omega^*)\psi(m))$

$= \left(\mathbf{p}(m), \sum_{\alpha\beta j=0}^3 g^{\alpha\beta}\omega_\alpha P_\beta'^j ([\gamma C(\varepsilon_j)] \mathbf{e}_p(m)) \otimes \mathbf{f}_q(m)\right)$

and as the tetrad defines the metric g :

$\sum_\beta g^{\alpha\beta} P_\beta'^j = \sum_{\beta kl} \eta^{kl} P_k^\alpha P_l^\beta P_\beta'^j = \sum_k \eta^{kj} P_k^\alpha$

$\sum_{\alpha\beta j=0}^3 g^{\alpha\beta}\omega_\alpha P_\beta'^j [\gamma C(\varepsilon_j)] \mathbf{e}_p(m) \otimes \mathbf{f}_q(m) = \sum_{\alpha\beta=0}^3 g^{\alpha\beta}\omega_\alpha [\gamma C(\partial\xi_\beta)] \mathbf{e}_p(m) \otimes \mathbf{f}_q(m)$

$= \sum_{\alpha=0}^3 \omega_\alpha [\gamma C(d\xi^\alpha)] \mathbf{e}_p(m) \otimes \mathbf{f}_q(m)$

iv) the covariant derivative is a one form on M so one can take $\varpi = \nabla_\alpha$ and the Dirac operator is :

$$D : \mathfrak{X}\left(J^1 Q[E \otimes F, \vartheta]\right) \rightarrow \mathfrak{X}\left(J^1 Q[E \otimes F, \vartheta]\right) :: D\psi = \sum_{\alpha=0}^3 [\gamma C(d\xi^\alpha)] [\nabla_\alpha \psi] \tag{6.1}$$

$D\psi = \sum_{\alpha=0}^3 [P]_i^\alpha \left[\gamma C\left(\varepsilon^i\right)\right] [\nabla_\alpha \psi]$

$\varepsilon^i(\varepsilon_j) = \delta_j^i \Rightarrow \gamma C\left(\varepsilon^i\right) = \gamma C(\varepsilon_i)^{-1}$

$D\psi = \sum_{\alpha=0}^3 [P]_a^\alpha [\gamma C(\varepsilon_a)] [\nabla_\alpha \psi]$

So the Dirac operator can be seen as the trace of the covariant operator, which averages the action of the covariant derivative along the directions $\alpha = 0...3$ which are put on the same footing This is mathematically convenient, and consistent with the notion of undifferentiated matter field, but has no real physical justification : it is clear that one direction is privileged on the world line.

$\langle\psi, \nabla_\alpha \psi\rangle = i \operatorname{Im}\langle\psi, \nabla_\alpha \psi\rangle$ which is convenient to define the energy of the particle in the system. But the Dirac's operator exchanges the chirality. The scalar product $\langle\psi, D\psi\rangle$ is not necessarily a real quantity and, with the matrices γ used in QTF, can be null, which is one of the reasons for the introduction of the Higgs boson (see Schücker).

Hamiltonian

In Classic Mechanics the time t is totally independent from the other geometric coordinates, so the most natural formulation of the Principle of Least Action takes the form (Maths.2606) :

$\ell(Z) = \int_0^T L\left(t, q^i, y^i\right) dt$

where y^i stands for $\frac{dq^i}{dt}$ in the 1-jet formalism, and the change of variable with the conjugate momenta :

$p^i = \frac{\partial L}{\partial q^i}$

$H = \sum_{i=1}^n p^i y^i - L$

leads to the Hamilton equations :

$\frac{dq^i}{dt} = \frac{\partial H}{\partial p^i}; \frac{dp^i}{dt} = -\frac{\partial H}{\partial q^i}$

which are the translation of the Euler-Lagrange equations with the new variables.

In QM the operator in the Schrödinger equation is assumed to be the Hamiltonian : $i\hbar\frac{\partial\psi}{\partial t} = H\psi$ and this has been an issue at the origin of Quantum Physics, because of the specific role played by the time, which seemed to be inconsistent with the covariance required by Relativity. After many attempts it has lead to the path integral formalism, which uses the lagrangian and is viewed as compatible both with Relativity and QM.

However, even if in a relativist lagrangian the coordinates are masked by a chart, it is not true that the coordinate time is banalized. To study, in a consistent manner, any system, we need a single time, and this is necessarily the time of an observer. We have to check that the formulation of the lagrangian is consistent with the Principle of Relativity : the equilibrium must be an equilibrium for any observer, but the definition of the system itself is observer-dependant. This is obvious with the folliation : the geometric area Ω of the Universe encompassed by the system during its evolution is not the same as the one of another observer. The covariance must be assured in any change of chart which respects this folliation, but that does not mean that the time itself is not specific. The Hamiltonian formulation is certainly not appropriate in the relativist context, but for many other reasons (for instance the Maxwell's equations, and more generally the concept of fields are not compatible with the Galilean Geometry) than the distinction of a privileged time.

Internal and external interactions

In the implementation of the Principle of Least Action the variables are assumed to be free, and this condition is required in the usual methods for the computation of a solution [1]. However they can appear as parameters, whose value is given, for instance if the trajectories of particles are known. In the case of fields, whose values are additive, we can have a known external field which adds up to the field generated by the particles of the system. The Principle applies to the total field, internal + external, considered as a free variable. In the usual case the field generated by the particles is neglected, and the fields variables are then totally dropped. If not the field generated by the particle is computed by substraction of the external field from the computed value.

Similarly if the observer is subjected to a specific movement, such as the rotation of his basis with respect to a chart, this motion must be accounted for the tetrad.

[1] However there are computational methods to find a solution under constraints. But the physical meaning of the Principle itself is clear : the underlying physical laws are such that the system reaches an equilibrium, in the scope of the freedom that it is left.

Conservation of Momentum and Energy

In a system comprised of interacting particles and field, each object has its own momentum and energy, which is not necessarily conserved. So to define a conservation of momentum and energy of the system we should find a way to aggregate these components, and it would be with respect to an observer who defines the time of the system.

Usually the lagrangian is built around the energy (or more precisely the exchange of energy between the components of the system), which is scalar function L and can be summed with respect to a volume form ϖ_4. The energy of the system can then be defined as :

$\mathcal{E} = \int_\Omega L\left(z^i, z^i_\alpha\right)\varpi_4$

and the conservation of energy for the observer means that [2]:

$\mathcal{E}(t) = \int_{\Omega(t)} L\left(z^i, z^i_\alpha\right)\varpi_3 = Ct = \int_{\Omega(t)} i_{\varepsilon_0}\left(L\left(z^i, z^i_\alpha\right)\varpi_4\right)$

Consider the manifold $\Omega([t_1,t_2])$ with borders $\Omega(t_1), \Omega(t_2)$:

$\mathcal{E}(t_2) - \mathcal{E}(t_1) = \int_{\partial\Omega([t_1,t_2])} i_{\varepsilon_0}(L\varpi_4) = \int_{\Omega([t_1,t_2])} d\left(i_{\varepsilon_0}L\varpi_4\right)$

$d\left(i_{\varepsilon_0}L\varpi_4\right) = \pounds_{\varepsilon_0}(L\varpi_4) - i_{\varepsilon_0}d(L\varpi_4)$

$= \left(\pounds_{\varepsilon_0}L\right)\varpi_4 + L\pounds_{\varepsilon_0}\varpi_4 - i_{\varepsilon_0}(dL\wedge\varpi_4) - i_{\varepsilon_0}Ld\varpi_4$

$= L'(\varepsilon_0)\varpi_4 + L(div\varepsilon_0)\varpi_4 - i_{\varepsilon_0}(dL\wedge\varpi_4)$

$= div(L\varepsilon_0)\varpi_4$

$\mathcal{E}(t_2) - \mathcal{E}(t_1) = \int_{\Omega([t_1,t_2])} div(L\varepsilon_0)\varpi_4$

(Maths.1517,1587)

The conservation of energy for the observer imposes a specific condition : $div(L\varepsilon_0) = 0$ which is not necessarily met for a stationary solution.

The momenta are vectorial quantities, defined in different vector spaces, and at different points, so their aggregation has no meaning and they are not involved, as such, in the implementation of the principle of least action.

6.1.2 Specification of a General Lagrangian

We will use the precise notation :

L denotes the scalar lagrangian $L\left(z^i, z^i_\alpha\right)$ function of the variables z^i, expressed by the components in the gauge of the observer, and their partial derivatives which, in the jets bundle formalism, are considered as independent variables z^i_α.

$\mathcal{L} = L\left(z^i, z^i_\alpha\right)(\det P')$

$L\varpi_4 = L\left(z^i, z^i_\alpha\right)(\det P')\, d\xi^0\wedge d\xi^1\wedge d\xi^2\wedge d\xi^3$ is the 4-form

$\frac{\partial\mathcal{L}}{\partial z}$ to denote the usual partial derivative with respect to the variable z

$\frac{d\mathcal{L}}{dz}$ to denote the total derivative with respect to the variable z, meaning accounting for the composite expressions in which it is an argument.

We will illustrate how to compute the rules of equivariance and covariance for a general lagrangian, using the variables that we have defined previously, expressed by their coordinates : $\psi^{ij}, G^a_\alpha, \grave{A}^a_\alpha, P^\alpha_i, \partial_\beta\psi^{ij}, \partial_\beta G^a_\alpha, \partial_\beta\grave{A}^a_\alpha, \partial_\beta P^\alpha_i, V^\alpha$.

So in this section :

$L\left(\psi^{ij}, G^a_\alpha, \grave{A}^a_\alpha, P^\alpha_i, \partial_\beta\psi^{ij}, \partial_\beta G^a_\alpha, \partial_\beta\grave{A}^a_\alpha, \partial_\beta P^\alpha_i, V^\alpha\right)$ in an action such as : $\int_\Omega L\mu\det P' d\xi^0\wedge d\xi^1\wedge d\xi^2\wedge d\xi^3$

All variables are represented by their coordinates in relevant bases, by real or complex scalars. L is not supposed to be holomorphic, so the real and imaginary part of the variables $\psi^{ij}, \partial_\alpha\psi^{ij}$

[2]Notice the difference with a similar computation done for material bodies : material bodies are characterized by a unique vector field V, but in a general system the unique reference is ε_0.

must appear explicitly. We will use the convenient notation for complex variables z and their conjugates $\overline{z}$, by introducing the holomorphic complex valued functions :

$$\frac{\partial L}{\partial z}=\frac{1}{2}\left(\frac{\partial L}{\partial \operatorname{Re} z}+\frac{1}{i}\frac{\partial L}{\partial \operatorname{Im} z}\right);\frac{\partial L}{\partial \overline{z}}=\frac{1}{2}\left(\frac{\partial L}{\partial \operatorname{Re} z}-\frac{1}{i}\frac{\partial L}{\partial \operatorname{Im} z}\right) \tag{6.2}$$

$\Leftrightarrow$
$\frac{\partial L}{\partial \operatorname{Re} z}=\frac{\partial l}{\partial z}+\frac{\partial L}{\partial \overline{z}}$
$\frac{\partial L}{\partial \operatorname{Im} z}=i\left(\frac{\partial L}{\partial z}-\frac{\partial L}{\partial \overline{z}}\right)$

The partial derivatives $\frac{\partial L}{\partial \operatorname{Re} z},\frac{\partial L}{\partial \operatorname{Im} z}$ are real valued functions, so $\frac{\partial L}{\partial \overline{z}}=\overline{\frac{\partial L}{\partial z}}$. And we have the identities for any complex valued function u :

$$\frac{\partial L}{\partial \operatorname{Re} z}\operatorname{Re} u+\frac{\partial L}{\partial \operatorname{Im} z}\operatorname{Im} u=2\operatorname{Re}\frac{\partial L}{\partial z}u;-\frac{\partial L}{\partial \operatorname{Re} z}\operatorname{Im} u+\frac{\partial L}{\partial \operatorname{Im} z}\operatorname{Re} u=-2\operatorname{Im}\frac{\partial L}{\partial z}u \tag{6.3}$$

To find a solution we need the explicit presence of the variables and their partial derivatives. But as our goal is to precise the specification of L, we can, without loss of generality, make the replacements :

$\partial_\alpha \psi^{ij}\rightarrow \nabla_\alpha \psi^{ij}=\partial_\alpha \psi^{ij}+\sum_{k=1}^{4}\sum_{a=1}^{6}\left[\gamma C\left(G_\alpha^a\right)\right]_k^i\psi^{kj}+\sum_{k=1}^{n}\psi^{ik}\left[\grave{A}_\alpha\right]_j^k$
$\partial_\beta G_\alpha^a\rightarrow \mathcal{F}_{G\alpha\beta}^a=\partial_\alpha G_\beta^a-\partial_\beta G_\alpha^a+2\left[G_\alpha,G_\beta\right]^a$ and $F_{G\alpha\beta}=\partial_\alpha G_\beta^a+\partial_\beta G_\alpha^a$
$\partial_\beta \grave{A}_\alpha^a\rightarrow \mathcal{F}_{A\alpha\beta}^a=\partial_\alpha \grave{A}_\beta^a-\partial_\beta \grave{A}_\alpha^a+2\left[\grave{A}_\alpha,\grave{A}_\beta\right]^a$ and $F_{G\alpha\beta}=\partial_\alpha \grave{A}_\beta^a+\partial_\beta \grave{A}_\alpha^a$

And the lagrangian is then a function :
$\mathcal{L}\left(\psi^{ij},G_\alpha^a,\grave{A}_\alpha^a,P_i^\alpha,\nabla_\alpha\psi^{ij},\mathcal{F}_{G\alpha\beta},F_{G\alpha\beta}^a,\mathcal{F}_{A\alpha\beta},F_{A\alpha\beta},\partial_\beta P_i^\alpha,u^\alpha\right)$

Most of the variables above are defined up to some transformation : for instance the components of the tetrad are defined up to a matrix of $SO(3,1)$. The function L should be intrinsic, meaning invariant by :

- a change of gauge in the principal bundles P_G,P_U and their associated bundles
- a change of chart in the manifold M

The operations below will give the relations which must exist between the variables and the partial derivatives of L, and some precious information about the presence or the absence of some variables.

6.1.3 Equivariance in a change of gauge

The mechanism is exposed below to find the relations between variables, but it is also at the foundation of Noether's currents.

One parameter group of change of gauge

One parameter groups of change of trivialization on a principal bundle are defined by sections of their adjoint bundle (Maths.2070) :

$\kappa\in\mathfrak{X}\left(P_G\left[T_1Spin(3,1),\mathbf{Ad}\right]\right)$
$\theta\in\mathfrak{X}\left(P_U\left[T_1U,Ad\right]\right)$

$\kappa = \upsilon(\kappa_r, \kappa_w), \theta$ are maps from M to the Lie algebras. At each point m, for a given value of a scalar parameter τ, the exponential on the Lie algebra defines an element of the groups at m (Maths.22.2.6) :

$\exp : \mathbb{R} \times T_1 Spin(3,1) \to Spin(3,1) :: \exp(\tau\kappa(m))$

$\exp : \mathbb{R} \times T_1 U \to U :: \exp(\tau\theta(m))$

The exponential on $T_1 Spin(3,1)$ is expressed by (see Annex) :

$\exp t\kappa = \exp \tau\upsilon(\kappa_r, \kappa_w) = \sigma_w(\tau, \kappa_w) \cdot \sigma_r(\tau, \kappa_r)$

$\sigma_w(\tau, \kappa_w) = a_w(\tau, \kappa_w) + \sinh \frac{1}{2}\tau\sqrt{\kappa_w^t \kappa_w}\upsilon(0, \kappa_w)$

$a_w(\tau, \kappa_w) = \sqrt{1 + \frac{1}{4}\left(\kappa_w^t \kappa_w \sinh^2 \frac{1}{2}\tau\sqrt{\kappa_w^t \kappa_w}\right)}$

$\sigma_r(\tau, \kappa_r) = a_r(\tau, \kappa_w) + \sin t\frac{1}{2}\sqrt{\kappa_r^t \kappa_r}\upsilon(\kappa_r, 0)$

$a_r(\tau, \kappa_w) = \sqrt{1 - \frac{1}{4}\kappa_r^r \kappa_r \sin^2 t\frac{1}{2}\sqrt{\kappa_r^t \kappa_r}}$

It is actually multivalued (because of the double cover) so we assume that one of the value has been chosen (for instance $a > 0$). This does not matter here.

By definition the derivative of these exponential for $\tau = 0$ gives back the elements of the Lie algebras :

$\frac{d}{d\tau}\exp(\tau\kappa(m))|_{\tau=0} = \kappa(m)$

$\frac{d}{d\tau}\exp(\tau\theta(m))|_{\tau=0} = \theta(m)$

With the change of gauge :

$\mathbf{p}_G(m) \to \widetilde{\mathbf{p}}_G(m, \tau) \cdot \exp(-\tau\kappa(m))$

$\mathbf{p}_U(m) \to \widetilde{\mathbf{p}}_U(m) \cdot \exp(-\tau\theta(m))$

The components of the variables become :

$P_i^\alpha \to \widetilde{P}_i^\alpha(m, \tau) = \sum_{j=0}^3 [h(\exp(-\tau\kappa))]_i^j P_j^\alpha$ where $[h]$ is the $SO(3,1)$ corresponding matrix

$\psi^{ij} \to \widetilde{\psi}^{ij}(m, \tau) = \sum_{k=1}^4 \sum_{l=1}^n [\gamma C(\exp(\tau\kappa))]_k^i [\varrho(\exp(\tau\theta))]_l^j \psi^{kl}$

$G_\alpha(m) \to \widetilde{G}_\alpha(m) = \mathbf{Ad}_{\exp\tau\kappa}\left(G_\alpha - \exp(-\tau\kappa)(\exp\tau\kappa)' \tau\partial_\alpha\kappa\right)$

$\grave{A}_\alpha \to \widetilde{\grave{A}}_\alpha(m, \tau) = Ad_{\exp\tau\theta}\left(\grave{A}_\alpha - \exp(-\tau\theta)\exp(\tau\theta)' \tau\partial_\alpha\theta\right)$

$\nabla_\alpha\psi \to \widetilde{\nabla_\alpha\psi^{ij}}(m, \tau) = \sum_{k=1}^4 \sum_{l=1}^n [\gamma C(\exp(\tau\kappa))]_k^i [\varrho(\exp(\tau\theta))]_l^j \nabla_\alpha\psi^{kl}$

All these expressions depend on m, as well as $\kappa(m), \theta(m)$, so they can be differentiated with respect to the coordinates of m to get :

$\partial_\beta P_i^\alpha \to \widetilde{\partial_\beta P}_i^\alpha(m, \tau) = \sum_{j=0}^3 \left(\left[h\left(\exp(-\tau\kappa)' \partial_\beta\kappa\right)\right]_i^j P_j^\alpha + [h(\exp(-\tau\kappa))]_i^j \partial_\beta P_i^\alpha\right)$

$\partial_\beta \widetilde{G}_\alpha(m, \tau)$

$= \left[(\exp -\tau\kappa)(\exp\tau\kappa)' \tau\partial_\beta\kappa, G_\alpha - \tau\partial_\alpha\kappa\right]$

$+Ad_{\exp\tau\kappa}\{\partial_\beta G_\alpha -$

$\{(\exp -\tau\kappa)' \tau\partial_\beta\kappa \circ (\exp\tau\kappa)' \tau\partial_\alpha\kappa$

$+\exp(-\tau\kappa) \circ (\exp\tau\kappa)"(\tau\partial_\beta\kappa, \tau\partial_\alpha\kappa) + \exp(-\tau\kappa) \circ \exp(\tau\kappa)' \tau\partial_{\alpha\beta}^2\kappa\}\}$

$\partial_\beta \widetilde{\grave{A}}_\alpha(m, \tau) = \left[(\exp -\tau\theta)(\exp\tau\theta)' \tau\partial_\beta\theta, \grave{A}_\alpha - \tau\partial_\alpha\theta\right]$

$+Ad_{\exp\tau\theta}(\partial_\beta\grave{A}_\alpha - \left(\exp(-\tau\theta)' \tau\partial_\beta\theta \circ (\exp\tau\theta)' \tau\partial_\alpha\theta\right)$

$+\exp(-\tau\theta) \circ (\exp\tau\kappa)"(\tau\partial_\beta\theta, \tau\partial_\alpha\theta) + \exp(-\tau\theta) \circ \exp(\tau\theta)' \tau\partial_{\alpha\beta}^2\theta)$

$\mathcal{F}_{G\alpha\beta}^a \to \widetilde{\mathcal{F}}_{G\alpha\beta}(\tau) = Ad_{\exp\tau\kappa}\mathcal{F}_{G\alpha\beta}$

$\mathcal{F}_{A\alpha\beta}^a \to \widetilde{\mathcal{F}}_{A\alpha\beta}(\tau) = Ad_{\exp\tau\theta}\mathcal{F}_{A\alpha\beta}$

$F_{G\alpha\beta} \to Ad_{\exp\tau\kappa}F_{G\alpha\beta}$

$+\left[(\exp -\tau\kappa)(\exp\tau\kappa)' \tau\partial_\beta\kappa, G_\alpha - t\partial_\alpha\kappa\right] + \left[(\exp -\tau\kappa)(\exp\tau\kappa)' \tau\partial_\alpha\kappa, G_\beta - \tau\partial_\beta\kappa\right]$

$-Ad_{\exp\tau\kappa}((\exp -\tau\kappa)' \tau\partial_\beta\kappa \circ (\exp\tau\kappa)' \tau\partial_\alpha\kappa + \exp(-\tau\kappa) \circ (\exp\tau\kappa)"(\tau\partial_\beta\kappa, \tau\partial_\alpha\kappa)$

$+\exp(-\tau\kappa) \circ \exp(\tau\kappa)' \tau\partial_{\alpha\beta}^2\kappa)$

$-Ad_{\exp \tau\kappa}((\exp -\tau\kappa)' \tau\partial_\alpha\kappa \circ (\exp \tau\kappa)' \tau\partial_\beta\kappa + \exp(-\tau\kappa) \circ (\exp \tau\kappa)" (\tau\partial_\alpha\kappa, \tau\partial_\beta\kappa)$
$+\exp(-\tau\kappa) \circ \exp(\tau\kappa)' \tau\partial^2_{\alpha\beta}\kappa)$
$F_{A\alpha\beta} \to Ad_{\exp(\tau\theta}F_{A\alpha\beta}$
$+\left[(\exp -\tau\theta)(\exp \tau\theta)' \tau\partial_\beta\theta, \grave{A}_\alpha - \tau\partial_\alpha\theta\right] + \left[(\exp -\tau\theta)(\exp \tau\theta)' \tau\partial_\alpha\theta, \grave{A}_\beta - \tau\partial_\beta\theta\right]$
$-Ad_{\exp \tau\theta}(\exp(-\tau\theta)' \tau\partial_\beta\theta \circ (\exp \tau\theta)' \tau\partial_\alpha\theta + \exp(-\tau\theta) \circ (\exp \tau\kappa)" (\tau\partial_\beta\theta, \tau\partial_\alpha\theta)$
$+\exp(-\tau\theta) \circ \exp(\tau\theta)' \tau\partial^2_{\alpha\beta}\theta)$
$-Ad_{\exp \tau\theta}(\exp(-\tau\theta)' \tau\partial_\alpha\theta \circ (\exp \tau\theta)' \tau\partial_\beta\theta + \exp(-\tau\theta) \circ (\exp \tau\kappa)" (\tau\partial_\alpha\theta, \tau\partial_\beta\theta)$
$+\exp(-\tau\theta) \circ \exp(\tau\theta)' \tau\partial^2_{\alpha\beta}\theta)$

The vector V is defined in the holonomic basis $\partial\xi_\alpha$ so its components are not impacted.

The determinant $\det P'$ is invariant, because we have a change of orthonormal basis, so the scalar lagrangian L is invariant :

$\forall\tau, (\kappa, \partial_\lambda\kappa, \partial_{\lambda\mu}\kappa), (\theta, \partial_\lambda\theta, \partial_{\lambda\mu}\theta)$:
$L\left(z^i, z^i_\alpha\right) = L\left(\widetilde{z}^i(\tau, \kappa, \partial_\lambda\kappa, \partial_{\lambda\mu}\kappa), \widetilde{z}^i_\alpha(\tau, \kappa, \partial_\lambda\kappa, \partial_{\lambda\mu}\kappa)\right)$
$L\left(z^i, z^i_\alpha\right) = L\left(\widetilde{z}^i(\tau, \theta, \partial_\lambda\theta, \partial_{\lambda\mu}\theta), \widetilde{z}^i_\alpha(\tau, \theta, \partial_\lambda\theta, \partial_{\lambda\mu}\theta)\right)$

If we take the derivative of this identity for $\tau = 0$:
$\frac{dL}{d\tau}|_{\tau=0} = \sum_{i,\alpha} \frac{\partial L}{\partial z^i}\left(z^i, z^i_\alpha\right) \frac{d\widetilde{z}^i}{d\tau}|_{\tau=0}$

$\frac{d\widetilde{z}^i}{d\tau}|_{\tau=0}$ depends on the value of $(\kappa, \partial_\lambda\kappa, \partial_{\lambda\mu}\kappa), (\theta, \partial_\lambda\theta, \partial_{\lambda\mu}\theta)$. So we have identities between the partial derivatives of L which must hold for any value of $(\kappa, \partial_\lambda\kappa, \partial_{\lambda\mu}\kappa), (\theta, \partial_\lambda\theta, \partial_{\lambda\mu}\theta)$. From a mathematical point of view this derivative with respect to τ is the Lie derivative of the lagrangian along the vertical vector fields generated by the derivative $\frac{dz^i_\alpha}{d\tau}|_{\tau=0}$ for each variable. These vector fields are the **Noether currents** (Maths.34.3.4). Here we will not explicit these currents, but simply deduce some compatibilities between the partial derivatives.

Moreover the formulas : $z^i \to \widetilde{z}^i$ can also be written : $\widetilde{z}^i(z^p, \kappa, \partial_\lambda\kappa, \partial_{\lambda\mu}\kappa), ...$ and we have :
$L\left(z^i, z^i_\alpha\right) = \widetilde{L}\left(\widetilde{z}^i, \widetilde{z}^i_\alpha\right) = \widetilde{L}\left(\widetilde{z}^i(z^p), \widetilde{z}^i_\alpha\left(z^j_p\right)\right)$

thus by taking the derivative with respect to the variables $\left(z^i, z^i_\alpha\right)$ at $\tau = 0$ we get identities between the values of the partial derivatives $\Pi^i = \frac{\partial L}{\partial z^i}\left(z^i, z^i_\alpha\right)$ and $\widetilde{\Pi}^i = \frac{\partial \widetilde{L}}{\partial \widetilde{z}^i}\left(z^i, z^i_\alpha\right)$ which tells if they transform as tensors.

Equivariance on P_G

The computation for $\exp(\tau\kappa(m))$ gives :

$\frac{d}{d\tau}\widetilde{P}^\alpha_i(m,\tau)|_{\tau=0} = -\sum_a \kappa^a([P][\kappa_a])^\alpha_i$
$\frac{d}{d\tau}\operatorname{Re}\widetilde{\psi}^{ij}(m,\tau)|_{\tau=0}$
$= \sum_a \kappa^a \sum_{k=1}^4 \left(\operatorname{Re}\left([\gamma C(\kappa_a)]^i_k\right)\operatorname{Re}\psi^{kj} - \operatorname{Im}\left([\gamma C(\kappa_a)]^i_k\right)\operatorname{Im}\psi^{kj}\right)$
$= \sum_a \kappa^a \operatorname{Re}([\gamma C(\kappa_a)][\psi])^i_j$
$\frac{d}{d\tau}\operatorname{Im}\widetilde{\psi^{ij}}(m,\tau)|_{\tau=0}$
$= \sum_a \kappa^a \sum_{k=1}^4 \left(\operatorname{Re}\left([\gamma C(\kappa_a)]^i_k\right)\operatorname{Im}\psi^{kj} + \operatorname{Im}\left([\gamma C(\kappa_a)]^i_k\right)\operatorname{Re}\psi^{kj}\right)$
$= \sum_a \kappa^a \operatorname{Im}([\gamma C(\kappa_a)][\psi])^i_j$
$\frac{d}{d\tau}\partial_\beta\widetilde{P}(m,t)^\alpha_j|_{\tau=0} = -\sum_a \kappa^a([\partial_\beta P][\kappa_a])^\alpha_i + \partial_\beta\kappa^a([P][\kappa_a])^\alpha_i$
$\frac{d}{d\tau}\operatorname{Re}\widetilde{\nabla_\alpha\psi}^{ij}(m,\tau)|_{\tau=0}$
$= \sum_a \kappa^a \sum_{k=1}^4 \left(\operatorname{Re}\left([\gamma C(\kappa_a)]^i_k\right)\operatorname{Re}\nabla_\alpha\psi^{kj} - \operatorname{Im}\left([\gamma C(\kappa_a)]^i_k\right)\operatorname{Im}\nabla_\alpha\psi^{kj}\right)$
$= \sum_a \kappa^a \operatorname{Re}([\gamma C(\kappa_a)][\nabla_\alpha\psi])^i_j$

$\frac{d}{d\tau}\operatorname{Im}\widetilde{\nabla_\alpha \psi^{ij}}(m,\tau)|_{\tau=0}$
$=\sum_a \kappa^a \sum_{k=1}^4 \left(\operatorname{Re}\left([\gamma C(\kappa_a)]_k^i\right)\operatorname{Im}\nabla_\alpha\psi^{kj}+\operatorname{Im}\left([\gamma C(\kappa_a)]_k^i\right)\operatorname{Re}\nabla_\alpha\psi^{kj}\right)$
$=\sum_a \kappa^a \operatorname{Im}\left([\gamma C(\kappa_a)][\nabla_\alpha\psi]\right)_j^i$
$\frac{d}{d\tau}\widetilde{G_\alpha^a}(m)|_{\tau=0}=\sum_b \kappa^b\left[\overrightarrow{\kappa}_b,G_\alpha\right]^a-\partial_\alpha\kappa^a$
$\frac{d}{d\tau}\partial_\beta\widetilde{G_\alpha^a}(m,\tau)|_{\tau=0}=\sum_b \kappa^b\left[\overrightarrow{\kappa}_b,\partial_\beta G_\alpha\right]^a+\partial_\beta\kappa^b\left[\overrightarrow{\kappa}_b,G_\alpha\right]^a-\partial_{\alpha\beta}\kappa^a$
$\frac{d}{d\tau}\widetilde{\mathcal{F}}_{G\alpha\beta}^a(\tau)|_{\tau=0}=\sum_b \kappa^b\left[\overrightarrow{\kappa}_b,\mathcal{F}_{G\alpha\beta}\right]^a$
$\frac{d}{d\tau}\widetilde{F}_{G\alpha\beta}^a|_{\tau=0}=\sum_b \kappa^b\left[\overrightarrow{\kappa}_b,F_{G\alpha\beta}\right]^a+\partial_\beta\kappa^b\left[\overrightarrow{\kappa}_b,G_\alpha\right]^a+\partial_\alpha\kappa^b\left[\overrightarrow{\kappa}_b,G_\beta\right]^a-2\partial_{\alpha\beta}\kappa^a$

So we have the identity :
$\forall\kappa_a,\partial_\beta\kappa^a,\partial_{\alpha\beta}\kappa^a$:
$0=$
$\sum_a\kappa^a\{\sum_{ij}\frac{\partial L}{\partial\operatorname{Re}\psi^{ij}}\operatorname{Re}\left([\gamma C(\kappa_a)][\psi]\right)_j^i+\frac{\partial L}{\partial\operatorname{Im}\psi^{ij}}\operatorname{Im}\left([\gamma C(\kappa_a)][\psi]\right)_j^i$
$+\sum_{\alpha ij}\frac{\partial L}{\partial\operatorname{Re}\nabla_\alpha\psi^{ij}}\operatorname{Re}\left([\gamma C(\kappa_a)][\nabla_\alpha\psi]\right)_j^i+\frac{\partial L}{\partial\operatorname{Im}\nabla_\alpha\psi^{ij}}\operatorname{Im}\left([\gamma C(\kappa_a)][\nabla_\alpha\psi]\right)_j^i\}$
$+\sum_{i\alpha}\frac{\partial L}{\partial P_i^\alpha}\left(-\sum_a\kappa^a\left([P][\kappa_a]\right)_i^\alpha\right)$
$+\sum_{i\alpha\beta}\frac{\partial\mathcal{L}}{\partial\partial_\beta P_i^\alpha}\left(-\sum_a\kappa^a\left([\partial_\beta P][\kappa_a]\right)_j^\alpha+\partial_\beta\kappa^a\left([P][\kappa_a]\right)_j^\alpha\right)$
$+\sum_{a\alpha}\frac{\partial L}{\partial G_\alpha^a}\left(\sum_b\kappa^b\left[\overrightarrow{\kappa}_b,G_\alpha\right]^a-\partial_\alpha\kappa^a\right)$
$+\sum_{a\alpha\beta}\frac{\partial L}{\partial\mathcal{F}_{G\alpha\beta}^a}\left(\sum_b\kappa^b\left[\overrightarrow{\kappa}_b,\mathcal{F}_{G\alpha\beta}\right]^a\right)$
$+\frac{\partial L}{\partial F_{G\alpha\beta}^a}\left(\sum_b\kappa^b\left[\overrightarrow{\kappa}_b,F_{G\alpha\beta}\right]^a+\partial_\beta\kappa^b\left[\overrightarrow{\kappa}_b,G_\alpha\right]^a+\partial_\alpha\kappa^b\left[\overrightarrow{\kappa}_b,G_\beta\right]^a-2\partial_{\alpha\beta}\kappa^a\right)$

With the component in $\partial_{\alpha\beta}\kappa^a$ we have immediately : $\forall a,\alpha,\beta:\frac{\partial L}{\partial F_{G\alpha\beta}^a}=0$
With the component in $\partial_\alpha\kappa^a:\forall a,\alpha:\sum_{\beta i}\frac{\partial L}{\partial\partial_\alpha P_i^\beta}\left([P][\kappa_a]\right)_i^\beta=-\frac{\partial L}{\partial G_\alpha^a}$

And we are left with :
$\forall a=1..6$:
$0=$
$\sum_{ij}\frac{\partial L}{\partial\psi^{ij}}\left([\gamma C(\kappa_a)][\psi]\right)_j^i+\sum_{\alpha ij}\frac{\partial L}{\partial\nabla_\alpha\psi^{ij}}\left([\gamma C(\kappa_a)][\nabla_\alpha\psi]\right)_j^i$
$-\sum_{i\alpha}\frac{\partial L}{\partial P_i^\alpha}\left([P][\kappa_a]\right)_i^\alpha-\sum_{i\alpha\beta}\frac{\partial\mathcal{L}}{\partial\partial_\beta P_i^\alpha}\left([\partial_\beta P][\kappa_a]\right)_j^\alpha$
$+\sum_{b\alpha}\frac{\partial L}{\partial G_\alpha^b}\left[\overrightarrow{\kappa}_a,G_\alpha\right]^b+\sum_{a\alpha\beta}\frac{\partial L}{\partial\mathcal{F}_{G\alpha\beta}^b}\left[\overrightarrow{\kappa}_a,\mathcal{F}_{G\alpha\beta}\right]^b$

Moreover, by taking the derivative with respect to the initial variables we get :
$\sum_{k=1}^4\left[\gamma C(\exp(\tau\kappa(m)))\right]_i^k\frac{\partial\widetilde{L}}{\partial\widetilde{\psi^{kj}}}=\frac{\partial L}{\partial\psi^{ij}}$
$\sum_{k=1}^4\left[\gamma C(\exp(\tau\kappa(m)))\right]_i^k\frac{\partial\widetilde{L}}{\partial\widetilde{\nabla_\alpha\psi^{kj}}}=\frac{\partial L}{\partial\nabla\psi^{ij}}$
$\sum_j\left[h(\exp(-\tau\kappa(m)))\right]_i^j\frac{\partial\widetilde{L}}{\partial P_j^\alpha}=\frac{\partial L}{\partial P_i^\alpha}$
$L\left(\left[Ad_{\exp\tau\kappa}\right]_b^a\mathcal{F}_{G\alpha\beta}^b\right)=L\left(\mathcal{F}_{G\alpha\beta}\right)$
$\sum_b\left[Ad_{\exp\tau\kappa}\right]_a^b\frac{\partial\widetilde{L}}{\partial\widetilde{\mathcal{F}_{G\alpha\beta}^b}}=\frac{\partial L}{\partial\mathcal{F}_{G\alpha\beta}^a}$
and other similar identities, which show that the partial derivatives are tensors, with respect to the dual vector bundles :
$\sum_i\frac{\partial L}{\partial\psi^{ij}}\mathbf{e}^i,\sum_i\frac{\partial L}{\partial\nabla_\alpha\psi^{ij}}\mathbf{e}^i,\frac{\partial L}{\partial\mathcal{F}_{G\alpha\beta}^a}\overrightarrow{\kappa}^a,\sum_i\frac{\partial L}{\partial\partial_\beta P_i^\alpha}\varepsilon^i$ with $\overrightarrow{\kappa}^a$ the basis vector of the dual of $T_1Spin(3,1)$: $\overrightarrow{\kappa}^a\left(\overrightarrow{\kappa}_b\right)=\delta_b^a$.

Equivariance on P_U

We have similarly :

$\frac{d}{d\tau}\widetilde{\psi^{ij}}(m,\tau)|_{\tau=0} = \sum_{k=1}^{n}\sum_{a=1}^{m}\theta^a\psi^{ik}[\theta_a]_j^k$
$\frac{d}{d\tau}\operatorname{Re}\widetilde{\psi^{ij}}(m,\tau)|_{\tau=0} = \sum_{a=1}^{m}\theta^a\operatorname{Re}(\psi[\theta_a])^{ij}$
$\frac{d}{d\tau}\operatorname{Im}\widetilde{\psi^{ij}}(m,\tau)|_{\tau=0} = \sum_{a=1}^{m}\theta^a\operatorname{Im}(\psi[\theta_a])^{ij}$
$\frac{d}{d\tau}\widetilde{\grave{A}}_\alpha^a(m,\tau)|_{\tau=0} = \sum_{b=1}^{m}\theta^b\left[\overrightarrow{\theta}_b,\grave{A}_\alpha\right]^a - \partial_\alpha\theta^a$
$\frac{d}{d\tau}\operatorname{Re}\widetilde{\nabla_\alpha\psi^{ij}}(m,\tau)|_{\tau=0} = \sum_{a=1}^{m}\theta^a\operatorname{Re}(\nabla_\alpha\psi[\theta_a])^{ij}$
$\frac{d}{d\tau}\operatorname{Im}\widetilde{\nabla_\alpha\psi^{ij}}(m,\tau)|_{\tau=0} = \sum_{a=1}^{m}\theta^a\operatorname{Im}(\nabla_\alpha\psi[\theta_a])^{ij}$
$\frac{d}{d\tau}\partial_\beta\widetilde{\grave{A}}_\alpha^a(m,\tau)|_{\tau=0} = \sum_{b=1}^{m}\theta^b\left[\theta_b,\partial_\beta\grave{A}_\alpha\right]^a + \partial_\beta\theta^b\left[\theta_b,\grave{A}_\alpha\right]^a - \partial_{\alpha\beta}\theta^a$
$\frac{d}{d\tau}\ \widetilde{\mathcal{F}}_{A\alpha\beta}(\tau)|_{\tau=0} = \sum_{b=1}^{m}\theta^b\left[\overrightarrow{\theta}_b,\mathcal{F}_{A\alpha\beta}\right]^a$
$\frac{d}{d\tau}\widetilde{F}_{A\alpha\beta}|_{\tau=0} = \sum_{b=1}^{m}\theta^b[\theta_b,F_{A\alpha\beta}]^a + \partial_\beta\theta^b\left[\theta_b,\grave{A}_\alpha\right]^a + +\partial_\alpha\theta^b\left[\theta_b,\grave{A}_\beta\right]^a - 2\partial_{\alpha\beta}\theta^a$

$\sum_{ij}\frac{\partial L}{\partial\operatorname{Re}\psi^{ij}}\sum_{a=1}^{m}\theta^a\operatorname{Re}(\psi[\theta_a])^{ij} + \frac{\partial L}{\partial\operatorname{Im}\psi^{ij}}\sum_{a=1}^{m}\theta^a\operatorname{Im}(\psi[\theta_a])^{ij}$
$+\sum_{ij\alpha}\frac{\partial L}{\partial\operatorname{Re}\nabla_\alpha\psi^{ij}}\sum_{a=1}^{m}\theta^a\operatorname{Re}(\psi[\theta_a])^{ij} + \sum_{ij\alpha}\frac{\partial L}{\partial\operatorname{Im}\nabla_\alpha\psi^{ij}}\sum_{a=1}^{m}\theta^a\operatorname{Im}(\psi[\theta_a])^{ij}$
$+\sum_{a\alpha}\frac{\partial L}{\partial\grave{A}_\alpha^a}\left(\sum_{b=1}^{m}\theta^b\left[\overrightarrow{\theta}_b,\grave{A}_\alpha\right]^a - \partial_\alpha\theta^a\right) + \sum_{a\alpha\beta}\frac{\partial L}{\partial\mathcal{F}_{A\alpha\beta}^a}\left(\sum_{b=1}^{m}\theta^b\left[\overrightarrow{\theta}_b,\mathcal{F}_{A\alpha\beta}\right]^a\right)$
$+\sum_{a\alpha\beta}\frac{\partial L}{\partial F_{A\alpha\beta}^a}\left(\sum_{b=1}^{m}\theta^b[\theta_b,F_{A\alpha\beta}]^a + \partial_\beta\theta^b\left[\theta_b,\grave{A}_\alpha\right]^a + \partial_\alpha\theta^b\left[\theta_b,\grave{A}_\beta\right]^a - 2\partial_{\alpha\beta}\theta^a\right)$
$=0$

Which implies :
$\forall a,\alpha,\beta : \frac{\partial L}{\partial F_{A\alpha\beta}^a} = 0, \frac{\partial L}{\partial\grave{A}_\alpha^a} = 0$
$\forall a = 1..m$:
$\sum_{ij}\frac{\partial L}{\partial\psi^{ij}}(\psi[\theta_a])^{ij} + \sum_{ij\alpha}\frac{\partial L}{\partial\nabla_\alpha\psi^{ij}}(\nabla_\alpha\psi[\theta_a])^{ij} + \sum_{b\alpha\beta}\frac{\partial L}{\partial\mathcal{F}_{A\alpha\beta}^b}\left(\left[\overrightarrow{\theta}_a,\mathcal{F}_{A\alpha\beta}\right]^b\right) = 0$

By taking the derivative with respect to the initial variables we check that the partial derivatives are tensors, with respect to the dual vector bundles : $\sum_i\frac{\partial L}{\partial\psi^{ij}}\mathbf{f}^j, \sum_i\frac{\partial L}{\partial\nabla_\alpha\psi^{ij}}\mathbf{f}^j, \sum_a\frac{\partial L}{\partial\mathcal{F}_{A\alpha\beta}^a}\overrightarrow{\theta}^a$ with $\overrightarrow{\theta}^a$ the basis vector of the dual of T_1U : $\overrightarrow{\theta}^a\left(\overrightarrow{\theta}_b\right) = \delta_b^a$

6.1.4 Covariance

In a change of charts on M with the jacobian : $J = \left[J_\beta^\alpha\right] = \left[\frac{\partial\widetilde{\xi}^\alpha}{\partial\xi^\beta}\right]$ and $K = J^{-1}$ the 4-form on M which defines the action changes as :
$L\mu\det[P]\,d\xi^0\wedge d\xi^1\wedge d\xi^2\wedge d\xi^3 = \widetilde{L}\widetilde{\mu}\det\left[\widetilde{P}\right]d\widetilde{\xi}^0\wedge d\widetilde{\xi}^1\wedge d\widetilde{\xi}^2\wedge d\widetilde{\xi}^3$
and because :
$\widetilde{\mu}\det\left[\widetilde{P}\right]d\widetilde{\xi}^0\wedge d\widetilde{\xi}^1\wedge d\widetilde{\xi}^2\wedge d\widetilde{\xi}^3 = \mu\det[P]\,d\xi^0\wedge d\xi^1\wedge d\xi^2\wedge d\xi^3$
the scalar lagrangian L should be invariant.

The variables change as (Maths.16.1.2) :
ψ^{ij} do not change

The covariant derivatives are one form :
$\nabla_\alpha \psi^{ij} \to \widetilde{\nabla_\alpha \psi^{ij}} = \sum_\beta K_\alpha^\beta \nabla_\beta \psi^{ij}$
P, V are vectors, but their components are functions :
$V^\alpha \to \widetilde{V}^\alpha = \sum_\gamma J_\gamma^\alpha u^\gamma$
$P_i^\alpha \to \widetilde{P}_i^\alpha = \sum_\gamma J_\gamma^\alpha P_i^\gamma$
$\widetilde{\partial_\beta P_i^\alpha} = \frac{\partial}{\partial \widetilde{\xi}^\beta} \left(\sum_\gamma J_\gamma^\alpha (\xi) P_i^\gamma (\xi) \right) = \sum_\gamma \left(\frac{\partial}{\partial \widetilde{\xi}^\beta} J_\gamma^\alpha (\xi) \right) P_i^\gamma (\xi) + J_\gamma^\alpha (\xi) \frac{\partial}{\partial \widetilde{\xi}^\beta} P_i^\gamma (\xi)$
$\widetilde{\partial_\beta P_i^\alpha} = \sum_{\gamma\eta} \left(\partial_\eta J_\gamma^\alpha \right) K_\beta^\eta P_i^\gamma + \left(\left(\partial_\eta P_i^\gamma \right) J_\gamma^\alpha K_\beta^\eta \right)$
The potentials are 1-form :
$G_\alpha^a \to \widetilde{G}_\alpha^a = \sum_\beta K_\alpha^\beta G_\beta^a$
$\grave{A}_\alpha^a \to \widetilde{\grave{A}}_\alpha^a = \sum_\beta K_\alpha^\beta \grave{A}_\beta^a$
The strengths of the fields are 2-forms. They change as :
$\mathcal{F}_{G\alpha\beta}^a \to \widetilde{\mathcal{F}}_{G\alpha\beta}^a = \sum_{\{\gamma\eta\}=0}^3 \mathcal{F}_{G\gamma\eta}^a \det [K]_{\{\alpha\beta\}}^{\{\gamma\eta\}} = \sum_{\gamma\eta=0}^3 \mathcal{F}_{G\gamma\eta}^a K_\alpha^\gamma K_\beta^\eta$

So we have the identity :
$L \left(z^i, z_\alpha^i, z_{\alpha\beta}^i \right) = \widetilde{L} \left(\widetilde{z}^i, \widetilde{z}_\alpha^i, \widetilde{z}_{\alpha\beta}^i \right)$
$= \widetilde{L} \left(\widetilde{z}^i \left(z_\lambda^i, J_\mu^\lambda \right), \widetilde{z}_\alpha^i \left(z_\lambda^i, J_\mu^\lambda, \partial_\gamma J_\mu^\lambda \right), \widetilde{z}_{\alpha\beta}^i \left(z_\lambda^i, J_\mu^\lambda, \partial_\gamma J_\mu^\lambda, \partial_{\gamma\varepsilon}^2 J_\mu^\lambda \right) \right).$
In a first step we take the derivative with respect to the components of the Jacobian.
If we take the derivative of this identity with respect to $\left(\partial_\eta J_\mu^\lambda \right)$:
$0 = \sum_{i\alpha\beta} \frac{\partial L}{\partial \partial_\beta P_i^\alpha} \sum_{\gamma\eta} K_\beta^\eta P_i^\gamma \delta_\lambda^\alpha \delta_\gamma^\mu = \sum_{\alpha\beta\eta i} \frac{\partial L}{\partial \partial_\beta P_i^\lambda} P_i^\mu K_\beta^\eta$
take $J_\mu^\lambda = \delta_\mu^\lambda \Rightarrow K_\beta^\eta = \delta_\beta^\eta$
$\sum_i \frac{\partial L}{\partial \partial_\eta P_i^\lambda} P_i^\mu = 0$
$\forall \alpha, \beta, \gamma : \sum_i \frac{\partial L}{\partial \partial_\alpha P_i^\beta} P_i^\gamma = 0$
by product with $P_\gamma^{j\prime}$ and summation : $\forall \alpha, \beta, j : \frac{\partial L}{\partial \partial_\alpha P_j^\beta} = 0$
and as we had :
$\forall a, \alpha : \sum_{\beta i} \frac{\partial L}{\partial \partial_\alpha P_i^\beta} \left([P] [\kappa_a] \right)_i^\beta = -\frac{\partial L}{\partial G_\alpha^a} \Rightarrow \forall a, \alpha : \frac{\partial L}{\partial G_\alpha^a} = 0$

The derivative with respect to J_μ^λ :
$\sum_{i\alpha} \frac{\partial L}{\partial P_i^\alpha} \sum_\gamma P_i^\gamma \delta_\alpha^\lambda \delta_\gamma^\mu + \sum_{i\alpha} \frac{\partial L}{\partial \operatorname{Re} \nabla_\alpha \psi^{ij}} \sum_\beta \left(\frac{\partial}{\partial J_\mu^\lambda} K_\alpha^\beta \right) \operatorname{Re} \nabla_\beta \psi^{ij}$
$+ \sum_{i\alpha} \frac{\partial L}{\partial \operatorname{Im} \nabla_\alpha \psi^{ij}} \sum_\beta \left(\frac{\partial}{\partial J_\mu^\lambda} K_\alpha^\beta \right) \operatorname{Im} \nabla_\beta \psi^{ij}$
$+ \sum_{a\alpha\beta} \frac{\partial L}{\partial \mathcal{F}_{G\alpha\beta}^a} \sum_{\gamma\eta} \left(\left(\frac{\partial}{\partial J_\mu^\lambda} K_\alpha^\gamma \right) K_\beta^\eta + K_\alpha^\gamma \frac{\partial}{\partial J_\mu^\lambda} K_\beta^\eta \right) \mathcal{F}_{G\gamma\eta}^a$
$+ \frac{\partial L}{\partial \mathcal{F}_{A\alpha\beta}^a} \sum_{\gamma\eta} \left(\left(\frac{\partial}{\partial J_\mu^\lambda} K_\alpha^\gamma \right) K_\beta^\eta + K_\alpha^\gamma \frac{\partial}{\partial J_\mu^\lambda} K_\beta^\eta \right) \mathcal{F}_{A\gamma\eta}^a + \frac{\partial L}{\partial V^\alpha} \sum_\gamma V^\gamma \delta_\alpha^\lambda \delta_\gamma^\mu = 0$
with $\frac{\partial}{\partial J_\mu^\lambda} K_\alpha^\beta = -K_\lambda^\beta K_\alpha^\mu$
$\sum_{i\alpha} \frac{\partial L}{\partial P_i^\lambda} P_i^\mu + \sum_{ij\alpha} \frac{\partial L}{\partial \nabla_\alpha \psi^{ij}} \sum_\beta \left(-K_\lambda^\beta K_\alpha^\mu \right) \nabla_\beta \psi^{ij}$
$+ \sum_{a\alpha\beta} \frac{\partial L}{\partial \mathcal{F}_{G\alpha\beta}^a} \sum_{\gamma\eta} \left(\left(\left(-K_\lambda^\gamma K_\alpha^\mu \right) \right) K_\beta^\eta + K_\alpha^\gamma \left(-K_\lambda^\eta K_\beta^\mu \right) \right) \mathcal{F}_{G\gamma\eta}^a$
$+ \frac{\partial L}{\partial \mathcal{F}_{A\alpha\beta}^a} \sum_{\gamma\eta} \left(\left(\left(-K_\lambda^\gamma K_\alpha^\mu \right) \right) K_\beta^\eta + K_\alpha^\gamma \left(-K_\lambda^\eta K_\beta^\mu \right) \right) \mathcal{F}_{A\gamma\eta}^a + \frac{\partial L}{\partial V^\lambda} V^\mu = 0$
Let us take $J_\mu^\lambda = \delta_\mu^\lambda \Rightarrow K_\mu^\lambda = \delta_\mu^\lambda$
$\sum_i \frac{\partial L}{\partial P_i^\lambda} P_i^\mu - \sum_{i\alpha} \frac{\partial L}{\partial \nabla_\mu \psi^{ij}} \nabla_\lambda \psi^{ij} - \sum_{a\eta} \frac{\partial L}{\partial \mathcal{F}_{G\mu\eta}^a} \mathcal{F}_{G\lambda\eta}^a - \sum_{a\gamma} \frac{\partial L}{\partial \mathcal{F}_{G\gamma\mu}^a} \mathcal{F}_{G\gamma\lambda}^a - \sum_{a\eta} \frac{\partial L}{\partial \mathcal{F}_{A\mu\eta}^a} \mathcal{F}_{A\lambda\eta}^a - \sum_{a\gamma} \frac{\partial L}{\partial \mathcal{F}_{A\gamma\mu}^a} \mathcal{F}_{A\gamma\lambda}^a + \frac{\partial L}{\partial V^\lambda} V^\mu = 0$

that is :
$\forall \alpha, \beta : \sum_{ij} \frac{\partial L}{\partial \nabla_\beta \psi^{ij}} \nabla_\alpha \psi^{ij} + \sum_{a\gamma} \frac{\partial L}{\partial \mathcal{F}^a_{G\beta\gamma}} \mathcal{F}^a_{G\alpha\gamma} + \frac{\partial L}{\partial \mathcal{F}^a_{A\beta\gamma}} \mathcal{F}^a_{A\alpha\gamma} = \sum_i \frac{\partial L}{\partial P^\alpha_i} P^\beta_i + \frac{\partial L}{\partial V^\alpha} V^\beta$

In the second step we can take the derivative with respect to the initial variable in the identity :

$\widetilde{L}\left(\widetilde{P^\alpha_i}, \widetilde{\psi^{ij}}, \widetilde{\nabla_\alpha \psi^{ij}}, \widetilde{\mathcal{F}^a_{A\alpha\beta}}, \widetilde{\mathcal{F}^a_{G\alpha\beta}}, \widetilde{V}^\alpha\right)$

$= \widetilde{L}\left(\widetilde{P^\alpha_i}\left(P^\lambda_i\right), \psi^{ij}, \widetilde{\nabla_\alpha \psi^{ij}}\left(\nabla_\lambda \psi^{pq}\right), \widetilde{\mathcal{F}^a_{A\alpha\beta}}\left(\mathcal{F}^b_{A\lambda\mu}\right), \widetilde{\mathcal{F}^a_{G\alpha\beta}}\left(\mathcal{F}^b_{G\lambda\mu}\right), \widetilde{V}^\alpha\left(V^\lambda\right)\right)$

$= L\left(P^\alpha_i, \psi^{ij}, \nabla_\alpha \psi^{ij}, \mathcal{F}^a_{A\alpha\beta}, \mathcal{F}^a_{G\alpha\beta}, V^\alpha\right)$

$\frac{\partial \widetilde{L}}{\partial \widetilde{P^\alpha_i}} \frac{\partial \widetilde{P^\alpha_i}}{\partial P^\lambda_i} = \frac{\partial \widetilde{L}}{\partial \widetilde{P^\alpha_i}} J^\alpha_\lambda = \frac{\partial L}{\partial P^\lambda_i}$

$\frac{\partial \widetilde{L}}{\partial \widetilde{\nabla_\alpha \psi^{ij}}} \frac{\partial \widetilde{\nabla_\alpha \psi^{ij}}}{\partial \nabla_\lambda \psi^{ij}} = \frac{\partial \widetilde{L}}{\partial \widetilde{\nabla_\alpha \psi^{ij}}} K^\lambda_\alpha = \frac{\partial L}{\partial \nabla_\lambda \psi^{ij}}$

$\frac{\partial \widetilde{L}}{\partial \widetilde{\mathcal{F}^a_{G\alpha\beta}}} \frac{\partial \widetilde{\mathcal{F}^a_{G\alpha\beta}}}{\partial \mathcal{F}^a_{B\lambda\mu}} = \frac{\partial \widetilde{L}}{\partial \widetilde{\mathcal{F}^a_{G\alpha\beta}}} K^\lambda_\alpha K^\mu_\beta = \frac{\partial L}{\partial \mathcal{F}^a_{G\lambda\mu}}$

$\frac{\partial \widetilde{L}}{\partial \widetilde{\mathcal{F}^a_{A\alpha\beta}}} \frac{\partial \widetilde{\mathcal{F}^a_{G\alpha\beta}}}{\partial \mathcal{F}^a_{A\lambda\mu}} = \frac{\partial \widetilde{L}}{\partial \widetilde{\mathcal{F}^a_{A\alpha\beta}}} K^\lambda_\alpha K^\mu_\beta = \frac{\partial L}{\partial \mathcal{F}^a_{A\lambda\mu}}$

which shows that the corresponding quantities are tensors : in TM^* for $\frac{\partial L}{\partial P^\lambda_i}$ and in $TM \otimes TM$ for $\frac{\partial L}{\partial \nabla_\lambda \psi^{ij}}, \frac{\partial L}{\partial \mathcal{F}^a_{G\lambda\mu}}, \frac{\partial L}{\partial \mathcal{F}^a_{A\lambda\mu}}$.

6.1.5 Conclusion

i) The potentials $\grave{A}, G$, and the derivatives $\partial_\beta P^\alpha_i$ do not figure explicitly, the derivatives of the potential $\grave{A}, G$ factor in the strength.

The lagrangian is a function of 6 variables only :

$$L = L\left(\psi, \nabla_\alpha \psi, P^\alpha_i, \mathcal{F}_{G\alpha\beta}, \mathcal{F}_{A\alpha\beta}, V^\alpha\right) \tag{6.4}$$

ii) The following quantities are tensors :
$\Pi_\psi = \sum_{ij} \frac{\partial L}{\partial \psi^{ij}} \mathbf{e}^i \otimes \mathbf{f}^i$
$\Pi_\nabla = \sum_\alpha \frac{\partial L}{\partial \nabla_\alpha \psi^{ij}} \partial \xi_\alpha \otimes \mathbf{e}^i \otimes \mathbf{f}^i$
$\Pi_P = \sum_\alpha \frac{\partial L}{\partial P^\alpha_i} d\xi^\alpha \otimes \varepsilon^i$
$\Pi_A = \sum_{\alpha\beta} \frac{\partial L}{\partial \mathcal{F}^a_{A\alpha\beta}} \partial \xi_\alpha \wedge \partial \xi_\beta \otimes \overrightarrow{\theta}^a$
$\Pi_G = \sum_{\alpha\beta} \frac{\partial L}{\partial \mathcal{F}^a_{G\alpha\beta}} \partial \xi_\alpha \wedge \partial \xi_\beta \otimes \overrightarrow{\kappa}^a$
and similarly $\sum_{\alpha\beta} \upsilon^*\left(\frac{\partial L}{\partial \mathcal{F}_{r\alpha\beta}}, \frac{\partial L}{\partial \mathcal{F}_{w\alpha\beta}}\right) \partial \xi_\alpha \wedge \partial \xi_\beta$

This result is useful because we can express cumbersome equations with simpler and more flexible geometric quantities. Notice that these quantities, when $\det[P']$ is added to L, are no longer covariant.

iii) We have the identities
$\forall a = 1..6 : \Pi_\psi\left[\gamma C\left(\overrightarrow{\kappa}_a\right)\right]\psi + \sum_\alpha \Pi^\alpha_\nabla\left[\gamma C\left(\overrightarrow{\kappa}_a\right)\right]\nabla_\alpha \psi - \Pi_P[P][\kappa_a] + \sum_{b\alpha\beta} \Pi^{\alpha\beta}_{Gb}\left[\overrightarrow{\kappa}_a, \mathcal{F}_{G\alpha\beta}\right]^b = 0$
$\forall a = 1..m : \left(\Pi_\psi \psi + \sum_\alpha \Pi^\alpha_\nabla \nabla_\alpha \psi\right)[\theta_a] + \sum_{b\alpha\beta} \Pi^{\alpha\beta}_{Ab}\left[\overrightarrow{\theta}_a, \mathcal{F}_{A\alpha\beta}\right]^b = 0$
$\forall \alpha, \beta : \Pi^\beta_\nabla \nabla_\alpha \psi + \sum_{a\gamma} \Pi^{\beta\gamma}_{Ga} \mathcal{F}^a_{G\alpha\gamma} + \Pi^{\beta\gamma}_{Aa} \mathcal{F}^a_{A\alpha\gamma} - \sum_i \Pi^i_{P\alpha} P^\beta_i = \frac{\partial L}{\partial V^\alpha} V^\beta$

These identities are minimal necessary conditions for the lagrangian : the calculations could be continued to higher derivatives. They do not depend on the signature. Whenever the lagrangian is expressed with the geometrical quantities, these identities are automatically satisfied.

6.2 THE POINT PARTICLE ISSUE

A lagrangian must suit the case of particles alone, fields alone and interacting fields and particles. So it comprises a part for the fields, and another one for the particles and their interactions. If we consider a population of N particles interacting with the fields the action is :

$\int_\Omega L_1 \left(P_i^\alpha, \mathcal{F}_{G\alpha\beta}, \mathcal{F}_{A\alpha\beta}\right) \varpi_4 + \sum_{p=1}^N \int_{\tau_p^1}^{\tau_p^2} L_2 \left(\psi_p, \nabla_\alpha \psi_p, P_i^\alpha, V_p^\alpha\right) cd\tau_p$

because the induced volume form on the world line is $cd\tau_p$

And this raises several issues, mathematical and physical, depending on the system considered.

6.2.1 Propagation of Fields

If we consider a system without any particle, focus on the fields and aim at knowing their propagation in Ω, the variables are just the components of the tetrad P, and the strength of the fields $\mathcal{F}_A, \mathcal{F}_G$. We have a unique integral over Ω and the Euler-Lagrange equations give general solutions which are matched to the initial conditions. A direct and simple answer can be found [3], at least for the variation with respect to t. In the usual models the propagation at the speed of light is postulated, and introduced separately in a linearized approximation. The classic examples are, in General Relativity (with the Levi-Civita connection) the Einstein equation :

$Ric_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R = 0$

and the Maxwell equations :

$\sum_{\alpha\beta} \partial_\alpha \left(\mathcal{F}^{\alpha\beta} \sqrt{|\det P'|}\right) = 0$

with the lagrangian : $L = \sum_{\alpha\beta} G g^{\alpha\beta} Ric_{\alpha\beta} + \mu_0 \mathcal{F}_{\alpha\beta} \mathcal{F}^{\alpha\beta}$

6.2.2 Particles moving in known Fields

When the system is comprised of particles moving in known fields actually only the second part of the action is involved. Replacing the proper time of each particle with the time of the observer, using the relations above, we have a classic variational problem over the interval $[0, T]$ of the experiment.

If the fields induced by the particle are negligible we can expect a solution, but it will be at best expressed as general conditions that the trajectories must meet. The main example is the trajectory of free particles, that is particles which are not submitted to a field. With the simple lagrangian $L_1 = 1$ and the Levi-Civita connection one finds that the trajectory must be a geodesic, and there is a unique geodesic passing through any point m with a given tangent $V(0)$. But the equation does not give by itself the coordinates of the geodesic (which require the knowledge of G) or the value of the field. For the electromagnetic field, if we know the value of the field and we neglect the field induced by the particle, we get similarly a solution :

$\nabla_u u = \mu_0 \frac{q}{mc} \sum_\alpha \mathcal{F}^{\alpha\beta} u_\beta$ with $u = \frac{c}{\sqrt{-\langle V, V\rangle}} V$

If we want to account for the field induced by the particle we have a problem. As the field propagates, we need to know the field out of the trajectory. It could be computed by the more general model, and the results reintegrated in the single particle model. The resulting equation for the trajectory is known, for the electromagnetic field, as the "Lorentz-Dirac equation" (see Poisson and Quinn). The procedure is not simple, and there are doubts about the physical meaning of the equation itself.

[3] With the restrictions noted before about the meaning of the results.

6.2.3 Particles and Fields interacting

The fundamental issue is that the particles are not present everywhere, so even if we can represent the states of the particles by a matter field, that is a section of a vector bundle, we have to account for the actual presence of the particles : virtual particles do not interact [4]. There are different solutions.

Common solutions

If the trajectories of the particles are known, a direct computation gives usually the field that they induce. This is useful for particles which are bonded (such as in condensed matter).

In QTF the introduction of matter fields in the lagrangian is in part formal, as most of the computations, notably when they address the problem of the creation / annihilation of particles, is done through Feynman's diagram, which is a way to reintroduce the action at a distance between identified particles.

In the classical picture the practical solutions which have been implemented with the Principle of Least Action have many variants, but share the following assumptions :

- they assume that the particles follow some kind of continuous trajectories and keep their physical characteristics (this condition adds usually a separate constraint)

- the trajectory is the key variable, but the model gives up the concept of point particle, replaced by some form of density of particles.

These assumptions makes sense when we are close to the equilibrium, and we are concerned not by the behavior of each individual particle but by global results about distinguished populations, measured as cross sections over an hypersurface. They share many characteristics with the models used in fluid mechanics. In the usual QM interpretation the density of particles can be seen as a probability of presence, but these models are used in the classical picture, and actually the state of the particles is represented as sections of the vector bundle TM (with a constraint imposed by the mass), combined with a density function. So the density has a direct, classic interpretation.

The simplest solution is, assuming that the particles have the same physical characteristics, to take as key variable a density $\mu\varpi_4$. Then the application of the principle of least action with a 4 dimensional integral gives the equations relating the fields and the density of charge.

The classic examples are :

- the 2nd Maxwell equation in GR :

$\nabla^{\beta}\mathcal{F}_{\beta\alpha} = -\mu_0 J_\alpha \Leftrightarrow \mu_0 J^\alpha \sqrt{-\det g} = \sum_\beta \partial_\beta \left(\mathcal{F}^{\alpha\beta}\sqrt{-\det g}\right)$

with the current : $J = \mu(m)\, qu$ and the lagrangian

$L = \mu_0 \sum_\alpha \grave{A}_\alpha J^\alpha + \frac{1}{2}\sum_{\alpha\beta}\mathcal{F}_{\alpha\beta}\mathcal{F}^{\alpha\beta}$

- the Einstein Equation in GR :

$Ric_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta} = \frac{8\pi G}{\sqrt{c}}T_{\alpha\beta}$

with the momentum energy tensor $T_{\alpha\beta} = \frac{\partial T}{\partial g^{\alpha\beta}} - \frac{1}{2}g_{\alpha\beta}T$

and the lagrangian $L = T\left(g, z^i, z^i_\alpha\right) + \frac{\sqrt{c}}{8\pi G}R$

The distribution of charges is defined independently, but it must meet a conservation law. In the examples above we must have :

$\sum_\alpha \partial_\alpha J^\alpha = 0$

$\sum_\alpha \nabla^\alpha T_{\alpha\beta} = 0$

The Einstein-Vlasov systems are also based on a distribution function $f(m,p)$ depending on the localization m and the linear momentum p, which must follow a conservation law, expressed

[4]Virtual : existing or resulting in essence or effect though not in actual fact, form, or name (American Heritage Dictionary). An interacting virtual particle is an oximoron.

as a differential equation (the Vlasov equation). The particles are generally assumed to have the same mass, so there is an additional constraint on the momentum as above. When only the gravitational field is considered the particles follow geodesics, to which the conservation law is adjusted. These systems have been extensively studied for plasmas and Astrophysics (see Andréasson).

This kind of model has been adjusted to Yang-Mills fields (Choquet-Bruhat) : the particles have different physical characteristics (similar to the vector ϕ seen previously), and must follow an additional conservation law given by $\nabla_V \phi = 0$ (the Wong equation).

In all these solutions the 4 dimensional action, with a lagrangian adapted to the fields considered, gives an equation relating the field and the distribution of charges.

So the situation is not satisfying. These difficulties have physical roots. The concept of field is aimed at removing the idea of action at a distance, but, as the example of the motion of a single particle in its own field shows, it seems difficult to circumvent the direct consideration of mutual interactions between particles, which needs to identify separately each of them.

However, from these classic examples, two results seem quite clear :

- the trajectories should belong to some family of curves, defined by the interactions

- the initial conditions, that is the beginning x of the curve and its initial tangent, should determine the curve in the family.

They are consistent with our description of the motions by sections of fiber bundles and matter fields. Moreover the Spinor formalism avoids the introduction of constraints on the state (or momentum) of the particles : the conservation law is satisfied by construct. Indeed the particles keep their intrinsic properties through ψ_0. And it can deal with the two components of motion : translation and rotation.

Finite number of particles

For a finite, fixed, number of particles, with known fundamental state ψ_{0p}, the second integral of the action reads :

$$\sum_{p=1}^{N} \int_{\tau_p^1}^{\tau_p^2} L_2 \left(\left(\psi_p, \nabla_{u_p} \psi_p, P_i^{\alpha}, V_p^{\alpha} \right) \right) d\tau_p$$

$$= \sum_{p=1}^{N} \int_0^T L_2 \left(\left(\psi_p, \nabla_{u_p} \psi_p, P_i^{\alpha} \right) (t) \right) \sqrt{- \langle V_p, V_p \rangle} dt$$

The states of the particles are represented by maps : $\psi_p : [0, T] \to E \otimes F$ with a fundamental state ψ_{0p}. The key variables are then $r_p(t), w_p(t)$ which are related to V_p , and the value of the potential along the world lines $\widehat{G}, \widehat{\grave{A}}$.

There is an obvious mathematical problem : the fields and the particles are defined over domains which are manifolds of different dimensions, which precludes the usual method by Euler-Lagrange equations. It is common to put a Dirac's function in the second part, but this, naive, solution is just a formal way to rewrite the same integral without any added value.

If the model considers only a finite, fixed, number of particles, there is a rigorous mathematical solution, by functional derivatives, that we will see in the next chapter.

Density of particles

The system is a population of identical particles (they have the same fundamental state ψ_0) which follow trajectories without collisions (the trajectories do not cross) and are observed in an area Ω over the period $[0, T]$. It is then possible to represent the system by a matter field $\psi \in \mathfrak{X}(\psi_0)$ with a density μ with respect to the volume form ϖ_4.

The key variables are :

- maps $r, w \in C_1 \left(\Omega; \mathbb{R}^3 \right)$ which define a section $\sigma \in \mathfrak{X}(P_G)$, and $\varkappa \in \mathfrak{X}(P_U)$

- the density $\mu \in C(\Omega;\mathbb{R}_+) : \mu(m) = \sqrt{\frac{\langle\psi,\psi\rangle}{\langle\psi_0,\psi_0\rangle}}$

- the fields represented by the potentials (if there are external fields they are accounted for in addition to the fields generated by the particles).

The lagrangian for the interactions is then :

$\int_\Omega L_2\left(\psi, \nabla_u\psi, P_i^\alpha, V^\alpha\right)\mu\varpi_4$

With the conditions :

$\psi = \mu(m)\,\vartheta(\sigma(m), \varkappa(m))\,\psi_0$

$\mu divV + \frac{d\mu}{dt} = 0$

and the initial conditions, defined on $\Omega_3(0)$.

The variables $r(m), w(m)$ are the coordinates of the section representing the matter field. Whenever the derivatives of a variable Y are taken along the trajectory $m(t) : \sum_{\alpha=0}^{3} V^\alpha \partial_\alpha Y = \frac{dY}{dt}(m(t))$ and usually one gets differential equations where the key variables are $\frac{dr}{dt}, \frac{dw}{dt}$.

6.3 FINDING A SOLUTION

The implementation of the Principle of Least Action leads to the problem of finding sections on vector bundles for which the action is stationary. There are two general methods, depending if the action is defined by a unique integral, or by several integrals on domains of different dimensions.

6.3.1 Variational calculus with Euler-Lagrange Equations

This is the most usual problem : find a section Z for which the integral $\int_\Omega L(z^i, z^i_\alpha)\varpi_4$ is stationary. This is a classic problem of variational calculus, and the solution is given by the Euler-Lagrange equation, for each variable (Maths.34.3).

L denotes the scalar lagrangian $L\left(z^i, z^i_\alpha\right)$ function of the variables z^i, expressed by the components in the gauge of the observer, and their partial derivatives which, in the jets bundle formalism, are considered as independent variables z^i_α.

$\mathcal{L} = L\left(z^i, z^i_\alpha\right)(\det P')$

$L\varpi_4 = L\left(z^i, z^i_\alpha\right)(\det P')\, d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$ is the 4-form

$\frac{\partial \mathcal{L}}{\partial z}$ denote the usual partial derivative with respect to the variable z

$\frac{d\mathcal{L}}{dz}$ denote the total derivative with respect to the variable z, meaning accounting for the composite expressions in which it is an argument.

For an action $\int_\Omega L\left(z^i, z^i_\alpha\right)\varpi_4$where $\left(z^i, z^i_\alpha\right)$ is a 1-jet section of a vector bundle, the Euler-Lagrange equations read :

$$\forall z^i : \frac{d\left(L \det P'\right)}{dz^i} - \sum_\beta \frac{d}{d\xi^\beta}\frac{d\left(L \det P'\right)}{dz^i_\beta} = 0 \tag{6.5}$$

where $\frac{d}{d\xi^\beta}$ is the derivative with respect to the coordinates in M. $\det[P']$ is necessary to account for ϖ_4 which involves $P\prime$.

The equation holds pointwise for any $m \in \Omega$. However when one considers a point along a trajectory : $p(t) = m(\Phi_V(t,x))$ then the expressions like : $\sum_\beta V^\beta \frac{d}{d\xi^\beta}(X(p(t)))$ read : $\frac{dX}{dt}(p(t))$

The divergence of a vector field $X = \sum_\alpha X^\alpha \partial\xi_\alpha$ is the function $div(X) : \pounds_X \varpi_4 = div(X)\varpi_4$ and its expression in coordinates is (Maths.17.2.4) :

$div(X) = \frac{1}{\det P'}\sum_\alpha \partial_\alpha\left(X^\alpha \det P'\right)$ which reads in the SR geometry : $div(X) = \sum_\alpha \partial_\alpha\left(X^\alpha\right)$

$\frac{dL}{dz^i_\beta}$ is a vector : $Z_i = \frac{dL}{dz^i_\beta}\partial\xi_\beta$ and $\frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\left(\frac{dL}{dz^i_\beta}\det P'\right) = div\left(Z_i\right)$

$\frac{dL \det P'}{dz^i}\det P' + L\frac{d\det P'}{dz^i} = \sum_\beta \frac{d}{d\xi^\beta}\left(\frac{\partial L \det P'}{\partial z^i_\beta}\right)$

$\frac{dL \det P'}{\partial dz^i} + L\frac{1}{\det P'}\frac{d\det P'}{dz^i} = div\left(Z_i\right)$

thus, when P' does not depend on z^i the equation reads : $\frac{dL \det P'}{dz^i} = div\left(Z_i\right)$

Complex variables :

Whenever complex variables are involved the derivatives of the real and imaginary parts must be computed separately.

We have then two families of real valued equations :

$\frac{\partial L \det P'}{\partial \operatorname{Re} z^i} - \sum_\beta \frac{d}{d\xi^\beta}\frac{\partial L \det P'}{\partial \operatorname{Re} z^i_\beta} = 0$

$\frac{\partial L \det P'}{\partial \operatorname{Im} z^i} - \sum_\beta \frac{d}{d\xi^\beta}\frac{\partial L \det P'}{\partial \operatorname{Im} z^i_\beta} = 0$

and by defining the holomorphic complex valued functions :
$\frac{\partial L\det P'}{\partial z^i}=\frac{\partial L\det P'}{\partial \operatorname{Re} z^i}+\frac{1}{i}\frac{\partial L\det P'}{\partial \operatorname{Im} z^i}$
$\frac{\partial L\det P'}{\partial \overline{z^i}}=\frac{\partial L\det P'}{\partial \operatorname{Re} z^i}-\frac{1}{i}\frac{\partial L\det P'}{\partial \operatorname{Im} z^i}$
the equations read :
$\frac{\partial L\det P'}{\partial z^i}=\sum_\beta \frac{d}{d\xi^\beta}\left(\frac{\partial L\det P'}{\partial \operatorname{Re} z^i_\beta}+\frac{1}{i}\frac{\partial L\det P'}{\partial \operatorname{Im} z^i_\beta}\right)=\sum_\beta \frac{d}{d\xi^\beta}\left(\frac{\partial L\det P'}{\partial z^i_\beta}\right)$
$\frac{\partial L\det P'}{\partial \overline{z^i}}=\overline{\frac{\partial L\det P'}{\partial z^i}}=\sum_\beta \frac{d}{d\xi^\beta}\left(\frac{\partial L\det P'}{\partial \operatorname{Re} z^i_\beta}-\frac{1}{i}\frac{\partial L\det P'}{\partial \operatorname{Im} z^i_\beta}\right)$
$=\sum_\beta \frac{d}{d\xi^\beta}\left(\frac{\partial L\det P'}{\partial \overline{z^i_\beta}}\right)=\overline{\sum_\beta \frac{d}{d\xi^\beta}\left(\frac{\partial L\det P'}{\partial z^i_\beta}\right)}$
and we are left with the unique equation :
$\frac{\partial L\det P'}{\partial z^i}=\sum_\beta \frac{d}{d\xi^\beta}\left(\frac{\partial L\det P'}{\partial z^i_\beta}\right)$

6.3.2 Functional derivatives

Whenever the system is comprised of force fields or matter fields on one hand, and of individual particles on the other hand, such as :

$\int_\Omega L_1\left(P^\alpha_i,\mathcal{F}_{G\alpha\beta},\mathcal{F}_{A\alpha\beta}\right)\varpi_4+\sum_{p=1}^N\int_{\tau_p^1}^{\tau_p^2} L_2\left(\psi_p,\nabla_\alpha\psi_p,P^\alpha_i,V^\alpha_p\right)cd\tau_p$

the action is the sum of integrals on domains which do not have the same dimension. The Euler-Lagrange equations do not hold any longer. It is common to introduce Dirac's functions, but this formal and naive method is mathematically wrong : the Euler-Lagrange equations are proven in precise conditions, which are no longer met. However there is another method : functional derivatives (derivative with respect to a function). It is commonly used by physicists, but in an uncertain mathematical rigor. Actually their theory can be done in a very general context, using the extension of distributions on vector bundles (see Maths.30.3.2 and 34.1). The method provides solutions for variational problems, but is also a powerful tool to study the neighborhood of an equilibrium.

A functional : $\ell : J^rE\to\mathbb{R}$ defined on a normed subspace of sections $\mathfrak{X}\left(J^rE\right)$ of a vector bundle E has a functional derivative $\frac{\delta\ell}{\delta z}\left(Z_0\right)$ with respect to a section $Z\in\mathfrak{X}\left(E\right)$ in Z_0 if there is a distribution $\frac{\delta\ell}{\delta z}$ such that for any smooth, compactly supported $\delta Z\in\mathfrak{X}_{c\infty}\left(E\right)$:

$\lim_{\|\delta Z\|\to 0}\left\|\ell\left(Z_0+\delta Z\right)-\ell\left(Z_0\right)-\frac{\delta\ell}{\delta z}\left(Z_0\right)Z\right\|=0$

Because Z and δZ are sections of E their r-jets extensions are computed by taking the partial derivatives. The key point in the definition is that only δZ, and not its derivatives, is involved. It is clear that the functional is stationary in Z_0 if $\frac{\delta\ell}{\delta z}\left(Z_0\right)=0$.

When the functional is linear in Z then $\frac{\delta\ell}{\delta z}=\ell$.

When the functional is given by an integral : $\int_\Omega\lambda\left(J^rZ\right)d\xi^0\wedge d\xi^1\wedge d\xi^2\wedge d\xi^3$ the functional derivative is the distribution :

$\frac{\delta\ell}{\delta z}\left(\delta Z\right)=\int_\Omega\sum_{s=0}^r\sum_{\alpha_1\ldots\alpha_s}\left(-1\right)^s D_{\alpha_1\ldots\alpha_s}\frac{\partial\lambda}{\partial Z_{\alpha_1\ldots\alpha_s}}\delta Z d\xi^0\wedge d\xi^1\wedge d\xi^2\wedge d\xi^3$

so that we get back the Euler- Lagrange equations if all the functionals are integral similarly defined, but this formula can be used when the integrals are of different order.

We will see how to implement this method in the next chapter.

6.4 PERTURBATIVE LAGRANGIAN

In a perturbative approach, meaning close to the equilibrium, which are anyway the conditions in which the principle of least action applies, the lagrangian can be estimated by a development in Taylor series, meaning that each term is represented by polynomials. Because all the variables are derivatives at most of the second order and are vectorial, it is natural to look for scalar products.

6.4.1 Interactions Fields / Fields

It is generally assumed that there is no direct interaction gravitation / other fields (the deviation of light comes from the fact that the Universe, as seen in an inertial frame, is curved). So we have two distinct terms, which can involve only the strength of the field. They are two forms on M valued in the Lie algebra, which transform in a change of gauge by the adjoint map, thus the scalar product must be invariant by Ad.

We have such quantities, the density of energy of the field, defined by scalar products. So this is the obvious choice. However for the gravitational field there is the usual solution of the scalar curvature R which can be computed with our variables. It is invariant by a change of gauge or chart. The action with the scalar curvature is then the Hilbert action $\int_\Omega R\varpi_4$. Any scalar constant added to a lagrangian leads to a lagrangian which is still covariant, however the Lagrange equations give the same solutions, so the cosmological constant is added ex-post to the Einstein equation. The models use traditionally the scalar curvature, with the Levi-Civita connection. The application of the principle of least action leads then in the vacuum to the Einstein equation : $Ric_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R = 0$. In our formalism the Hilbert action leads to linear equations : R is a linear function of $\mathcal{F}_G$, so it leads to much simpler computations than the usual method (and of course they provide the same solutions).

In all the, difficult, experimental verifications, the models are highly simplified, and to tell that the choice of R is validated by facts would be a bit excessive. We have seen that its computation, mathematically legitimate, has no real physical justification : the contraction of indices is actually similar to the procedure used to define the Dirac's operator.

It seems logical to use the same quantity for the gravitational field as for the other fields. This is the option that we will follow. It is more pedagogical, and opens the possibility to study a dissymmetric gravitational field.

It will be useful to single out the EM field, so, to represent the interactions fields / fields we will take in a perturbative lagrangian (with cosmological constant) :

$$\int_\Omega \left(\sum_{\alpha\beta} C_G \left(\sum_{a=1}^{3} \mathcal{F}^a_{G\alpha\beta}\mathcal{F}^{a\alpha\beta}_G - \sum_{a=4}^{6} \mathcal{F}^a_{G\alpha\beta}\mathcal{F}^{a\alpha\beta}_G \right) + C_A \sum_{a=1}^{m} \mathcal{F}^a_{A\alpha\beta}\mathcal{F}^{a\alpha\beta}_A + +C_{EM}\mathcal{F}_{EM\alpha\beta}\mathcal{F}^{\alpha\beta}_{EM} \right) \varpi_4(m) \tag{6.6}$$

where C_G, C_A, C_{EM} are real constant scalars.

6.4.2 Interactions Particles /Fields

The derivative of the fields appear through the covariant derivative, and that it must be taken along the world line. We have to choose the scalar product. As the lagrangian involves only variables at the first order, the key quantity is the scalar product $\langle \psi, \nabla_\alpha \psi \rangle$.

We have already introduced the quantity :

$\frac{1}{i}\langle\psi,\nabla_\alpha\psi\rangle = \operatorname{Im}\langle\psi,\nabla_\alpha\psi\rangle = \frac{1}{i}\left(\langle\psi,\partial_\alpha\psi\rangle + \left\langle\psi,[\psi]\left[\grave{A}_\alpha\right]\right\rangle + \langle\psi,\gamma C(G_\alpha)\psi\rangle\right)$

So $\frac{1}{i}\langle\psi,\nabla_V\psi\rangle$ can be seen as the energy of the particle in the system, as measured by the observer.

$\frac{1}{i}\langle\psi,\nabla_V\psi\rangle = k^t\widehat{X} + \frac{1}{i}\left\langle\psi_0,[\psi_0]\left[Ad_\varkappa\widehat{\grave{A}}\right]\right\rangle$

where :

$(k_a)_{a=1}^3 = -\epsilon\left(Tr\left(\psi_R^*\sigma_a\psi_R\right)\right)_{a=1}^3$

$\sum_{\alpha=0}^3 V^\alpha X_\alpha = \sum_{\alpha=0}^3\sum_{a=1}^3 k_a X_\alpha^a V^\alpha = \widehat{X}$

V depends on the vectors r,w by :

$V = \frac{dp}{dt} = \overrightarrow{v} + c\varepsilon_0(m) = c\left(\varepsilon_0 + \epsilon\frac{a_w}{2a_w^2-1}\sum_{i=1}^3 w_i\varepsilon_i\right) = \sum_{i=0}^3 V^i\varepsilon_i = \sum_{\alpha=0}^3 V^\alpha\partial\xi_\alpha$

$V^\alpha = \sum_{0=1}^3 P_i^\alpha V^i$

$\sqrt{-\langle V,V\rangle} = \frac{c}{2a_w^2-1}$

and $\widehat{X}$ depends on r,w and linearly on their first derivative by :

$X_\alpha = [C(r)]^t\left([D(r)]\frac{dr}{dt} + \frac{1}{2}j(w)\frac{dw}{dt} + [A(w)]G_{r\alpha} + [B(w)]G_{w\alpha}\right)$

The part of the action representing the interactions is then with a density of particles :

$$\int_\Omega C_I\frac{1}{i}\langle\psi,\nabla_V\psi\rangle\mu\varpi_4 = \int_\Omega C_I\left(k^t\widehat{X} + \frac{1}{i}\left\langle\psi_0,[\psi_0]\left[Ad_\varkappa\widehat{\grave{A}}\right]\right\rangle\right)\mu\varpi_4 \tag{6.7}$$

and for a collection of particles :

$$\sum_{p=1}^N\int_0^T C_I\left(k_p^t\widehat{X}_p + \frac{1}{i}\left\langle\psi_{0p},[\psi_{0p}]\left[Ad_\varkappa\widehat{\grave{A}}_p\right]\right\rangle\right)\sqrt{-\langle V_p,V_p\rangle}dt \tag{6.8}$$

In the lagrangian used in the Standard Model [5] there is a similar expression, with the Dirac operator and $\grave{A}$ is identified with the bosons as force carriers (which requires the introduction of the Higgs boson).

There could be a part expressing the energy at rest of the particles. With the assumption of the existence of a fundamental state if takes the form $C_P\langle\psi_0,\psi_0\rangle$. In a model with a density of particle this would just be the addition of a constant.

[5]Of course the tools used in QTF are quite different (the key variables are located operators), but the comparison with the lagrangian of the Standard Model makes sense.

Chapter 7

CONTINUOUS MODELS

Continuous models represent systems where no discontinuous process occurs : the variables are mainly defined as differentiable sections of vector bundles, the trajectories do not cross, there is no creation or annihilation of particles and the bosons are absent. The application of the Principle of Least Action with a lagrangian provides usually a set of differential equations for the variables involved, which can be restricted to 6 sections on vector bundles. Then the problem is well posed : the evolution is determinist, and the solutions depend linearly on the initial conditions. However the solutions are not necessarily unique, and the Principle does not tell us how to go from one equilibrium to another.

Continuous models correspond to an ideal situation, they are nevertheless useful to study the basic relations between the variables. We will study 2 models : matter field with a density, individual particles. The main purpose is to show the computational methods, and introduce important variables : currents and energy-momentum tensor. The equations for the fields show that 2 special vector fields, the currents, one linked to the particles and the other to the fields, play a specific role and explain how the particles are the sources of the fields. The energy-momentum tensor measures the resistance that a given system opposes to a change in the equilibrium.

Continuous models provide PDE. Their solutions represent continuous evolutions, adjusted to the initial conditions. Whenever the variables meet the conditions described in the chapter 2, the theorems of QM tell us that the solutions must belong to some classes of maps, as well as the observables : they must belong to some finite or infinite dimensional vector spaces. These additional constraints provide a tool to find solutions, but also restrict the set of possible solutions. In many cases one looks for stationary solutions, which can be easily found from the PDE. Quite often only a finite number of stationary solutions are possible : the states of the system are quantized. We will not view this aspect, as there are too many different cases, and we will focus on the PDE. But it must be clear that, when the conditions are met, they are not the final point of the study.

7.1 MODEL WITH A DENSITY OF PARTICLES

We will do the computation with the perturbative lagrangian :

$\int_\Omega \{ \left(\sum_{\alpha\beta} C_G \left(\sum_{a=1}^3 \mathcal{F}^a_{G\alpha\beta} \mathcal{F}^{a\alpha\beta}_G - \sum_{a=4}^6 \mathcal{F}^a_{G\alpha\beta} \mathcal{F}^{a\alpha\beta}_G \right) + C_A \sum_{a=1}^m \mathcal{F}^a_{A\alpha\beta} \mathcal{F}^{a\alpha\beta}_A + C_{EM} \mathcal{F}_{EM\alpha\beta} \mathcal{F}^{\alpha\beta}_{EM} \right) + C_I \frac{1}{i} \langle \psi, \nabla_V \psi \rangle \mu \} \varpi_4$

7.1.1 Equation for the Matter Field

The lagrangian

The state equation can be made explicit with a perturbative lagrangian where only the EM and gravitational field are present. The interaction term is then : $C_I \mu \frac{1}{i} \langle \psi, \nabla_V \psi \rangle = C_I \mu \left(k^t \widehat{X} + \widehat{\grave{A}} \right)$ where $\widehat{\grave{A}} = \sum_{\alpha=0}^3 V^\alpha \grave{A}_\alpha$ is a scalar and k is a constant matrix 3×1. The variables are the vectors r, w which define V. We will denote $r(t) = r(\varphi_o(t,x)), w(t) = w(\varphi_o(t,x))$: the particles are followed along their trajectories, given by integral curves.

It is assumed that the density μ meets the continuity equation : $\frac{d\mu}{dt} + \mu div V = 0$.

$X_\alpha = [C(r)]^t \left([D(r)] \partial_\alpha r + \frac{1}{2} j(w) \partial_\alpha w + \left[1 - \frac{1}{2} j(w) j(w) \right] G_{r\alpha} + a_w [j(w)] G_{w\alpha} \right)$ is a matrix column 3×1

with

$[C(r)] = \left[1 + a_r j(r) + \frac{1}{2} j(r) j(r) \right]$

$[D(r)] = \left[\frac{1}{a_r} + \frac{1}{2} j(r) + \frac{1}{4a_r} j(r) j(r) \right]$

$[C(r)]^t [D(r)] = [D(r)]^t$

$r^t r = 4 \left(1 - a_r^2 \right)$

$w^t w = 4 \left(a_w^2 - 1 \right)$

Equations

The equations are :

$\forall a = 1, 2, 3 :$

$\frac{dL_I}{dr_a} = \frac{1}{\det P'} \sum_{\alpha=0}^3 \frac{d}{d\xi^\alpha} \left(\frac{\partial L_I}{\partial \partial_\alpha r_a} \det P' \right)$

$\frac{dL_I}{dw_a} = \frac{1}{\det P'} \sum_{\alpha=0}^3 \frac{d}{d\xi^\alpha} \left(\frac{\partial L_I}{\partial \partial_\alpha w_a} \det P' \right)$

The equations read for r :

$\forall a = 1, 2, 3 :$

$\frac{\partial L_I}{\partial \partial_\alpha r_a} = \sum_{\alpha=0}^3 V^\alpha \frac{\partial}{\partial \partial_\alpha r_a} C_I \mu \left(k^t X_\alpha + \grave{A}_\alpha \right) = C_I \mu V^\alpha \left(k^t [C(r)]^t [D(r)] \varepsilon_a \right) = C_I \mu V^\alpha \left(k^t [D(r)]^t \varepsilon_a \right)$

$\frac{dL_I}{dr_a} = \frac{d}{dr_a} C_I \mu \left(k^t \widehat{X} + \widehat{\grave{A}} \right) = C_I \mu \frac{d}{dr_a} \left(k^t \widehat{X} \right)$

$\mu \frac{d}{dr_a} \left(k^t \widehat{X} \right) = \frac{1}{\det P'} \sum_{\alpha=0}^3 \frac{d}{d\xi^\alpha} \left(\mu V^\alpha \left(k^t [D(r)]^t \varepsilon_a \right) \det P' \right)$

$= \mu \sum_{\alpha=0}^3 V^\alpha \frac{d}{d\xi^\alpha} \left(k^t [D(r)]^t \varepsilon_a \right)$

$+ \left(k^t [D(r)]^t \varepsilon_a \right) \left(\sum_{\alpha=0}^3 V^\alpha \frac{d\mu}{d\xi^\alpha} + \mu \frac{1}{\det P'} \sum_{\alpha=0}^3 \frac{d}{d\xi^\alpha} \left(V^\alpha \det P' \right) \right)$

$= \mu \frac{d}{dt} \left(k^t [D(r)]^t \varepsilon_a \right) + \left(k^t [D(r)]^t \varepsilon_a \right) \left(\frac{d\mu}{dt} + \mu div V \right)$

$= \mu \frac{d}{dt} \left(k^t [D(r)]^t \varepsilon_a \right)$

with the continuity equation we are left with :

$\forall a=1,2,3:\frac{d}{dr_a}\left(k^t\widehat{X}\right)=\frac{d}{dt}\left(k^t\left[D\left(r\right)\right]^t\varepsilon_a\right)$
And similarly for w, but V depends on w :
$\frac{dL_I}{dw_a}=C_I\mu\frac{d}{dw_a}\left(\sum_{\alpha=0}^3 V^\alpha\left(k^tX_\alpha+\grave{A}_\alpha\right)\right)$
$\frac{\partial L_I}{\partial\partial_\alpha w_a}=C_I\mu\sum_{\alpha=0}^3 V^\alpha\frac{\partial}{\partial\partial_\alpha w_a}\left(k^tX_\alpha+\grave{A}_\alpha\right)=\frac{1}{2}V^\alpha k^t\left[C\left(r\right)\right]^t j\left(w\right)\varepsilon_a$
Thus the equation reads :
$\forall a=1,2,3:\frac{d}{dw_a}\left(\sum_{\alpha=0}^3 V^\alpha\left(k^tX_\alpha+\grave{A}_\alpha\right)\right)=\frac{d}{dt}\left(\frac{1}{2}k^t\left[C\left(r\right)\right]^t j\left(w\right)\varepsilon_a\right)$

Derivatives

In order to get equations expressed with vectors r,w and not only their components some work must be done involving the properties of the operator j which makes the matrices C,D.

By definition :
$j\in\mathcal{L}\left(\mathbb{R}^3;\mathcal{L}\left(\mathbb{R}^3;\mathbb{R}^3\right)\right)\Rightarrow\frac{dj}{dr}=j'=j:\left(j\left(r\right)\right)'\left(u\right)=j\left(u\right)$
$j^2\in\mathcal{L}^2\left(\mathbb{R}^3,\mathbb{R}^3;\mathcal{L}\left(\mathbb{R}^3;\mathbb{R}^3\right)\right)\Rightarrow\left(j\left(x\right)j\left(y\right)\right)'\left(u,v\right)=j\left(u\right)j\left(y\right)+j\left(x\right)j\left(v\right)$
For any matrix P such that :
$P\left(x\right)=a+bj\left(x\right)+cj\left(x\right)j\left(x\right)$ with $a,b,c=Ct$
$\frac{\partial P}{\partial x}\left(y\right)=bj\left(y\right)+c\left(j\left(y\right)j\left(x\right)+j\left(x\right)j\left(y\right)\right)$ is a matrix
$z^t\frac{\partial P}{\partial x}\left(y\right)$
$=z^t\left(bj\left(y\right)+c\left(j\left(y\right)j\left(x\right)+j\left(x\right)j\left(y\right)\right)\right)$
$=bz^tj\left(y\right)+cz^t\left(xy^t+yx^t-2x^ty\right)$
$=-by^tj\left(z\right)+c\left(z^tx\right)y^t+c\left(y^tz\right)x^t-2cz^t\left(y^tx\right)$
$=y^t\left[-bj\left(z\right)+c\left(\left(z^tx\right)+zx^t-2xz^t\right)\right]$
$=y^t\left[-bj\left(z\right)+c\left(\left(z^tx\right)+j\left(x\right)j\left(z\right)+z^tx-2j\left(z\right)j\left(x\right)-2z^tx\right)\right]$
$=y^t\left[-bj\left(z\right)+c\left(j\left(x\right)j\left(z\right)-2j\left(z\right)j\left(x\right)\right)\right]$
and we will denote :

$$\left[\delta P\left(x,z\right)\right]=\left[-bj\left(z\right)+c\left(j\left(x\right)j\left(z\right)-2j\left(z\right)j\left(x\right)\right)\right] \tag{7.1}$$

$$z^t\frac{\partial P}{\partial x}\left(y\right)=y^t\left[\delta P\left(x,z\right)\right]\Leftrightarrow\left[\frac{\partial P}{\partial x}\left(y\right)\right]z=\left[\delta P\left(x,z\right)\right]^t y \tag{7.2}$$

In particular, by definition : $\frac{\partial P}{\partial r_a}=\frac{\partial P}{\partial r}\left(\varepsilon_a\right)$

$$z^t\left[\frac{\partial P}{\partial r_a}\right]=\varepsilon_a^t\left[\delta P\left(r,z\right)\right]\Leftrightarrow\left[\delta P\left(r,z\right)\right]^t\varepsilon_a=\left[\frac{\partial P}{\partial r_a}\right]^t z \tag{7.3}$$

If the coefficients a,b,c depend on a_r then :
$\frac{d\left(z^tP\right)}{dr_a}\left(r\right)=z^t\frac{\partial P}{\partial a_r}\frac{da_r}{dr_a}+z^t\frac{\partial P}{\partial r_a}=\left(-\frac{1}{4}r_a\right)z^t\frac{\partial P}{\partial a_r}+\varepsilon_a^t\left[\delta P\left(r,z\right)\right]=\varepsilon_a^t\left(\left[\delta P\left(r,z\right)\right]-\frac{1}{4}rz^t\frac{\partial P}{\partial a_r}\right)$
and similarly for $P\left(w\right)$ and a,b,c depending on a_w :
$\frac{d\left(z^tP\right)}{dw_a}\left(w\right)=z^t\frac{\partial P}{\partial a_w}\frac{da_r}{dw_a}+z^t\frac{\partial P}{\partial r_a}=\left(\frac{1}{4}w_a\right)z^t\frac{\partial P}{\partial a_w}+\varepsilon_a^t\left[\delta P\left(w,z\right)\right]=\varepsilon_a^t\left(\left[\delta P\left(w,z\right)\right]+\frac{1}{4}wz^t\frac{\partial P}{\partial a_w}\right)$

Equation for r

Using these objects, the equations read for r :

$\frac{d}{dr_a}\left(k^t\widehat{X}\right)$

$=\frac{\partial}{\partial a_r}\left(k^t\widehat{X}\right)\frac{\partial a_r}{\partial r_a}+\frac{\partial}{\partial r_a}\left(k^t\widehat{X}\right)$

$=\varepsilon_a^t\{\left(\left[\delta D^t\left(r,k\right)\right]-\frac{1}{4a_r}rk^t\frac{\partial D^t}{\partial a_r}\right)\frac{dr}{dt}$

$+\left(\left[\delta C^t\left(r,k\right)\right]-\frac{1}{4a_r}rk^t\frac{\partial C^t}{\partial a_r}\right)\left(\frac{1}{2}\left[j\left(w\right)\right]\frac{dw}{dt}+\left[1-\frac{1}{2}j\left(w\right)j\left(w\right)\right]\widehat{G}_r+a_w\left[j\left(w\right)\right]\widehat{G}_w\right)\}$

$\frac{d}{dt}\left(k^t\left[D\right]^t\varepsilon_a\right)=\frac{d}{dt}\left(\varepsilon_a^t\left[D\right]k\right)=\varepsilon_a^t\frac{dD}{dt}k$

Thus the 3 equations are equivalent to the matrices equation :

$\left(\left[\delta D^t\left(r,k\right)\right]\frac{dr}{dt}-\frac{1}{4a_r}rk^t\frac{\partial D^t}{\partial a_r}\right)\frac{dr}{dt}$

$+\left(\left[\delta C^t\left(r,k\right)\right]-\frac{1}{4a_r}rk^t\frac{\partial C^t}{\partial a_r}\right)\left(\frac{1}{2}\left[j\left(w\right)\right]\frac{dw}{dt}+\left[1-\frac{1}{2}j\left(w\right)j\left(w\right)\right]\widehat{G}_r+a_w\left[j\left(w\right)\right]\widehat{G}_w\right)$

$=\frac{dD}{dt}k$

$\frac{dD}{dt}k=\left(k^t\frac{dD^t}{dt}\right)^t$

$k^t\frac{dD^t}{dt}=\frac{da_r}{dt}k^t\frac{\partial D^t}{\partial a_r}+k^t\frac{\partial D^t}{dr}\frac{dr}{dt}$

$=\left(-\frac{1}{4a_r}\left(\frac{dr}{dt}\right)^tr\right)k^t\frac{\partial D^t}{\partial a_r}+\left(\frac{dr}{dt}\right)^t\left[\delta D^t\left(r,k\right)\right]$

$=\left(\frac{dr}{dt}\right)^t\left(-\frac{1}{4a_r}rk^t\frac{\partial D^t}{\partial a_r}+\left[\delta D^t\left(r,k\right)\right]\right)$

$\frac{dD}{dt}k=\left(-\frac{1}{4a_r}\frac{\partial D}{\partial a_r}kr^t+\left[\delta D^t\left(r,k\right)\right]^t\right)\frac{dr}{dt}$

The equation reads :

$\left(\left[\delta D^t\left(r,k\right)\right]\frac{dr}{dt}-\frac{1}{4a_r}rk^t\frac{\partial D^t}{\partial a_r}\right)\frac{dr}{dt}$

$+\left(\left[\delta C^t\left(r,k\right)\right]-\frac{1}{4a_r}rk^t\frac{\partial C^t}{\partial a_r}\right)\left(\frac{1}{2}\left[j\left(w\right)\right]\frac{dw}{dt}+\left[1-\frac{1}{2}j\left(w\right)j\left(w\right)\right]\widehat{G}_r+a_w\left[j\left(w\right)\right]\widehat{G}_w\right)$

$=\left(-\frac{1}{4a_r}\frac{\partial D}{\partial a_r}kr^t+\left[\delta D^t\left(r,k\right)\right]^t\right)\frac{dr}{dt}$

With :

$A_1=\left(\left[\delta D^t\left(r,k\right)\right]-\frac{1}{4a_r}rk^t\frac{\partial D^t}{\partial a_r}\right)-\left(\left[\delta D^t\left(r,k\right)\right]-\frac{1}{4a_r}rk^t\frac{\partial D^t}{\partial a_r}\right)^t$

$\left[\delta D^t\left(r,k\right)\right]=\frac{1}{2}j\left(k\right)+\frac{1}{4a_r}\left(j\left(r\right)j\left(k\right)-2j\left(k\right)j\left(r\right)\right)$

$\left[\delta D^t\left(r,k\right)\right]-\left[\delta D^t\left(r,k\right)\right]^t=j\left(k\right)+\frac{3}{4a_r}\left(j\left(r\right)j\left(k\right)-j\left(k\right)j\left(r\right)\right)$

$A_2\left(r\right)=\left[\delta C^t\left(r,k\right)\right]-\frac{1}{4a_r}rk^t\frac{\partial C^t}{\partial a_r}$

$\left[\delta C^t\left(r,k\right)\right]=a_rj\left(k\right)+\frac{1}{2}j\left(r\right)j\left(k\right)-j\left(k\right)j\left(r\right)$

$\frac{\partial C^t}{\partial a_r}=-j\left(r\right)$

$$\left[\begin{array}{l} A_1\left(r\right)\frac{dr}{dt}+A_2\left(r\right)\left(\frac{1}{2}\left[j\left(w\right)\right]\frac{dw}{dt}+\left[1-\frac{1}{2}j\left(w\right)j\left(w\right)\right]\widehat{G}_r+a_w\left[j\left(w\right)\right]\widehat{G}_w\right)=0\\ A_1\left(r\right)=j\left(k\right)+\left(\frac{3}{4a_r}-1\right)\left[j\left(r\right)j\left(k\right)-j\left(k\right)j\left(r\right)\right]-\frac{1}{4a_r^2}j\left(r\right)\left[j\left(r\right)j\left(k\right)-j\left(k\right)j\left(r\right)\right]j\left(r\right)\\ A_2\left(r\right)=\frac{1}{a_r}\left(2a_r^2-1\right)j\left(k\right)+\frac{1}{2}j\left(r\right)j\left(k\right)-j\left(k\right)j\left(r\right)-\frac{1}{4a_r}j\left(r\right)j\left(r\right)j\left(k\right) \end{array}\right] \tag{7.4}$$

Equation for w

$V=\sum_{\alpha=0}^{3}V^{\alpha}\partial\xi_{\alpha}$ is related to w by $V^{\alpha}=\sum_{i=0}^{3}V^{i}P_{i}^{\alpha}$ with : $V=\sum_{i=0}^{3}V^{i}\varepsilon_{i}=c\left(\varepsilon_{0}+\frac{a_{w}}{2a_{w}^{2}-1}\sum_{i=1}^{3}w_{i}\varepsilon_{i}\right)$

$k^{t}\widehat{X}+\widehat{\grave{A}}=\sum_{\alpha=0}^{3}\sum_{i=0}^{3}V^{i}P_{i}^{\alpha}\left(k^{t}X_{\alpha}+\grave{A}_{\alpha}\right)$
$=V^{0}P_{0}^{0}k^{t}X_{0}+\sum_{i=1}^{3}V^{i}P_{i}^{0}\left(k^{t}X_{\alpha}+\grave{A}_{\alpha}\right)+\sum_{\alpha=1}^{3}V^{0}P_{0}^{\alpha}\left(k^{t}X_{\alpha}+\grave{A}_{\alpha}\right)$
$+\sum_{\alpha=1}^{3}\sum_{i=1}^{3}V^{i}P_{i}^{\alpha}\left(k^{t}X_{\alpha}+\grave{A}_{\alpha}\right)$
$=V^{0}P_{0}^{0}\left(k^{t}X_{0}+\grave{A}_{0}\right)+\sum_{\alpha=1}^{3}\sum_{i=1}^{3}V^{i}P_{i}^{\alpha}\left(k^{t}X_{\alpha}+\grave{A}_{\alpha}\right)$
$=ck^{t}\left(k^{t}X_{0}+\grave{A}_{0}\right)+\sum_{\alpha=1}^{3}\sum_{i=1}^{3}c\frac{a_{w}}{2a_{w}^{2}-1}w_{i}P_{i}^{\alpha}\left(k^{t}X_{\alpha}+\grave{A}_{\alpha}\right)$
$\frac{d}{dw_{a}}\left(k^{t}\widehat{X}+\widehat{\grave{A}}\right)$
$=\sum_{\alpha=1}^{3}\sum_{i=1}^{3}\frac{dV^{i}}{dw_{a}}P_{i}^{\alpha}\left(k^{t}X_{\alpha}+\grave{A}_{\alpha}\right)+\sum_{\alpha=0}^{3}\sum_{i=1}^{3}V^{i}P_{i}^{\alpha}\frac{d}{dw_{a}}\left(k^{t}X_{\alpha}+\grave{A}_{\alpha}\right)$
$\sum_{\alpha=1}^{3}\sum_{i=1}^{3}\frac{dV^{i}}{dw_{a}}P_{i}^{\alpha}\left(k^{t}X_{\alpha}+\grave{A}_{\alpha}\right)$
$=\sum_{\alpha=1}^{3}\sum_{i=1}^{3}\frac{d}{dw_{a}}\left(\frac{\partial V^{i}}{\partial a_{w}}\frac{\partial a_{w}}{\partial w_{a}}+\frac{\partial V^{i}}{\partial w_{a}}\right)P_{i}^{\alpha}\left(k^{t}X_{\alpha}+\grave{A}_{\alpha}\right)$
$=\sum_{\alpha=1}^{3}\sum_{i=1}^{3}\left(-\frac{1}{4a_{w}}w_{a}\frac{c\left(2a_{w}^{2}+1\right)}{(2a_{w}^{2}-1)^{2}}w_{i}+\delta_{i}^{a}\frac{ca_{w}}{2a_{w}^{2}-1}\right)P_{i}^{\alpha}\left(k^{t}X_{\alpha}+\grave{A}_{\alpha}\right)$
$=-w_{a}\frac{1}{4a_{w}^{2}}\frac{2a_{w}^{2}+1}{2a_{w}^{2}-1}\sum_{\alpha=1}^{3}\sum_{i=1}^{3}V^{i}P_{i}^{\alpha}\left(k^{t}X_{\alpha}+\grave{A}_{\alpha}\right)+\sum_{\alpha=1}^{3}\sum_{i=1}^{3}\delta_{i}^{a}\frac{ca_{w}}{2a_{w}^{2}-1}P_{i}^{\alpha}\left(k^{t}X_{\alpha}+\grave{A}_{\alpha}\right)$
$=-\varepsilon_{a}^{t}\frac{1}{4a_{w}^{2}}\frac{2a_{w}^{2}+1}{2a_{w}^{2}-1}w\left(k^{t}\widehat{X}+\widehat{\grave{A}}-c\left(k^{t}X_{0}+\grave{A}_{0}\right)\right)+\varepsilon_{a}^{t}\sum_{\alpha=1}^{3}\sum_{i=1}^{3}\varepsilon_{i}\frac{ca_{w}}{2a_{w}^{2}-1}P_{i}^{\alpha}\left(k^{t}X_{\alpha}+\grave{A}_{\alpha}\right)$

$k^{t}\frac{dX_{\alpha}}{dw_{a}}=k^{t}\left(\frac{dX_{\alpha}}{da_{w}}\frac{da_{w}}{dw_{a}}+\frac{\partial X_{\alpha}}{\partial w_{a}}\right)=k^{t}\left(\frac{dX_{\alpha}}{da_{w}}\frac{1}{4a_{w}}w_{a}+\frac{\partial X_{\alpha}}{\partial w_{a}}\right)=\left(\varepsilon_{a}^{t}w\right)\left(k^{t}\frac{dX_{\alpha}}{da_{w}}\right)\frac{1}{4a_{w}}+k^{t}\frac{\partial X_{\alpha}}{\partial w_{a}}$
$X_{\alpha}=\left[C\left(r\right)\right]^{t}\left(\left[D\left(r\right)\right]\partial_{\alpha}r+\frac{1}{2}j\left(w\right)\partial_{\alpha}w+\left[1-\frac{1}{2}j\left(w\right)j\left(w\right)\right]G_{r\alpha}+a_{w}\left[j\left(w\right)\right]G_{w\alpha}\right)$
$k^{t}\frac{dX_{\alpha}}{da_{w}}=k^{t}C^{t}j\left(w\right)G_{w\alpha}=-w^{t}j\left(Ck\right)G_{w\alpha}$
$\frac{\partial}{\partial w_{a}}\left(k^{t}X_{\alpha}\right)$
$=k^{t}C^{t}\frac{\partial}{\partial w_{a}}\left(\frac{1}{2}\left[j\left(w\right)\right]\partial_{\alpha}w+\left[1-\frac{1}{2}j\left(w\right)j\left(w\right)\right]G_{r\alpha}+a_{w}\left[j\left(w\right)\right]G_{w\alpha}\right)$
$=\frac{1}{2}k^{t}C^{t}j\left(\varepsilon_{a}\right)\partial_{\alpha}w-\frac{1}{2}k^{t}B\frac{\partial}{\partial w_{a}}\left[j\left(w\right)j\left(w\right)\right]G_{r\alpha}+a_{w}k^{t}Bj\left(\varepsilon_{a}\right)G_{w\alpha}$
$=-\frac{1}{2}\varepsilon_{a}^{t}j\left(Ck\right)\partial_{\alpha}w-\frac{1}{2}\varepsilon_{a}^{t}\left[\delta j\left(w\right)j\left(w\right)\left(w,Ck\right)\right]G_{r\alpha}-a_{w}\varepsilon_{a}^{t}j\left(Ck\right)G_{w\alpha}$
$=\varepsilon_{a}^{t}\left[-\frac{1}{2}j\left(Ck\right)\partial_{\alpha}w-\frac{1}{2}\left(j\left(w\right)j\left(Ck\right)-2j\left(Ck\right)j\left(w\right)\right)G_{r\alpha}-a_{w}j\left(Ck\right)G_{w\alpha}\right]$
$\sum_{\alpha=0}^{3}V^{\alpha}k^{t}\frac{dX_{\alpha}}{dw_{a}}=-\varepsilon_{a}^{t}\{\frac{1}{2}j\left(Ck\right)\frac{dw}{dt}+\frac{1}{2}\left(j\left(w\right)j\left(Ck\right)-2j\left(Ck\right)j\left(w\right)\right)\widehat{G}_{r}$
$+\left(2a_{w}-\frac{1}{a_{w}}+\frac{1}{4a_{w}}j\left(w\right)j\left(w\right)\right)j\left(Ck\right)\widehat{G}_{w}\}$

$\frac{d}{dt}\left(\frac{1}{2}k^{t}\left[C\left(r\right)\right]^{t}j\left(w\right)\varepsilon_{a}\right)=-\frac{1}{2}\frac{d}{dt}\left(\varepsilon_{a}^{t}j\left(w\right)\left[C\left(r\right)\right]k\right)=-\varepsilon_{a}^{t}\frac{1}{2}\frac{d}{dt}\left(j\left(w\right)\left[C\left(r\right)\right]k\right)$

And the equation is equivalent to the matrices equation :

$\frac{1}{2}j\left(Ck\right)\frac{dw}{dt}+\frac{1}{2}\left(j\left(w\right)j\left(Ck\right)-2j\left(Ck\right)j\left(w\right)\right)\widehat{G}_{r}+\left(2a_{w}-\frac{1}{a_{w}}+\frac{1}{4a_{w}}j\left(w\right)j\left(w\right)\right)j\left(Ck\right)\widehat{G}_{w}-\frac{1}{2}\frac{d}{dt}\left(j\left(w\right)\left[C\left(r\right)\right]k\right)$
$=-\frac{1}{4a_{w}^{2}}\frac{2a_{w}^{2}+1}{2a_{w}^{2}-1}w\left(k^{t}\widehat{X}+\widehat{\grave{A}}-c\left(k^{t}X_{0}+\grave{A}_{0}\right)\right)+\sum_{\alpha=1}^{3}\sum_{i=1}^{3}\varepsilon_{i}\frac{ca_{w}}{2a_{w}^{2}-1}P_{i}^{\alpha}\left(k^{t}X_{\alpha}+\grave{A}_{\alpha}\right)$

$\frac{d}{dt}\left(j\left(w\right)\left[C\left(r\right)\right]k\right)$
$=-\left(\frac{d}{dt}\left(k^{t}\left[C\left(r\right)\right]^{t}\left[j\left(w\right)\right]\right)\right)^{t}$
$\frac{d}{dt}\left(k^{t}\left[C\left(r\right)\right]^{t}\left[j\left(w\right)\right]\right)$

$= k^t \frac{dC^t}{dt} j(w) + k^t C^t j\left(\frac{dw}{dt}\right)$
$= \left(k^t \frac{\partial C^t}{\partial a_r}\frac{da_r}{dt} + k^t \frac{\partial C^t}{\partial r}\left(\frac{dr}{dt}\right)\right) j(w) - \left(\frac{dw}{dt}\right)^t j(Ck)$
$= \left(\left(-\frac{1}{4a_r}\left(\frac{dr}{dt}\right)^t r\right) k^t \frac{\partial C^t}{\partial a_r} + \left(\frac{dr}{dt}\right)^t \left[\delta C^t(r,k)\right]\right) j(w) - \left(\frac{dw}{dt}\right)^t j\left(B^t k\right)$
$= \left(\frac{dr}{dt}\right)^t \left(-\frac{1}{4a_r} r k^t \frac{\partial C^t}{\partial a_r} + \left[\delta C^t(r,k)\right]\right) j(w) - \left(\frac{dw}{dt}\right)^t j(Ck)$
$\frac{d}{dt}\left(j(w)\left[C(r)\right]k\right) = j(w)\left(-\frac{1}{4a_r}\frac{\partial C}{\partial a_r} k r^t + \left[\delta C^t(r,k)\right]^t\right)\frac{dr}{dt} - j(Ck)\frac{dw}{dt}$

The equation reads :
$\frac{1}{2} j(w)\left(\frac{1}{4a_r}\frac{\partial C}{\partial a_r} k r^t - \left[\delta C^t(r,k)\right]^t\right)\frac{dr}{dt} + j(Ck)\frac{dw}{dt}$
$+\frac{1}{2}\left(j(w)j(Ck) - 2j(Ck)j(w)\right)\widehat{G}_r + \left(2a_w - \frac{1}{a_w} + \frac{1}{4a_w} j(w)j(w)\right) j(Ck)\widehat{G}_w$
$= -\frac{1}{4a_w^2}\frac{2a_w^2+1}{2a_w^2-1} w\left(k^t\widehat{X} + \widehat{\grave{A}} - c\left(k^t X_0 + \grave{A}_0\right)\right) + \sum_{\alpha=1}^3\sum_{i=1}^3 \varepsilon_i \frac{ca_w}{2a_w^2-1} P_i^\alpha\left(k^t X_\alpha + \grave{A}_\alpha\right)$
With :
$\frac{1}{2}\left(\frac{1}{4a_r}\frac{\partial C}{\partial a_r} k r^t - \left[\delta C^t(r,k)\right]^t\right) = -\frac{1}{2}\left[A_2(r)\right]$
$\left[A_3(r)\right] = j(Ck)$
$j(Ck) = j\left(\left[1 + a_r j(r) + \frac{1}{2} j(r) j(r)\right] k\right)$
$= j(k) + a_r j(j(r)k) + \frac{1}{2} j(j(r)j(r)k)$
$= j(k) + a_r\left(j(r)j(k) - j(k)j(r)\right) + \frac{1}{2}\left(\left(k^t r\right) j(r) - 4\left(1 - a_r^2\right) j(k)\right)$
$= \left(2a_r^2 - 1\right) j(k) + \frac{1}{2}\left(k^t r\right) j(r) + a_r\left(j(r)j(k) - j(k)j(r)\right)$
and, as we will see :$k^t\widehat{X} + \widehat{\grave{A}} = 0$, the equation reads :

$$\left[\begin{array}{l} -\frac{1}{2} j(w)\left[A_2(r)\right]^t \frac{dr}{dt} + \left[A_3(r)\right]\frac{dw}{dt} \\ +\frac{1}{2}\left(j(w)\left[A_3(r)\right] - 2\left[A_3(r)\right] j(w)\right)\widehat{G}_r + \left(2a_w - \frac{1}{a_w} + \frac{1}{4a_w} j(w)j(w)\right)\left[A_3(r)\right]\widehat{G}_w \\ = \frac{c}{4a_w^2}\frac{2a_w^2+1}{2a_w^2-1} w\left(k^t X_0 + \grave{A}_0\right) + \frac{ca_w}{2a_w^2-1}\sum_{\alpha,i=1}^3 \varepsilon_i P_i^\alpha\left(k^t X_\alpha + \grave{A}_\alpha\right) \\ \left[A_3(r)\right] = \left(a_r^2 - \frac{1}{2}\right) j(k) + \frac{1}{4}\left(k^t r\right) j(r) + \frac{1}{2} a_r\left(j(r)j(k) - j(k)j(r)\right) \end{array}\right] \tag{7.5}$$

One can check that, for the equations in r, w :
if (a_r, r) is a solution, then $(-a_r, -r)$ is still a solution,
if (a_w, w) is a solution, then $(-a_w, -w)$ is still a solution.
We have a reversible process : so it is determinist.

7.1.2 Equations for the gravitational field

The equations are :
$\forall a = 1..6, \alpha = 0...3$:
$\frac{d\left(L\det P'\right)}{dG_\alpha^a} = \sum_\beta \frac{d}{d\xi^\beta}\frac{d\left(L\det P'\right)}{d\partial_\beta G_\alpha^a}$

Derivatives

i) $\frac{dL}{dG_\alpha^a} = C_I \mu \frac{1}{i}\frac{\partial}{\partial G_\alpha^a}\left\langle\psi, \nabla_V\psi\right\rangle + C_G \frac{\partial}{\partial G_\alpha^a}\sum_{\lambda\mu}\sum_{b=1}^3 \mathcal{F}_{r\lambda\mu}^b \mathcal{F}_r^{b\lambda\mu} - \mathcal{F}_{w\lambda\mu}^b \mathcal{F}_w^{b\lambda\mu}$
ii) $\frac{\partial}{\partial G_\alpha^a}\left\langle\psi, \nabla_V\psi\right\rangle = V^\alpha\left\langle\psi, \left[\gamma C\left(\overrightarrow{\kappa}_a\right)\right]\left[\psi\right]\right\rangle$
iii) $\frac{\partial}{\partial G_\alpha^a}\left(\sum_{\lambda\mu}\sum_{b=1}^3 \mathcal{F}_{r\lambda\mu}^b \mathcal{F}_r^{b\lambda\mu} - \mathcal{F}_{w\lambda\mu}^b \mathcal{F}_w^{b\lambda\mu}\right)$

$= \frac{\partial}{\partial G^a_\alpha}\left(\sum_{b=1}^3 \sum_{pr\lambda\mu} \mathcal{F}^b_{r\lambda\mu} g^{p\lambda} g^{q\mu} \mathcal{F}^b_{rpq} - \mathcal{F}^b_{w\lambda\mu} g^{p\lambda} g^{q\mu} \mathcal{F}^b_{wpq}\right)$
$= 2\sum_{b=1}^3 \sum_{\lambda\mu}\left(\frac{\partial}{\partial G^a_\alpha}\mathcal{F}^b_{r\lambda\mu}\right)\mathcal{F}^{b\lambda\mu}_r - \left(\frac{\partial}{\partial G^a_\alpha}\mathcal{F}^b_{w\lambda\mu}\right)\mathcal{F}^{b\lambda\mu}_w$
$\frac{\partial}{\partial G^a_\alpha}\mathcal{F}^b_{r\lambda\mu} = 2\frac{\partial}{\partial G^a_\alpha}\left[j\left(G_{r\lambda}\right)G_{r\mu} - j\left(G_{w\lambda}\right)G_{w\mu}\right]^b = 2\frac{\partial}{\partial G^a_\alpha}\sum_{p,q=1}^3 \epsilon\left(b,p,q\right)\left[G^p_{r\lambda}G^q_{r\mu} - G^p_{w\lambda}G^q_{w\mu}\right]$
$\frac{\partial}{\partial G^a_\alpha}\mathcal{F}^b_{w\lambda\mu} = 2\frac{\partial}{\partial G^a_\alpha}\left[j\left(G_{w\lambda}\right)G_{r\mu} + j\left(G_{r\lambda}\right)G_{w\mu}\right]^b = 2\frac{\partial}{\partial G^a_\alpha}\sum_{p,q=1}^3 \epsilon\left(b,p,q\right)\left[G^p_{w\lambda}G^q_{r\mu} + G^p_{r\lambda}G^q_{w\mu}\right]$
$a = 1,2,3:$
$\frac{\partial}{\partial G^a_{r\alpha}}\mathcal{F}^b_{r\lambda\mu} = 2\sum_{c=1}^3 \epsilon\left(b,a,c\right)\left(\delta^\lambda_\alpha G^c_{r\mu} - \delta^\mu_\alpha G^c_{r\lambda}\right)$
$\frac{\partial}{\partial G^a_\alpha}\mathcal{F}^b_{w\lambda\mu} = 2\sum_{c=1}^3 \epsilon\left(b,a,c\right)\left(-\delta^\mu_\alpha G^c_{w\lambda} + \delta^\lambda_\alpha G^c_{w\mu}\right)$
$\frac{\partial}{\partial G^a_\alpha}\left(\sum_{\lambda\mu}\sum_{b=1}^3 \mathcal{F}^b_{r\lambda\mu}\mathcal{F}^{b\lambda\mu}_r - \mathcal{F}^b_{w\lambda\mu}\mathcal{F}^{b\lambda\mu}_w\right) = -8\sum_\lambda\left(j\left(\mathcal{F}^{\alpha\lambda}_r\right)G_{r\lambda} - j\left(\mathcal{F}^{\alpha\lambda}_w\right)G_{w\lambda}\right)^a$
$a = 4,5,6:$
$\frac{\partial}{\partial G^a_{w\alpha}}\mathcal{F}^b_{r\lambda\mu} = -2\sum_{c=1}^3 \epsilon\left(b,a,c\right)\left(\delta^\lambda_\alpha G^c_{w\mu} - \delta^\mu_\alpha G^c_{w\lambda}\right)$
$\frac{\partial}{\partial G^a_\alpha}\mathcal{F}^b_{w\lambda\mu} = 2\sum_{c=1}^3 \epsilon\left(b,a,c\right)\left(\delta^\lambda_\alpha G^c_{r\mu} - G^c_{r\lambda}\delta^\mu_\alpha\right)$
$\frac{\partial}{\partial G^a_\alpha}\left(\sum_{\lambda\mu}\sum_{b=1}^3 \mathcal{F}^b_{r\lambda\mu}\mathcal{F}^{b\lambda\mu}_r - \mathcal{F}^b_{w\lambda\mu}\mathcal{F}^{b\lambda\mu}_w\right) = 8\sum_\lambda\left(j\left(\mathcal{F}^{\alpha\lambda}_r\right)G_{w\lambda} + j\left(\mathcal{F}^{\alpha\lambda}_w\right)G_{r\lambda}\right)^a$
iv) $\frac{dL}{d\partial_\beta G^a_\alpha} = C_G\frac{\partial}{\partial\partial_\beta G^a_\alpha}\left(\sum_{b=1}^3 \sum_{pr\lambda\mu} \mathcal{F}^b_{r\lambda\mu} g^{p\lambda} g^{q\mu} \mathcal{F}^b_{rpq} - \mathcal{F}^b_{w\lambda\mu} g^{p\lambda} g^{q\mu} \mathcal{F}^b_{wpq}\right)$
$= 2C_G\sum_{b=1}^3 \sum_{\lambda\mu}\left(\frac{\partial}{\partial\partial_\beta G^a_\alpha}\mathcal{F}^b_{r\lambda\mu}\right)\mathcal{F}^{b\lambda\mu}_r - \left(\frac{\partial}{\partial\partial_\beta G^a_\alpha}\mathcal{F}^b_{w\lambda\mu}\right)\mathcal{F}^{b\lambda\mu}_w$
$a = 1,2,3:$
$\frac{\partial}{\partial\partial_\beta G^a_\alpha}\mathcal{F}^b_{r\lambda\mu} = \frac{\partial}{\partial\partial_\beta G^a_\alpha}\left(\partial_\lambda G^b_{r\mu} - \partial_\mu G^b_{r\lambda}\right) = 0$
$\frac{dL}{d\partial_\beta G^a_\alpha} = -4C_G\mathcal{F}^{a\alpha\beta}_r$
$a = 4,5,6:$
$\frac{\partial}{\partial\partial_\beta G^a_{w\alpha}}\mathcal{F}^b_{r\lambda\mu} = \frac{\partial}{\partial\partial_\beta G^a_{w\alpha}}\left(\partial_\lambda G^b_{r\mu} - \partial_\mu G^b_{r\lambda}\right) = 0$
$\frac{\partial}{\partial\partial_\beta G^a_\alpha}\mathcal{F}^b_{w\lambda\mu} = 4C_G\mathcal{F}^{a\alpha\beta}_w$

Equations :

$\forall\alpha = 0,...3$
$\forall a = 1,2,3$
$C_I\mu\frac{1}{i}V^\alpha\left\langle\psi,\left[\gamma C\left(\overrightarrow{\kappa}_a\right)\right]\left[\psi\right]\right\rangle - 8C_G\sum_\lambda\left(j\left(\mathcal{F}^{\alpha\lambda}_r\right)G_{r\lambda} - j\left(\mathcal{F}^{\alpha\lambda}_w\right)G_{w\lambda}\right)^a = -\frac{1}{\det P'}4C_G\sum_\beta\frac{d}{d\xi^\beta}\left(\mathcal{F}^{a\alpha\beta}_r\det P'\right)$
$\forall a = 4,5,6$
$C_I\mu\frac{1}{i}V^\alpha\left\langle\psi,\left[\gamma C\left(\overrightarrow{\kappa}_a\right)\right]\left[\psi\right]\right\rangle + 8C_G\sum_{\lambda=0}^3\left[j\left(\mathcal{F}^{\alpha\lambda}_r\right)G_{w\lambda} + j\left(\mathcal{F}^{\alpha\lambda}_w\right)G_{r\lambda}\right]^a = \frac{1}{\det P'}4C_G\sum_\beta\frac{d}{d\xi^\beta}\left(\mathcal{F}^{a\alpha\beta}_w\det P'\right)$

by product with $\overrightarrow{\kappa}_a$ and summation :
$C_I\mu\frac{1}{i}V^\alpha\upsilon\left(-\left\langle\psi,\left[\gamma C\left(\overrightarrow{\kappa}_a\right)\right]\left[\psi\right]\right\rangle,\left\langle\psi,\left[\gamma C\left(\overrightarrow{\kappa}_{a+3}\right)\right]\left[\psi\right]\right\rangle\right)$
$+8C_G\sum_\lambda\upsilon\left(\left(j\left(\mathcal{F}^{\alpha\lambda}_r\right)G_{r\lambda} - j\left(\mathcal{F}^{\alpha\lambda}_w\right)G_{w\lambda}\right), j\left(\mathcal{F}^{\alpha\lambda}_r\right)G_{w\lambda} + j\left(\mathcal{F}^{\alpha\lambda}_w\right)G_{r\lambda}\right)$
$= \frac{1}{\det P'}4C_G\sum_\beta\frac{d}{d\xi^\beta}\upsilon\left(\mathcal{F}^{a\alpha\beta}_r\det P', \mathcal{F}^{a\alpha\beta}_w\det P'\right)$
$\frac{C_I}{8C_G}\mu\frac{1}{i}V^\alpha\upsilon\left(\left\langle\psi,\left[\gamma C\left(\overrightarrow{\kappa}_a\right)\right]\left[\psi\right]\right\rangle,\left\langle\psi,\left[\gamma C\left(\overrightarrow{\kappa}_{a+3}\right)\right]\left[\psi\right]\right\rangle\right)$
$+\sum_\lambda\left[\mathcal{F}^{\alpha\lambda}_r, G_{r\lambda}\right]^a\overrightarrow{\kappa}_a = \frac{1}{2}\frac{1}{\det P'}\sum_\beta\frac{d}{d\xi^\beta}\left(\mathcal{F}^{\alpha\beta}_G\det P'\right)$
That is :

$$\left[\begin{array}{l} \forall\alpha = 0,...3: \phi^\alpha_G - J^\alpha_G = \frac{1}{2}\frac{1}{\det P'}\sum_\beta\frac{d}{d\xi^\beta}\left(\mathcal{F}^{\alpha\beta}_G\det P'\right) \\ J^\alpha_G = \frac{C_I}{8C_G}\mu\frac{1}{i}V^\alpha\upsilon\left(\left\langle\psi,\left[\gamma C\left(\overrightarrow{\kappa}_a\right)\right]\left[\psi\right]\right\rangle, -\left\langle\psi,\left[\gamma C\left(\overrightarrow{\kappa}_{a+3}\right)\right]\left[\psi\right]\right\rangle\right) \\ \phi^\alpha_G = \sum_{a=1}^6\sum_{\beta=0}^3\left[\mathcal{F}^{\alpha\beta}_r, G_{r\beta}\right]^a\overrightarrow{\kappa}_a \end{array}\right] \quad (7.6)$$

7.1.3 Equation for the EM field

i) $\frac{dL}{d\grave{A}_\alpha} = C_I \mu \frac{1}{i} \left\langle \psi, \frac{\partial}{\partial \grave{A}^a_\alpha} \nabla_V \psi \right\rangle + C_{EM} \frac{\partial}{\partial \grave{A}_\alpha} \sum_{\lambda\mu} \mathcal{F}_{EM\lambda\mu} \mathcal{F}^{\lambda\mu}_{EM}$

$\left\langle \psi, \frac{\partial}{\partial \grave{A}_\alpha} \nabla_V \psi \right\rangle = V^\alpha \langle \psi, i\psi \rangle$

$\frac{\partial}{\partial \grave{A}^a_\alpha} \sum_{\lambda\mu} \mathcal{F}_{EM\lambda\mu} \mathcal{F}^{\lambda\mu}_{EM} = 0$

ii) $\frac{dL}{d\partial_\beta \grave{A}_\alpha} = C_{EM} \frac{\partial}{\partial \partial_\beta \grave{A}_\alpha} \sum_{\lambda\mu} \left(\mathcal{F}_{EM\lambda\mu} \sum_{p,q} g^{\lambda p} g^{\mu q} \mathcal{F}_{EMpq} \right)$

$= 2C_{EM} \sum_{\lambda\mu} \left(\frac{\partial}{\partial \partial_\beta \grave{A}_\alpha} \mathcal{F}_{EM\lambda\mu} \right) \mathcal{F}^{\lambda\mu}_{EM}$

$\frac{\partial}{\partial \partial_\beta \grave{A}_\alpha} \mathcal{F}_{EM\lambda\mu} = \frac{\partial}{\partial \partial_\beta \grave{A}_\alpha} \left(\partial_\lambda \grave{A}_\mu - \partial_\mu \grave{A}_\lambda \right) = \delta^\lambda_\beta \delta^\mu_\alpha - \delta^\lambda_\alpha \delta^\mu_\beta$

$\frac{dL}{d\partial_\beta \grave{A}_\alpha} == -4C_{EM} \mathcal{F}^{\alpha\beta}_{EM}$

iii) Equation :

$\forall \alpha : C_I \mu V^\alpha \langle \psi, \psi \rangle = -4C_{EM} \frac{1}{\det P'} \sum_\beta \frac{d}{d\xi^\beta} \mathcal{F}^{\alpha\beta}_{EM}$

$$\left[\begin{array}{l} \forall \alpha = 0...3 : -J^\alpha_{EM} = \frac{1}{2} \frac{1}{\det P'} \sum_\beta \frac{d}{d\xi^\beta} \mathcal{F}^{\alpha\beta}_{EM} \\ J^\alpha_{EM} = \frac{1}{8} \frac{C_I}{C_{EM}} \mu V^\alpha \langle \psi_0, \psi_0 \rangle \end{array} \right] \quad (7.7)$$

7.1.4 Equation for the other fields

Derivatives :

i) $\frac{dL}{d\grave{A}^a_\alpha} = C_I \mu \frac{1}{i} \left\langle \psi, \frac{\partial}{\partial \grave{A}^a_\alpha} \nabla_V \psi \right\rangle + C_A \frac{\partial}{\partial \grave{A}^a_\alpha} \sum_{\lambda\mu} \sum_{b=1}^m \left(\mathcal{F}^b_{A\lambda\mu} \mathcal{F}^{a\lambda\mu}_A \right)$

$\left\langle \psi, \frac{\partial}{\partial \grave{A}^a_\alpha} \nabla_V \psi \right\rangle = V^\alpha \langle \psi, [\theta_a][\psi] \rangle$

$\frac{\partial}{\partial \grave{A}^a_\alpha} \sum_{\lambda\mu} \sum_{b=1}^m \left(\mathcal{F}^b_{A\lambda\mu} \mathcal{F}^{a\lambda\mu}_A \right) = \frac{\partial}{\partial \grave{A}^a_\alpha} \left(\sum_{b=1}^m \sum_{pr\lambda\mu} \mathcal{F}^b_{A\lambda\mu} g^{p\lambda} g^{q\mu} \mathcal{F}^b_{Apq} \right)$

$= 2 \sum_{b=1}^m \sum_{\lambda\mu} \left(\frac{\partial}{\partial \grave{A}^a_\alpha} \mathcal{F}^b_{A\lambda\mu} \right) \mathcal{F}^{b\lambda\mu}_A$

$\frac{\partial}{\partial \grave{A}^a_\alpha} \mathcal{F}^b_{A\lambda\mu} = 2 \frac{\partial}{\partial \grave{A}^a_\alpha} \left[\grave{A}_\lambda, \grave{A}_\mu \right]^b$

$= 2 \frac{\partial}{\partial \grave{A}^a_\alpha} \left[\sum_{c=1}^m \grave{A}^c_\lambda \overrightarrow{\theta}_c, \sum_{d=1}^m \grave{A}^d_\mu \overrightarrow{\theta}_d \right]^b$

$= 2 \left(\left[\delta^\lambda_\alpha \overrightarrow{\theta}_a, \sum_{d=1}^m \grave{A}^d_\mu \overrightarrow{\theta}_d \right]^b + \left[\sum_{c=1}^m \grave{A}^c_\lambda \overrightarrow{\theta}_c, \delta^\mu_\alpha \overrightarrow{\theta}_a \right]^b \right)$

$= 2 \left(\delta^\lambda_\alpha \left[\overrightarrow{\theta}_a, \grave{A}_\mu \right]^b + \delta^\mu_\alpha \left[\grave{A}_\lambda, \overrightarrow{\theta}_a \right]^b \right)$

$\frac{\partial}{\partial \grave{A}^a_\alpha} \sum_{\lambda\mu} \sum_{b=1}^m \left(\mathcal{F}^b_{A\lambda\mu} \mathcal{F}^{a\lambda\mu}_A \right)$

$= 4 \sum_{b=1}^m \sum_{\lambda\mu} \left(\delta^\lambda_\alpha \left[\overrightarrow{\theta}_a, \grave{A}_\mu \right]^b \mathcal{F}^{b\lambda\mu}_A + \delta^\mu_\alpha \left[\grave{A}_\lambda, \overrightarrow{\theta}_a \right]^b \mathcal{F}^{b\lambda\mu}_A \right)$

$= 4 \sum_{b=1}^m \sum_{\lambda\mu} \left(\left[\overrightarrow{\theta}_a, \grave{A}_\mu \right]^b \mathcal{F}^{b\alpha\mu}_A + \left[\grave{A}_\lambda, \overrightarrow{\theta}_a \right]^b \mathcal{F}^{b\lambda\alpha}_A \right)$

$= 8 \sum_{b=1}^m \sum_\lambda \left(\left[\overrightarrow{\theta}_a, \grave{A}_\lambda \right]^b \mathcal{F}^{b\alpha\lambda}_A \right)$

$= 8 \sum_\lambda \left\langle \left[\overrightarrow{\theta}_a, \grave{A}_\lambda \right], \mathcal{F}^{\alpha\lambda}_A \right\rangle_{T_1 U}$

$= 8 \sum_\lambda \left\langle \overrightarrow{\theta}_a, \left[\grave{A}_\lambda, \mathcal{F}^{\alpha\lambda}_A \right] \right\rangle_{T_1 U}$

$\frac{\partial}{\partial \grave{A}^a_\alpha} \sum_{\lambda\mu} \sum_{b=1}^m \left(\mathcal{F}^b_{A\lambda\mu} \mathcal{F}^{a\lambda\mu}_A \right) = 8 \sum_\lambda \left[\grave{A}_\lambda, \mathcal{F}^{\alpha\lambda}_A \right]^a$

Using : $\forall X, Y, Z \in T_1 U : \langle X, [Y, Z] \rangle = \langle [X, Y], Z \rangle$ and the fact that the basis is orthonormal.

$\frac{dL}{d\grave{A}^a_\alpha} = C_I \mu \frac{1}{i} V^\alpha \langle \psi, [\theta_a] [\psi] \rangle - 8 C_A \sum_{\lambda=0}^3 \left[\mathcal{F}^{\alpha\lambda}_A, \grave{A}_\lambda \right]^a$

ii) $\frac{dL}{d\partial_\beta \grave{A}^a_\alpha} = C_A \frac{\partial}{\partial\partial_\beta \grave{A}^a_\alpha} \sum_{b=1}^3 \sum_{\lambda\mu} \mathcal{F}^b_{A\lambda\mu} \mathcal{F}^{b\lambda\mu}_A$

$= 2 C_A \sum_{b=1}^m \sum_{\lambda\mu} \left(\frac{\partial}{\partial\partial_\beta \grave{A}^a_\alpha} \mathcal{F}^b_{A\lambda\mu} \right) \mathcal{F}^{b\lambda\mu}_A$

$\frac{d\mathcal{F}^b_{A\lambda\mu}}{d\partial_\beta \grave{A}^a_\alpha} = -4 C_A \mathcal{F}^{a\alpha\beta}_A$

Equation

The equation reads :

$\forall \alpha = 0...3, \forall a = 1, ...m$

$C_I \mu \frac{1}{i} V^\alpha \langle \psi, [\theta_a] [\psi] \rangle - 8 C_A \sum_{\lambda=0}^3 \left[\mathcal{F}^{\alpha\lambda}_A, \grave{A}_\lambda \right]^a = -4 C_A \frac{1}{\det P'} \sum_\beta \frac{d}{d\xi^\beta} \left(\mathcal{F}^{a\alpha\beta}_A \right)$

That is :

$$\left[\begin{array}{l} \forall \alpha = 0...3 : \phi^\alpha_A - J^\alpha_A = \frac{1}{2} \frac{1}{\det P'} \frac{d}{d\xi^\beta} \left(\mathcal{F}^{a\alpha\beta}_A \det P' \right) \\ \phi^\alpha_A = \sum_{a=1}^m \sum_{\beta=0}^3 \left[\mathcal{F}^{\alpha\beta}_A, \grave{A}_\beta \right] \overrightarrow{\theta}_a \\ J^{a\alpha}_A = \frac{C_I}{8 C_A} \mu V^\alpha \sum_{a=1}^m \left\langle \psi, [\psi] \frac{1}{i} [\theta_a] \right\rangle \end{array} \right] \quad (7.8)$$

7.1.5 Equation for the tetrad

The equations are :

$\forall \alpha, i : \frac{dL \det P'}{dP^\alpha_i} = \sum_\beta \frac{d}{d\xi^\beta} \left(\frac{\partial L \det P'}{\partial\partial_\beta P^\alpha_i} \right) = 0$

The derivative of the determinant is (Maths.490) :

$\frac{\partial \det P'}{\partial P^\alpha_i} = -\left(\frac{1}{\det P} \right)^2 \frac{\partial \det P}{\partial P^\alpha_i} = -\left(\frac{1}{\det P} \right)^2 P'^i_\alpha \det P = -P'^i_\alpha \det P'$

So the equations read :

$\frac{dL}{dP^\alpha_i} \det P' - L (\det P') P'^i_\alpha = 0$

By product with P^β_i and summation on i :

$$\forall \alpha, \beta = 0...3 : \sum_i \frac{dL}{dP^\alpha_i} P^\beta_i - L \delta^\alpha_\beta = 0 \quad (7.9)$$

Derivatives

For the part related to the fields :

$\frac{dL_1}{dP^\alpha_i} =$

$\sum_{\rho\theta\lambda\mu} \frac{\partial}{\partial P^\alpha_i} \left(g^{\lambda\rho} g^{\mu\theta} \right) \left(C_G \sum_{a=1}^3 \left(\mathcal{F}^a_{r\lambda\mu} \mathcal{F}^a_{r\rho\theta} - \mathcal{F}^a_{w\lambda\mu} \mathcal{F}^a_{w\rho\theta} \right) + C_A \sum_{a=1}^m \mathcal{F}^a_{A\lambda\mu} \mathcal{F}^a_{A\rho\theta} + C_{EM} \mathcal{F}_{EM\lambda\mu} \mathcal{F}_{EM\rho\theta} \right)$

$\frac{dL_1}{dP^\alpha_i}$

$= 4 \sum_{\theta\lambda\mu} \sum_j \eta^{ij} g^{\mu\lambda} P^\theta_j \{ C_G \sum_{a=1}^3 \mathcal{F}^a_{r\alpha\mu} \mathcal{F}^a_{r\theta\lambda} - \mathcal{F}^a_{w\alpha\mu} \mathcal{F}^a_{w\theta\lambda} + C_A \sum_{a=1}^m \mathcal{F}^a_{A\alpha\mu} \mathcal{F}^a_{A\theta\lambda} + C_{EM} \mathcal{F}_{EM\alpha\mu} \mathcal{F}^a_{EM\theta\lambda} \}$

$\frac{d}{dP^\alpha_i} \left(\sum_{\alpha,i=0}^3 C_I \mu \frac{1}{i} V^i P^\alpha_i \langle \psi, \nabla_\alpha \psi \rangle \right) = C_I \mu \frac{1}{i} V^i \langle \psi, \nabla_\alpha \psi \rangle$

Equations :

$\forall \alpha, \beta = 0...3:$
$4\sum_{\theta\lambda\mu}\sum_{ij}\eta^{ij}g^{\mu\lambda}P_i^\beta P_j^\theta\{C_G\sum_{a=1}^3\mathcal{F}_{r\alpha\mu}^a\mathcal{F}_{r\theta\lambda}^a - \mathcal{F}_{w\alpha\mu}^a\mathcal{F}_{w\theta\lambda}^a + C_A\sum_{a=1}^m\mathcal{F}_{A\alpha\mu}^a\mathcal{F}_{A\theta\lambda}^a + C_{EM}\mathcal{F}_{EM\alpha\mu}\mathcal{F}_{EM\theta\lambda}^a\}$
$+C_I\mu\frac{1}{i}\sum_i V^iP_i^\beta\langle\psi,\nabla_\alpha\psi\rangle - L\delta_\beta^\alpha = 0$
$4\sum_\mu\{C_G\sum_{a=1}^3\mathcal{F}_{r\alpha\mu}^a\mathcal{F}_r^{a\beta\mu} - \mathcal{F}_{w\alpha\mu}^a\mathcal{F}_w^{a\beta\mu} + C_A\sum_{a=1}^m\mathcal{F}_{A\alpha\mu}^a\mathcal{F}_A^{a\beta\mu} + C_{EM}\mathcal{F}_{EM\alpha\mu}\mathcal{F}_{EM}^{\beta\mu}\}$
$+C_I\mu\frac{1}{i}\sum_i V^iP_i^\beta\langle\psi,\nabla_\alpha\psi\rangle - L\delta_\beta^\alpha = 0$
The equation reads :
$\forall \alpha, \beta = 0...3: C_I\mu\frac{1}{i}V^\beta\langle\psi,\nabla_\alpha\psi\rangle + 4\sum_\gamma\{4C_G\left\langle\mathcal{F}_{G\alpha\gamma},\mathcal{F}_G^{\beta\gamma}\right\rangle_{Cl} + C_A\left\langle\mathcal{F}_{A\alpha\gamma},\mathcal{F}_A^{\beta\gamma}\right\rangle_{T_1U} + C_{EM}\mathcal{F}_{EM\alpha\gamma}\mathcal{F}_{EM}^{\beta\gamma}\}$
$= \delta_\beta^\alpha\sum_{\lambda\mu}\{C_G\sum_{a=1}^3\mathcal{F}_{r\lambda\mu}^a\mathcal{F}_r^{a\lambda\mu} - \mathcal{F}_{w\lambda\mu}^a\mathcal{F}_w^{a\lambda\mu} + C_A\sum_{a=1}^m\mathcal{F}_{A\lambda\mu}^a\mathcal{F}_A^{a\lambda\mu} + C_{EM}\mathcal{F}_{EM\lambda\mu}\mathcal{F}_{EM}^{\lambda\mu}\}$
By taking $\alpha = \beta$ and summing :
$C_I\mu\frac{1}{i}\langle\psi,\nabla_V\psi\rangle + 4\sum_{\alpha\gamma}\{4C_G\langle\mathcal{F}_{G\alpha\gamma},\mathcal{F}_G^{\alpha\gamma}\rangle_{Cl} + C_A\langle\mathcal{F}_{A\alpha\gamma},\mathcal{F}_A^{\alpha\gamma}\rangle_{T_1U} + C_{EM}\mathcal{F}_{EM\alpha\gamma}\mathcal{F}_{EM}^{\alpha\gamma}\}$
$= 4\{C_I\mu\frac{1}{i}\langle\psi,\nabla_V\psi\rangle + \sum_{\lambda\mu}C_G\sum_{a=1}^3\mathcal{F}_{r\lambda\mu}^a\mathcal{F}_r^{a\lambda\mu} - \mathcal{F}_{w\lambda\mu}^a\mathcal{F}_w^{a\lambda\mu} + C_A\sum_{a=1}^m\mathcal{F}_{A\lambda\mu}^a\mathcal{F}_A^{a\lambda\mu} + C_{EM}\mathcal{F}_{EM\lambda\mu}\mathcal{F}_{EM}^{\lambda\mu}\}$
$\Rightarrow$

$$\langle\psi,\nabla_V\psi\rangle = 0 \tag{7.10}$$

The equation reads :

$$\left[\begin{array}{l} \forall \alpha, \beta = 0...3: \\ C_I\mu\frac{1}{i}V^\beta\langle\psi,\nabla_\alpha\psi\rangle + \\ 4\sum_\gamma\{4C_G\left\langle\mathcal{F}_{G\alpha\gamma},\mathcal{F}_G^{\beta\gamma}\right\rangle_{Cl} + C_A\left\langle\mathcal{F}_{A\alpha\gamma},\mathcal{F}_A^{\beta\gamma}\right\rangle_{T_1U} + C_{EM}\mathcal{F}_{EM\alpha\gamma}\mathcal{F}_{EM}^{\beta\gamma}\} \\ = 2\delta_\alpha^\beta\left(\sum_{\lambda\mu}\left(C_G\langle\mathcal{F}_G,\mathcal{F}_G\rangle + C_A\langle\mathcal{F}_A,\mathcal{F}_A\rangle + C_{EM}\langle\mathcal{F}_{EM},\mathcal{F}_{EM}\rangle\right)\right) \end{array}\right] \tag{7.11}$$

the factor 2 accounting for the fact that the indices λ, μ are not ordered in the lagrangian.
As a consequence, in the general case :
$\frac{1}{i}\langle\psi,\nabla_V\psi\rangle = k^t\widehat{X} + \frac{1}{i}\left\langle\psi_0,[\psi_0]\left[Ad_\varkappa\widehat{\grave{A}}\right]\right\rangle = 0$
and with the EM field :

$$k^t\widehat{X} + \widehat{\grave{A}} = 0 \tag{7.12}$$

which gives :
$\widehat{\grave{A}} = -k^t[C(r)]^t\left([D(r)]\frac{dr}{dt} + \frac{1}{2}[j(w)]\frac{dw}{dt} + \left[1 - \frac{1}{2}j(w)j(w)\right]\widehat{G}_r + a_w[j(w)]\widehat{G}_w\right)$

These general solutions must be adjusted to the fields equations to find the trajectory. The density is deduced from the continuity equation. Because
$w \simeq \left(1 + \frac{1}{8}\frac{v^2}{c^2}\right)\frac{v}{c}$
$a_w \simeq 1 + \frac{1}{8}\frac{v^2}{c^2}$
$\frac{2a_w^2 - 1}{a_w} \simeq 1 + \frac{1}{8}\frac{v^2}{c^2}$
one can take only the first orders terms in w.

7.1.6 Deformable solid

The computation of a the single spinor representing a deformable solid by the aggregation of particles (see Spinors) is a special case.

i) If the external fields are given, then r, w, μ are deduced from the state equation and the continuity equations, with the parameters $\widehat{G}, \widehat{\grave{A}}$, and adjustment to the initial conditions. This is the study of the deformation of the body submitted to given forces.

If the fields have the same value at any point of the material body ($\widehat{G}, \widehat{\grave{A}}$ do not depend on x) then the solutions r, w depend only on t and the initial conditions.

$N(t) = \int_{\omega(t)} \mu_3(x,t)\,\varpi_3(t,x)$

$\widehat{r}(t) = r(t)\,N(t), \widehat{w}(\tau) = w(t)\,N(t), \widehat{a}(t) = a(t)\,N(t), \widehat{b}(t) = b(t)\,N(t)$

The constraints are met :

$\widehat{a}(\tau)\,\widehat{b}(\tau) = a(t)\,b(t)\,N(t)^2 = -\frac{1}{4}r(t)^t\,w(t)\,N(t)^2 = -\frac{1}{4}R(t)^t\,W(t)$

$\widehat{a}^2 - \widehat{b}^2 = \left(a(t)^2 - b(t)^2\right)N(t)^2 = \left(1 + \frac{1}{4}\left(w(t)^t\,w(t) - r(t)^t\,r(t)\right)\right)N(t)^2 = N^2 + \frac{1}{4}\left(\widehat{w}^t\widehat{w} - \widehat{r}^t\widehat{r}\right)$

and there is always a solution.

If $\widehat{G}, \widehat{\grave{A}}$ depend on x the conditions are met if other fields are added, which represent the internal forces necessary to keep the cohesion of the body.

So actually the state equation can be used the other way around. The model is built explicitly on the assumption that the particles constituting the material body are represented by a matter field, so that its cohesion is kept. Then, for each value of the parameters $\widehat{G} = \widehat{G}_{ext} + \widehat{G}_{int}, \widehat{\grave{A}} = \widehat{\grave{A}}_{ext} + \widehat{\grave{A}}_{int}$, the equation gives the deformation of the solid under external fields, and the value of the internal fields which are necessary to keep its cohesion. In particular a rigid body is such that S_B,and thus σ, is constant. As can be seen in the equation $\frac{dr}{dt} = 0, \frac{dw}{dt} = 0 \nRightarrow r = 0, w = 0$. The constant solutions depend on the value of the fields, whose total must be constant : the internal forces counterbalance the external fields. And the conditions above are limiting conditions for the internal forces.

ii) In the previous case the fields equations are ignored : the deformation of the body does not change the value of the fields. But in some cases this effect cannot be ignored and the full model is required. Then usually there is no external field : the system is self-contained. However we need to precise what is the observer. In the previous case ε_0 is arbitrary (it represents the trajectory of the whole body in a more general model) and the tetrad (and the metric) are given. With the full model the tetrad is deformed with respect to a chart (through P), and this is this chart which provides the external reference for the full motion. This chart is arbitrary : it will be the chart of a model in which the body will be inserted as a single spinor.

7.2 MODEL WITH INDIVIDUAL PARTICLES

We consider a system of a fixed number N of particles p=1...N interacting with the fields, represented as above :

- the state of each particle is represented by a map : $\psi_p : [0,T] \to Q[E \otimes F, \vartheta]$

- each of them occupies a location, which is a point m_p in M, and in the chart φ_o of an observer : $m_p(t) = \varphi_0(t, x_p(t)) \in \Omega_3(t)$ with

$V_p(t) = \frac{dm_p}{dt} = \overrightarrow{v}_p + c\varepsilon_0(m_p(t))$

- the fields are represented by their potential $G_\alpha, \grave{A}_\alpha$ and their strength $\mathcal{F}_{G\alpha\beta}, \mathcal{F}_{A\alpha\beta}$

- the value of the potential of the fields at m_p is denoted as usual :

$\widehat{A}_p^a = \sum_{\alpha=0}^3 \grave{A}_\alpha^a(m_p(t)) V_p^\alpha(t)$

$\widehat{G}_p^a = \sum_{\alpha=0}^3 G_\alpha^a(m_p(t)) V_p^\alpha(t)$

If there is an external field it should be added as a parameter (the field which is computed is the total field : internal + external). Similarly if the trajectory is known, it should be incorporated in the model. What we have here is a collection of particles in equilibrium with their own field (for N=1 this is the "Lorentz-Dirac equation") and exterior fields.

We will study the equations with an action of the general form :

$\int_\Omega \left(\sum_{\alpha\beta} C_G \left(\mathcal{F}_{r\alpha\beta}^t \mathcal{F}_r^{\alpha\beta} - \mathcal{F}_{w\alpha\beta}^t \mathcal{F}_w^{\alpha\beta} \right) + C_A \mathcal{F}_{A\alpha\beta}^t \mathcal{F}_A^{\alpha\beta} + C_{EM} \mathcal{F}_{EM\alpha\beta}^t \mathcal{F}_{EM}^{\alpha\beta} \right) \varpi_4$

$+ \sum_{p=1}^N \int_0^T C_I \operatorname{Im} \left\langle \psi_p, \nabla_{V_p} \psi_p \right\rangle \sqrt{-\langle V_p, V_p \rangle} dt$

The particles are assumed to be of a type defined by a fundamental state $\psi_{0p}, p = 1...N$ so that :

$\psi_p(t) = \vartheta(\sigma_p(t), \varkappa_p(t)) \psi_{0p}$

There is no need for a density.

We will use a detailed version of the interactions in the state equation (ψ) only, as it is only there that $r(t), w(t)$ are involved :

$C_I \operatorname{Im} \left\langle \psi_p, \nabla_{V_p} \psi_p \right\rangle = C_I \sum_{\alpha=0}^3 V_{\alpha p} \left(k_p^t X_{\alpha p} + \grave{A}_\alpha \right) = C_I \left(k^t \widehat{X}_p + \widehat{A}_p \right)$

Of course there is no longer a continuity equation.

7.2.1 Equations for the particles

The variables ψ_p are involved in the last integral only :

$\int_0^T C_I \sum_{\alpha=0}^3 V^\alpha \left(k^t X_{\alpha p} + \grave{A}_\alpha \right) \sqrt{-\langle V_p, V_p \rangle} dt$

so the equations can be deduced from the Euler-Lagrange equations with the variables $r_p(t), w_p(t)$.

All the variables are valued at the location on their trajectories given by t, and specific to each particle. We will drop the index p.

X has the same definition as in the previous model.

Equation for r

The equations are :

$\forall a = 1,2,3 : \sqrt{-\langle V, V \rangle} \frac{dL_I}{dr_a} = \frac{d}{dt} \left(\frac{dL_I}{d\frac{dr_a}{dt}} \sqrt{-\langle V, V \rangle} \right)$

With :

$\frac{dL_I}{dr_a} = \frac{d}{dr_a} \left(C_I \left(k^t \widehat{X} + \widehat{A} \right) \right) = C_I k^t \frac{d\widehat{X}}{dr_a}$

$\frac{dL_I}{d\frac{dr_a}{dt}} = \frac{d}{d\frac{dr_a}{dt}} \left(C_I \left(k^t \widehat{X} + \widehat{A} \right) \right) = C_I k^t \frac{d\widehat{X}}{d\frac{dr_a}{dt}} = C_I \left(k^t [D]^t \varepsilon_a \right)$

the equations read :

$C_I k^t \frac{d\widehat{X}}{dr_a} = \frac{1}{\sqrt{-\langle V,V\rangle}} \frac{d}{dt}\left(C_I \left(k^t [D]^t \varepsilon_a\right) \sqrt{-\langle V,V\rangle}\right)$

$= C_I \left(k^t [D]^t \varepsilon_a\right) \frac{1}{\sqrt{-\langle V,V\rangle}} \frac{d}{dt}\left(\sqrt{-\langle V,V\rangle}\right) + \frac{d}{dt}\left(C_I \left(k^t [D]^t \varepsilon_a\right)\right)$

$\frac{1}{\sqrt{-\langle V,V\rangle}} \frac{d}{dt}\left(\sqrt{-\langle V,V\rangle}\right) = -\frac{4a_w}{(2a_w^2-1)^2} \frac{da_w}{dt}\left(2a_w^2 - 1\right) = -\frac{4a_w}{2a_w^2-1}\frac{da_w}{dt} = -\frac{1}{2a_w^2-1}\left(w^t \frac{dw}{dt}\right)$

$C_I k^t \frac{d\widehat{X}}{dr_a} = \frac{d}{dt}\left(C_I \left(k^t [D]^t \varepsilon_a\right)\right) - C_I \left(k^t [D]^t \varepsilon_a\right) \frac{1}{2a_w^2-1}\left(w^t \frac{dw}{dt}\right)$

$k^t \frac{d\widehat{X}}{dr_a} = \varepsilon_a^t \left(-\frac{1}{2a_w^2-1}\left(w^t \frac{dw}{dt}\right)[D]\,k + \frac{d}{dt}\left([D]\,k\right)\right)$

The quantities have been computed in the previous section.

$\frac{d}{dr_a}\left(k^t \widehat{X}\right) = \varepsilon_a^t \{\left(\left[\delta D^t (r,k)\right] - \frac{1}{4a_r} r k^t \frac{\partial D^t}{\partial a_r}\right) \frac{dr}{dt}$

$+ \left(\left[\delta C^t (r,k)\right] - \frac{1}{4a_r} r k^t \frac{\partial C^t}{\partial a_r}\right)\left(\frac{1}{2}\left[j(w)\right]\frac{dw}{dt} + \left[1 - \frac{1}{2} j(w) j(w)\right]\widehat{G}_r + a_w \left[j(w)\right]\widehat{G}_w\right)\}$

$\frac{dD}{dt}k = \left(-\frac{1}{4a_r}\frac{\partial D}{\partial a_r} k r^t + \left[\delta D^t (r,k)\right]^t\right)\frac{dr}{dt}$

The equation reads :

$\varepsilon_a^t \{\left(\left[\delta D^t (r,k)\right] - \frac{1}{4a_r} r k^t \frac{\partial D^t}{\partial a_r}\right) \frac{dr}{dt}$

$+ \left(\left[\delta C^t (r,k)\right] - \frac{1}{4a_r} r k^t \frac{\partial C^t}{\partial a_r}\right)\left(\frac{1}{2}\left[j(w)\right]\frac{dw}{dt} + \left[1 - \frac{1}{2} j(w) j(w)\right]\widehat{G}_r + a_w \left[j(w)\right]\widehat{G}_w\right)\}$

$= -\frac{1}{2a_w^2-1}\left(\varepsilon_a^t [D]\,k\right)\left(w^t \frac{dw}{dt}\right) + \varepsilon_a^t \left(-\frac{1}{4a_r}\frac{\partial D}{\partial a_r} k r^t + \left[\delta D^t (r,k)\right]^t\right)\frac{dr}{dt}$

$\left(\left[\delta D^t (r,k)\right] - \left[\delta D^t (r,k)\right]^t + \frac{1}{4a_r}\frac{\partial D}{\partial a_r} k r^t - \frac{1}{4a_r} r k^t \frac{\partial D^t}{\partial a_r}\right)\frac{dr}{dt}$

$+ \left(\left(\left[\delta C^t (r,k)\right] - \frac{1}{4a_r} r k^t \frac{\partial C^t}{\partial a_r}\right)\frac{1}{2}\left[j(w)\right] + \frac{1}{2a_w^2-1}[D]\,k w^t\right)\frac{dw}{dt}$

$+ \left(\left[\delta C^t (r,k)\right] - \frac{1}{4a_r} r k^t \frac{\partial C^t}{\partial a_r}\right)\left(\left[1 - \frac{1}{2} j(w) j(w)\right]\widehat{G}_r + a_w \left[j(w)\right]\widehat{G}_w\right) = 0$

$$[A_1(r)]\frac{dr}{dt} + \left([A_2(r)]\frac{1}{2}[j(w)] + [D(r)]\frac{1}{2a_w^2-1}\left(j(w)j(k) + k^t w\right)\right)\frac{dw}{dt}$$

$$+ [A_2(r)]\left(\left[1 - \frac{1}{2}j(w)j(w)\right]\widehat{G}_r + a_w [j(w)]\widehat{G}_w\right) = 0$$

with the same matrices as above. The difference is the term $[D(r)]\frac{1}{2a_w^2-1}\left(j(w)j(k) + k^t w\right)$ in the coefficient of $\frac{dw}{dt}$.

Equation for w

$\sqrt{-\langle V,V\rangle} = \frac{c}{2a_w^2-1}$

We have similarly the equation :

$\forall a = 1,2,3 : \sqrt{-\langle V,V\rangle}\frac{dL_I}{dw_a} = \frac{d}{dt}\left(\frac{dL_I}{d\frac{dw_a}{dt}}\sqrt{-\langle V,V\rangle}\right)$

with :

$\frac{dL_I}{dw_a} = \frac{d}{dw_a}\left(\sum_{\alpha=0}^3 V^\alpha C_I \left(k^t X_\alpha + \grave{A}_\alpha\right)\right) = C_I \left(\sum_{\alpha=0}^3 \frac{dV^\alpha}{dw_a}\left(k^t X_\alpha + \grave{A}_\alpha\right) + V^\alpha \frac{dk^t X_\alpha}{dw_a}\right)$

$\sum_{\alpha=0}^3 \frac{dV^\alpha}{dw_a}\left(k^t X_\alpha + \grave{A}_\alpha\right)$

$= -\varepsilon_a^t \frac{1}{4a_w^2}\frac{2a_w^2+1}{2a_w^2-1} w\left(k^t \widehat{X} + \widehat{\grave{A}} - c\left(k^t X_0 + \grave{A}_0\right)\right) + \varepsilon_a^t \sum_{\alpha=1}^3 \sum_{i=1}^3 \varepsilon_i \frac{c a_w}{2a_w^2-1} P_i^\alpha \left(k^t X_\alpha + \grave{A}_\alpha\right)$

$\sum_{\alpha=0}^3 V^\alpha k^t \frac{dX_\alpha}{dw_a}$

$= -\varepsilon_a^t \{ \frac{1}{2} j(Ck) \frac{dw}{dt} + \frac{1}{2} (j(w) j(Ck) - 2j(Ck) j(w)) \widehat{G}_r + \left(2a_w - \frac{1}{a_w} + \frac{1}{4a_w} j(w) j(w) \right) j(Ck) \widehat{G}_w \}$

$\frac{dL_I}{d\frac{dw_a}{dt}} = \frac{1}{2} C_I k^t C^t(r) j(w) \varepsilon_a = -\frac{1}{2} C_I \varepsilon_a^t j(w) C(r) k$

The equation is :

$-\varepsilon_a^t \frac{1}{4a_w^2} \frac{2a_w^2+1}{2a_w^2-1} w \left(k^t \widehat{X} + \widehat{\grave{A}} - c \left(k^t X_0 + \grave{A}_0 \right) \right) + \varepsilon_a^t \sum_{\alpha=1}^3 \sum_{i=1}^3 \varepsilon_i \frac{ca_w}{2a_w^2-1} P_i^\alpha \left(k^t X_\alpha + \grave{A}_\alpha \right)$

$-\varepsilon_a^t \{ \frac{1}{2} j(Ck) \frac{dw}{dt} + \frac{1}{2} (j(w) j(Ck) - 2j(Ck) j(w)) \widehat{G}_r + \left(2a_w - \frac{1}{a_w} + \frac{1}{4a_w} j(w) j(w) \right) j(Ck) \widehat{G}_w \}$

$= -\frac{1}{2} \varepsilon_a^t \frac{d}{dt} (j(w) Ck) - \frac{1}{2a_w^2-1} \varepsilon_a^t \left(w^t \frac{dw}{dt} \right) \left(-\frac{1}{2} j(w) Ck \right)$

$\frac{d}{dt} (j(w) [C(r)] k) = j(w) \left(-\frac{1}{4a_r} \frac{\partial C}{\partial a_r} k r^t + [\delta C^t(r,k)]^t \right) \frac{dr}{dt} - j(Ck) \frac{dw}{dt}$

That is :

$$-\frac{1}{2} j(w) [A_2(r)]^t \frac{dr}{dt} + [A_3(r)] \frac{1}{2a_w^2 - 1} \left(1 - \frac{1}{2} j(w) j(w) \right) \frac{dw}{dt}$$

$$+\frac{1}{2} (j(w) [A_3(r)] - 2 [A_3(r)] j(w)) \widehat{G}_r + \left(2a_w - \frac{1}{a_w} + \frac{1}{4a_w} j(w) j(w) \right) [A_3(r)] \widehat{G}_w$$

$$= -\frac{1}{4a_w^2} \frac{2a_w^2+1}{2a_w^2-1} w \left(k^t \widehat{X} + \widehat{\grave{A}} - c \left(k^t X_0 + \grave{A}_0 \right) \right) + \frac{ca_w}{2a_w^2 - 1} \sum_{\alpha,i=1}^3 \varepsilon_i P_i^\alpha \left(k^t X_\alpha + \grave{A}_\alpha \right)$$

with the same matrices as above. The only difference is in the coefficient of $\frac{dw}{dt}$.

We will see that, on shell : $k^t \widehat{X} + \widehat{\grave{A}} = 0$.

7.2.2 Equation for the fields

The equation for the fields is computed by the method of variational derivative. We use the more general expression for the interactions $C_I \frac{1}{i} \langle \psi, \nabla_V \psi \rangle$. This term vanishes at the points where there is no particle.

Let us consider a variation $\delta \grave{A}_\alpha^a$ of $\grave{A}_\alpha^a$, given by a compactly supported map (so it has a value anywhere, null outside its support).

The functional derivative of the first integral is (Maths.2601), with $d\xi = d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$:

$\frac{\delta}{\delta \grave{A}_\alpha^a} \left(\int_\Omega C_A \sum_{\alpha\beta} \mathcal{F}_{A\alpha\beta}^t \mathcal{F}_A^{\alpha\beta} \varpi_4 \right) \left(\delta \grave{A}_\alpha^a \right)$

$= \int_\Omega C_A \left(\frac{\partial}{\partial \grave{A}_\alpha^a} \sum_{\alpha\beta} \mathcal{F}_{A\alpha\beta}^t \mathcal{F}_A^{\alpha\beta} \det P' - \sum_\beta \frac{d}{d\xi^\beta} \frac{\partial}{\partial \partial_\beta \grave{A}_\alpha^a} \left(\sum_{\alpha\beta} \mathcal{F}_{A\alpha\beta}^t \mathcal{F}_A^{\alpha\beta} \det P' \right) \right) \left(\delta \grave{A}_\alpha^a \right) d\xi$

$= \int_\Omega C_A \left(8 \sum_\beta \left[\grave{A}_\beta, \mathcal{F}_A^{\alpha\beta} \right]^a \det P' - \sum_\beta \frac{d}{d\xi^\beta} \left(-4 C_A \mathcal{F}_A^{a\alpha\beta} \det P' \right) \right) \left(\delta \grave{A}_\alpha^a \right) d\xi$

$= C_A \int_\Omega \left(8 \sum_\beta \left[\grave{A}_\beta, \mathcal{F}_A^{\alpha\beta} \right]^a + \frac{1}{\det P'} \sum_\beta \frac{d}{d\xi^\beta} \left(4 \mathcal{F}_A^{a\alpha\beta} \det P' \right) \right) \left(\delta \grave{A}_\alpha^a \right) \varpi_4$

For the simple integral a direct computation gives the functional integral :

$\frac{\delta}{\delta \grave{A}_\alpha^a} \left(\sum_{p=1}^N \int_0^T C_I \frac{1}{i} \langle \psi_p, \nabla_{V_p} \psi_p \rangle \sqrt{-\langle V_p, V_p \rangle} dt \right) \left(\delta \grave{A}_\alpha^a \right)$

$= C_I \frac{1}{i} \sum_{p=1}^N \int_0^T \left\langle \psi_p, \psi_p V_p^\alpha \delta \grave{A}_\alpha^a (m_p) [\theta_a] \right\rangle \sqrt{-\langle V_p, V_p \rangle} dt$

$= C_I \frac{1}{i} \sum_{p=1}^N \int_0^T V_p^\alpha \delta \grave{A}_\alpha^a (m_p) \langle \psi_p, \psi_p [\theta_a] \rangle \sqrt{-\langle V_p, V_p \rangle} dt$

The equation reads :

$\forall \delta \grave{A}_\alpha^a$:

$C_A \int_\Omega \left(8\sum_\beta \left[\grave{A}_\beta, \mathcal{F}_A^{\alpha\beta}\right]^a + 4\frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{a\alpha\beta}\det P'\right)\right)\left(\delta\grave{A}_\alpha^a\right)\varpi_4$
$+\sum_{p=1}^N \int_0^T C_I \frac{1}{i}\left(V_p^\alpha \delta\grave{A}_\alpha^a(m_p)\right)\langle\psi_p, [\psi_p][\theta_a]\rangle\sqrt{-\langle V_p, V_p\rangle}dt = 0$

The equation holds for any compactly smooth $\delta\grave{A}_\alpha^a$. Take $\delta\grave{A}_\alpha^a$ null outside a small tube ∂C_p enclosing the trajectory of each particle. By shrinking ∂C_p the first integral converges to the integral along the trajectory, with its volume form $\sqrt{-\langle V_p, V_p\rangle}dt$:

$C_A \int_\Omega \left(8\sum_\beta \left[\grave{A}_\beta, \mathcal{F}_A^{\alpha\beta}\right]^a + 4\frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{a\alpha\beta}\det P'\right)\right)\left(\delta\grave{A}_\alpha^a\right)\varpi_4$
$\to C_A \int_0^T \left(8\sum_\beta \left[\grave{A}_\beta, \mathcal{F}_A^{\alpha\beta}\right]^a + 4\frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{a\alpha\beta}\det P'\right)\right)\left(\delta\grave{A}_\alpha^a(m_p(t))\right)\sqrt{-\langle V_p, V_p\rangle}dt$

and the equation reads :

$\forall\delta\grave{A}_\alpha^a$:

$C_A \int_0^T \left(8\sum_\beta \left[\grave{A}_\beta, \mathcal{F}_A^{\alpha\beta}\right]^{a\prime} + 4\frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{a\alpha\beta}\det P'\right)\right)\left(\delta\grave{A}_\alpha^a(m_p(t))\right)\sqrt{-\langle V_p, V_p\rangle}dt$
$+\int_0^T C_I \frac{1}{i}\left(\delta\grave{A}_\alpha^a(m_p(t))\right)V_p^\alpha\langle\psi_p, [\psi_p][\theta_a]\rangle\sqrt{-\langle V_p, V_p\rangle}dt = 0$

$\forall a, \alpha : C_A \sum_\beta 8\left[\grave{A}_\beta, \mathcal{F}_A^{\alpha\beta}\right]^a + C_I \frac{1}{i}V_p^\alpha\langle\psi_p, [\psi_p][\theta_a]\rangle + 4C_A\frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{\alpha\beta}\det P'\right) = 0$

$\sum_\beta \left[\mathcal{F}_A^{\alpha\beta}, \grave{A}_\beta\right]^a - \frac{C_I}{8C_A}\frac{1}{i}V_p^\alpha\langle\psi_p, [\psi_p][\theta_a]\rangle = \frac{1}{2}\frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{\alpha\beta}\det P'\right)$

$$\phi_A^\alpha - J_A^\alpha = \frac{1}{2}\frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{\alpha\beta}\det P'\right) \tag{7.13}$$

with

$$\phi_A^\alpha = \sum_{a=1}^m \sum_\beta \left[\mathcal{F}_A^{\alpha\beta}, \grave{A}_\beta\right]^a \overrightarrow{\theta}_a \tag{7.14}$$

$$J_{Ap}^\alpha = \frac{C_I}{8C_A}V_p^\alpha \sum_{a=1}^m \left\langle \psi_p, [\psi_p]\frac{1}{i}[\theta_a]\right\rangle \overrightarrow{\theta}_a \tag{7.15}$$

So we have the same equation as in the first model and μ disappears. We have similarly :

$\forall\alpha = 0, ...3 : \forall\alpha = 0, ...3 : \phi_G^\alpha - J_{Gp}^\alpha = \frac{1}{2}\frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\left(\mathcal{F}_G^{\alpha\beta}\det P'\right)$

$\phi_G^\alpha = \sum_{a=1}^6 \sum_{\beta=0}^3 \left[\mathcal{F}_r^{\alpha\beta}, G_{r\beta}\right]^a \overrightarrow{\kappa}_a$

$J_{Gp}^\alpha = \frac{C_I}{8C_G}\frac{1}{i}V_p^\alpha \upsilon\left(\langle\psi_p, [\gamma C(\overrightarrow{\kappa}_a)][\psi_p]\rangle, -\langle\psi_p, [\gamma C(\overrightarrow{\kappa}_{a+3})][\psi_p]\rangle\right)$

$\forall\alpha = 0...3 : -J_{EM}^\alpha = \frac{1}{2}\frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\mathcal{F}_{EM}^{\alpha\beta}$

$J_{EMp}^\alpha = \frac{1}{8}\frac{C_I}{C_{EM}}\mu V_p^\alpha\langle\psi_{0p}, \psi_{0p}\rangle$

These equations holds only on the trajectories : $m = m_p(t)$

7.2.3 Tetrad equation

We have to compute the functional derivative on both integrals. The terms related to the particles vanish in the vacuum.

The functional derivative reads for the first :

$\frac{\delta}{\delta P_i^\alpha}\left(\int_\Omega\left(\sum_{\lambda\mu} C_G\left(\mathcal{F}_{r\lambda\mu}^t \mathcal{F}_r^{\lambda\mu} - \mathcal{F}_{w\lambda\mu}^t \mathcal{F}_w^{\lambda\mu}\right) + C_A \mathcal{F}_{A\lambda\mu}^t \mathcal{F}_A^{\lambda\mu} + C_{EM}\mathcal{F}_{EM\lambda\mu}\mathcal{F}_{EM}^{\lambda\mu}\right)\varpi_4\right)(\delta P_i^\alpha)$
$= \int_\Omega 4\sum_{\theta\lambda\mu}\sum_j \eta^{ij}g^{\mu\lambda}P_j^\theta\{C_G\sum_{a=1}^3 \mathcal{F}_{r\alpha\mu}^a \mathcal{F}_{r\theta\lambda}^a - \mathcal{F}_{w\alpha\mu}^a \mathcal{F}_{w\theta\lambda}^a$
$+C_A\sum_{a=1}^m \mathcal{F}_{A\alpha\mu}^a \mathcal{F}_{A\theta\lambda}^a + C_{EM}\mathcal{F}_{EM\alpha\mu}\mathcal{F}_{EM\theta\lambda}\}(\delta P_i^\alpha)\varpi_4$

$-\int_\Omega \left(\sum_{\lambda\mu} C_G \left(\mathcal{F}^t_{r\lambda\mu}\mathcal{F}^{\lambda\mu}_r - \mathcal{F}^t_{w\lambda\mu}\mathcal{F}^{\lambda\mu}_w\right) + C_A \mathcal{F}^t_{A\lambda\mu}\mathcal{F}^{\lambda\mu}_A + C_{EM}\mathcal{F}_{EM\lambda\mu}\mathcal{F}^{\lambda\mu}_{EM}\right)\left(P'^i_\alpha\right)\left(\delta P^\alpha_i\right)\varpi_4$

(to account for the derivative with respect to det P')

For the part related to the interactions V is defined by $V^\alpha = \sum_{i=0}^3 P^\alpha_i V^i$ and $\sqrt{-\langle V_p, V_p\rangle} = \frac{c}{2a^2_{wp}-1}$

$C_I \frac{1}{i}\langle\psi, \nabla_V \psi\rangle = \sum_{\alpha,i=0}^3 C_I \frac{1}{i} V^i P^\alpha_i \langle\psi, \nabla_\alpha \psi\rangle$

$\frac{\delta}{\delta P^\alpha_i}\int_0^T \left(C_I \frac{1}{i}\langle\psi_p, \nabla_{V_p}\psi_p\rangle\right)\sqrt{-\langle V_p, V_p\rangle}dt$

$= C_I \frac{1}{i}\int_0^T V^i \langle\psi_p, \nabla_\alpha \psi_p\rangle \sqrt{-\langle V_p, V_p\rangle}dt$

Thus :

$\delta\mathcal{L} = \int_\Omega 4\sum_{\theta\lambda\mu}\sum_j \eta^{ij} g^{\mu\lambda} P^\theta_j \{C_G \sum_{a=1}^3 \mathcal{F}^a_{r\alpha\mu}\mathcal{F}^a_{r\theta\lambda} - \mathcal{F}^a_{w\alpha\mu}\mathcal{F}^a_{w\theta\lambda}$

$+C_A \sum_{a=1}^m \mathcal{F}^a_{A\alpha\mu}\mathcal{F}^a_{A\theta\lambda} + C_{EM}\mathcal{F}_{EM\alpha\mu}\mathcal{F}_{EM\theta\lambda}\}\left(\delta P^\alpha_i\right)\varpi_4$

$-\int_\Omega \left(\sum_{\lambda\mu} C_G \left(\mathcal{F}^t_{r\lambda\mu}\mathcal{F}^{\lambda\mu}_r - \mathcal{F}^t_{w\lambda\mu}\mathcal{F}^{\lambda\mu}_w\right) + C_A \mathcal{F}^t_{A\lambda\mu}\mathcal{F}^{\lambda\mu}_A + C_{EM}\mathcal{F}_{EM\lambda\mu}\mathcal{F}^{\lambda\mu}_{EM}\right)\left(P'^i_\alpha\right)\left(\delta P^\alpha_i\right)\varpi_4$

$+\sum_{p=1}^N C_I \frac{1}{i}\int_0^T V^i \langle\psi_p, \nabla_\alpha \psi_p\rangle \left(\delta P^\alpha_i\right)\sqrt{-\langle V_p, V_p\rangle}dt$

And the equation $\frac{\delta\mathcal{L}}{\delta P^\alpha_i}\left(\delta P^\alpha_i\right) = 0$ reads, for the solutions :

$\forall \delta P^\alpha_i$:

$\int_\Omega 4\sum_{\theta\lambda\mu}\sum_j \eta^{ij} g^{\mu\lambda} P^\theta_j \{C_G \sum_{a=1}^3 \mathcal{F}^a_{r\alpha\mu}\mathcal{F}^a_{r\theta\lambda} - \mathcal{F}^a_{w\alpha\mu}\mathcal{F}^a_{w\theta\lambda} + C_A \sum_{a=1}^m \mathcal{F}^a_{A\alpha\mu}\mathcal{F}^a_{A\theta\lambda} + C_{EM}\mathcal{F}_{EM\alpha\mu}\mathcal{F}_{EM\theta\lambda}\}\left(\delta P^\alpha_i\right)\varpi$

$-\int_\Omega \left(\sum_{\lambda\mu} C_G \left(\mathcal{F}^t_{r\lambda\mu}\mathcal{F}^{\lambda\mu}_r - \mathcal{F}^t_{w\lambda\mu}\mathcal{F}^{\lambda\mu}_w\right) + C_A \mathcal{F}^t_{A\lambda\mu}\mathcal{F}^{\lambda\mu}_A + C_{EM}\mathcal{F}_{EM\lambda\mu}\mathcal{F}^{\lambda\mu}_{EM}\right)\left(P'^i_\alpha\right)\left(\delta P^\alpha_i\right)\varpi_4$

$+\sum_{p=1}^N C_I \frac{1}{i}\int_0^T V^i \langle\psi_p, \nabla_\alpha \psi_p\rangle \left(\delta P^\alpha_i\right)\sqrt{-\langle V_p, V_p\rangle}dt = 0$

With the same reasoning as above :for each particle along its trajectory :

$\int_0^T 4\sum_{\theta\lambda\mu}\sum_j \eta^{ij} g^{\mu\lambda} P^\theta_j \{C_G \sum_{a=1}^3 \mathcal{F}^a_{r\alpha\mu}\mathcal{F}^a_{r\theta\lambda} - \mathcal{F}^a_{w\alpha\mu}\mathcal{F}^a_{w\theta\lambda}$

$+C_A \sum_{a=1}^m \mathcal{F}^a_{A\alpha\mu}\mathcal{F}^a_{A\theta\lambda} + C_{EM}\mathcal{F}_{EM\alpha\mu}\mathcal{F}_{EM\theta\lambda}\}\left(\delta P^\alpha_i\right)\sqrt{-\langle V_p, V_p\rangle}dt$

$-\int_0^T \left(\sum_{\lambda\mu} C_G \left(\mathcal{F}^t_{r\lambda\mu}\mathcal{F}^{\lambda\mu}_r - \mathcal{F}^t_{w\lambda\mu}\mathcal{F}^{\lambda\mu}_w\right) + C_A \mathcal{F}^t_{A\lambda\mu}\mathcal{F}^{\lambda\mu}_A + C_{EM}\mathcal{F}_{EM\lambda\mu}\mathcal{F}^{\lambda\mu}_{EM}\right)\left(P'^i_\alpha\right)\left(\delta P^\alpha_i\right)\sqrt{-\langle V_p, V_p\rangle}dt$

$+C_I \frac{1}{i}\int_0^T V^i \langle\psi_p, \nabla_\alpha \psi_p\rangle \left(\delta P^\alpha_i\right)\sqrt{-\langle V_p, V_p\rangle}dt = 0$

$4\sum_{\theta\lambda\mu}\sum_j \eta^{ij} g^{\mu\lambda} P^\theta_j \{C_G \sum_{a=1}^3 \mathcal{F}^a_{r\alpha\mu}\mathcal{F}^a_{r\theta\lambda} - \mathcal{F}^a_{w\alpha\mu}\mathcal{F}^a_{w\theta\lambda} + C_A \sum_{a=1}^m \mathcal{F}^a_{A\alpha\mu}\mathcal{F}^a_{A\theta\lambda} + C_{EM}\mathcal{F}_{EM\alpha\mu}\mathcal{F}_{EM\theta\lambda}\}$

$-\left(\sum_{\lambda\mu} C_G \left(\mathcal{F}^t_{r\lambda\mu}\mathcal{F}^{\lambda\mu}_r - \mathcal{F}^t_{w\lambda\mu}\mathcal{F}^{\lambda\mu}_w\right) + C_A \mathcal{F}^t_{A\lambda\mu}\mathcal{F}^{\lambda\mu}_A + C_{EM}\mathcal{F}_{EM\lambda\mu}\mathcal{F}^{\lambda\mu}_{EM}\right)\left(P'^i_\alpha\right)$

$+C_I \frac{1}{i} V^i \langle\psi_p, \nabla_\alpha \psi_p\rangle = 0$

The equation, by product with $P^\beta_i$and summation on i gives :

$4\sum_\gamma \{C_G \sum_{a=1}^3 \mathcal{F}^a_{r\alpha\gamma}\mathcal{F}^{a\beta\gamma}_r - \mathcal{F}^a_{w\alpha\gamma}\mathcal{F}^{a\beta\gamma}_w + C_A \sum_{a=1}^m \mathcal{F}^a_{A\alpha\gamma}\mathcal{F}^{a\beta\gamma}_A + C_{EM}\mathcal{F}_{EM\alpha\gamma}\mathcal{F}^{\beta\gamma}_{EM}\} + C_I \frac{1}{i} V^\beta \langle\psi_p, \nabla_\alpha \psi_p\rangle$

$= \delta^\beta_\alpha \left(\sum_{\lambda\mu}\left(C_G \sum_{a=1}^3 \left(\mathcal{F}^a_{r\lambda\mu}\mathcal{F}^{a\lambda\mu}_r - \mathcal{F}^a_{w\lambda\mu}\mathcal{F}^{a\lambda\mu}_w\right) + C_A \sum_{a=1}^m \mathcal{F}^a_{A\lambda\mu}\mathcal{F}^{a\lambda\mu}_A + C_{EM}\mathcal{F}_{EM\lambda\mu}\mathcal{F}^{\lambda\mu}_{EM}\right)\right)$

$\forall \alpha, \beta = 0...3 : C_I \frac{1}{i} V^\beta_p \langle\psi_p, \nabla_\alpha \psi_p\rangle + 4\sum_\gamma \{4C_G \left\langle \mathcal{F}_{G\alpha\gamma}, \mathcal{F}^{\beta\gamma}_G \right\rangle_{Cl} + C_A \left\langle \mathcal{F}_{A\alpha\gamma}, \mathcal{F}^{\beta\gamma}_A \right\rangle_{T_1 U} + C_{EM}\mathcal{F}_{EM\alpha\gamma}\mathcal{F}^{\beta\gamma}_{EM}\}$

$= \delta^\alpha_\beta \sum_{\lambda\mu}\left(4C_G \left\langle \mathcal{F}_{G\lambda\mu}, \mathcal{F}^{\lambda\mu}_G \right\rangle_{Cl} + C_A \left\langle \mathcal{F}_{A\lambda\mu}, \mathcal{F}^{\lambda\mu}_A \right\rangle_{T_1 U} + C_{EM}\mathcal{F}_{EM\lambda\mu}\mathcal{F}^{\lambda\mu}_{EM}\right)$

With $\alpha = \beta$ and summing :

$$\langle\psi_p, \nabla_V \psi_p\rangle = 0 \tag{7.16}$$

and as a consequence :

$$\forall\alpha,\beta=0...3:C_I\frac{1}{i}V_p^\beta\left\langle\psi_p,\nabla_\alpha\psi_p\right\rangle+4\sum_\gamma\{4C_G\left\langle\mathcal{F}_{G\alpha\gamma},\mathcal{F}_G^{\beta\gamma}\right\rangle_{Cl}+C_A\left\langle\mathcal{F}_{A\alpha\gamma},\mathcal{F}_A^{\beta\gamma}\right\rangle_{T_1U}+C_{EM}\mathcal{F}_{EM\alpha\gamma}\mathcal{F}_{EM}^{\beta\gamma}\} \quad (7.17)$$

$$=2\delta_\alpha^\beta\left(C_G\left\langle\mathcal{F}_G,\mathcal{F}_G\right\rangle+C_A\left\langle\mathcal{F}_A,\mathcal{F}_A\right\rangle+C_{EM}\left\langle\mathcal{F}_{EM},\mathcal{F}_{EM}\right\rangle\right) \quad (7.18)$$

$$k^t\widehat{X}_p+\widehat{\grave{A}}_p=0 \quad (7.19)$$

The equations are (up to the density μ) identical or similar to the previous model. They hold everywhere, but the terms related to the particles vanish in the vacuum, that is out of the trajectories.

7.2.4 Particle submitted to an external field

For each particle we have the set of 7 differential equations. If the field generated by the particles can be neglected, they give the behavior of the particle under a known field. Of particular interest is the case of bonded particles : $w=0$.

$[A_1(r)]\frac{dr}{dt}+[A_2(r)]\widehat{G}_r=0$

$[A_3(r)]\widehat{G}_w=\sum_{\alpha,i=1}^3\varepsilon_i cP_i^\alpha\left(k^tX_\alpha+\grave{A}_\alpha\right)$

$k^t[D(r)]^t\frac{dr}{dt}+k^t[C(r)]^t\widehat{G}_r+\widehat{\grave{A}}=0$

If the gravitational field can be neglected then the obvious solution has the form : $r=\lambda(t)k$ and, accounting for $a_r^2=1-\frac{1}{4}r^tr$ can be written :

$r(t)=k\frac{2}{K}\sin\varpi(t,x)$

$K^2=k^tk$

$a_r=\cos\varpi(t,x)$

We have :

$X_\alpha=\left[\frac{1}{\cos\varpi}-\left(\frac{1}{K}\sin\varpi\right)j(k)+\left(\frac{1}{K^2}\frac{\sin^2\varpi}{\cos\varpi}\right)j(k)j(k)\right]\partial_\alpha r$

$k^tX_\alpha=\frac{1}{\cos\varpi}k^t\partial_\alpha r$

$\sum_{\alpha,i=1}^3\varepsilon_i cP_i^\alpha\left(\frac{1}{\cos\varpi(t,x)}k^t\partial_\alpha r+\grave{A}_\alpha(t,x)\right)=0$ which, in the SR context, reads with $P_i^\alpha=\delta_i^\alpha$:

$\alpha=1,2,3:ck^t\partial_\alpha r=-(\cos\varpi(t,x))\frac{1}{c}\sum_{i=1}^3\grave{A}_i(t,x)$

The equation ruling the variation of ϖ with t is given by the last equation :

$\frac{\partial\varpi}{\partial t}=-\frac{1}{2K}\widehat{\grave{A}}$

The motion is a rotation with a speed $-\frac{1}{2K}\widehat{\grave{A}}$. This can be seen as a precession with gyromagnetic ratio $\gamma=-\frac{1}{2K}\sim-\frac{eg}{2m}$. The vector k can be seen as a magnetic moment. A more exact value can be computed by accounting for the field generated by the particle.

7.3 CURRENTS

The Noether currents are usually introduced through the equivariance of the Lagrange equations, by computing the effects of a change of gauge or chart on the lagrangian. This is exactly what we have done before, deducing some basic rules for the specification of the lagrangian, and identities which must be satisfied by the partial derivatives. Whenever the lagrangian is defined from geometric quantities these identities are met, and the Noether currents do not appear this way. But we have a more interesting, and more intuitive, view of the currents from the equations that we have computed with the perturbative lagrangian.

7.3.1 Definition

The equations for the force fields (usually called the equations of motion) are :

$\forall\alpha = 0, ...3$:

$\phi_G^\alpha - J_G^\alpha = \frac{1}{2}\frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\left(\mathcal{F}_G^{\alpha\beta}\det P'\right)$

$-J_{EM}^\alpha = \frac{1}{2}\frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\mathcal{F}_{EM}^{\alpha\beta}$

$\phi_A^\alpha - J_A^\alpha = \frac{1}{2}\frac{1}{\det P'}\frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{a\alpha\beta}\det P'\right)$

In the first model the equations hold everywhere in Ω. In the second model they hold only on the trajectories of the particles, however one can check that they hold also in the vacuum, with $J_G, J_A, J_{EM} = 0$. They give relations between components, but they have a geometric formulation, which is more illuminating. The quantities $J_G^\alpha, J_A^\alpha, J_{EM}^\alpha$ are proportional to V and are vectors. The quantities $\phi_G^\alpha, \phi_A^\alpha$ are, up to constant, the derivatives of the lagrangian $\frac{\partial L}{\partial \grave{A}_\alpha^a}, \frac{\partial L}{\partial G_\alpha^a}$ and, as such, are vectors (see covariance of lagrangians). Moreover all these quantities are valued in the Lie algebras. So, with the quantities on the left hand side of the equations, we can define, at any point $\varphi_o(t,x)$, the tensors :

$$\left[\begin{array}{l} \phi_G = \sum_{a=1}^6\sum_{\beta=0}^3\left[\mathcal{F}_r^{\alpha\beta}, G_{r\beta}\right]^a \overrightarrow{\kappa}_a \otimes \partial\xi_\alpha \in T_1 Spin(3,1)\otimes TM \\ \phi_A = \sum_{a=1}^m\sum_{\beta=0}^3\left[\mathcal{F}_A^{\alpha\beta}, \grave{A}_\beta\right]\overrightarrow{\theta}_a \otimes \partial\xi_\alpha \in T_1 U\otimes TM \end{array}\right] \quad (7.20)$$

For a continuous distribution of particles one can similarly define the tensors, at any point $\varphi_o(t,x)$:

$$\left[\begin{array}{l} J_G = \frac{C_I}{8C_G}\mu\frac{1}{i}\upsilon\left(\langle\psi,[\gamma C(\overrightarrow{\kappa}_a)][\psi]\rangle, -\langle\psi,[\gamma C(\overrightarrow{\kappa}_{a+3})][\psi]\rangle\right)\otimes V \in T_1 Spin(3,1)\otimes TM \\ J_A = \frac{C_I}{8C_A}\mu\sum_{a=1}^m\left\langle\psi,[\psi]\frac{1}{i}[\theta_a]\right\rangle\overrightarrow{\theta}_a\otimes V\in T_1U\otimes TM \\ J_{EM} = \frac{C_I}{8C_{EM}}\mu\langle\psi_0,\psi_0\rangle V\in TM \end{array}\right] \quad (7.21)$$

For a single particle these tensors read, along the trajectory $\varphi_o(t, x_p(t))$ of the particle :

$J_{Gp} = \frac{C_I}{8C_G}\frac{1}{i}\upsilon\left(\langle\psi_p,[\gamma C(\overrightarrow{\kappa}_a)][\psi_p]\rangle, -\langle\psi_p,[\gamma C(\overrightarrow{\kappa}_{a+3})][\psi_p]\rangle\right)\otimes V_p \in T_1 Spin(3,1)\otimes TM$

$J_{Ap} = \frac{C_I}{8C_A}\sum_{a=1}^m\left\langle\psi_p,[\psi_p]\frac{1}{i}[\theta_a]\right\rangle\overrightarrow{\theta}_a\otimes V_p\in T_1U\otimes TM$

$J_{EMp} = \frac{C_I}{8C_{EM}}\langle\psi_{0p},\psi_{0p}\rangle V_p\in TM$

$J_G, J_A, J_{EM}, \phi_G, \phi_A, \phi_{EM}$ are called the **currents** associated to the fields labeled by a.

There is one vector field for each value of a, 12 in the Standard Model for the other fields and 6 for gravitation. This latter result does not come from the choice of a connection different from the Levi-Civita connection ($\mathcal{F}$ corresponds to the Riemann tensor) but from the choice of the scalar product over the scalar curvature. In many ways this is more in accordance with the law of equivalence, and in a unified theory we would have strictly the same equations.

The currents ϕ_G, ϕ_A are defined at any point, including the vacuum. The currents J_G, J_A, J_{EM} for the particles have for support the trajectory of the particles, so they are non null whenever there is a particle (second model), or $\mu \neq 0$ (first model) and are clearly attached to the flow of the particles through V.

In a time reversal, given by the matrix

$T = \begin{bmatrix} 0 & i\sigma_0 \\ i\sigma_0 & 0 \end{bmatrix}$

particles are exchanged with antiparticles, and the scalar products in the currents take the opposite sign, so we have opposite currents.

For the EM field :

$\forall \alpha = 0...3 : -J^\alpha_{EM} = \frac{1}{2}\frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\mathcal{F}^{\alpha\beta}_{EM} \Leftrightarrow -\frac{C_I}{4C_{EM}}\mu \langle \psi_0, \psi_0 \rangle V^\alpha = \frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\mathcal{F}^{\alpha\beta}_{EM}$

is the second Maxwell equation in GR, and $-\frac{C_I}{4C_{EM}}\mu \langle \psi_0, \psi_0 \rangle = \mu_0 q$ with a universal constant μ_0 and a charge q.

The meaning of the currents is more obvious by rewriting the lagrangian with them. The interaction term in the lagrangian reads :

$\frac{1}{i}C_I\mu \langle \psi, \nabla_V \psi \rangle = C_I \mu k^t \left([C(r)]^t \left([D(r)] \frac{dr}{dt} + \frac{1}{2} j(w) \frac{dw}{dt} \right) \right) + \frac{1}{i} C_I \mu \left\langle \psi, [\psi] \left[\widehat{A} \right] \right\rangle$

$+ C_I \mu \frac{1}{i} \left\langle \psi, \gamma C \left(\widehat{G} \right) \psi \right\rangle + C_{EM}\mu \left\langle \psi_0, \psi_0 \widehat{A} \right\rangle$

For the gravitational field :

$C_I \mu \frac{1}{i} \left\langle \psi, \gamma C \left(\widehat{G} \right) \psi \right\rangle = C_I \mu \frac{1}{i} \left\langle \psi, \sum_{\beta=0}^3 V^\beta \gamma C \left(\upsilon \left(G_{r\beta}, G_{w\beta} \right) \right) \psi \right\rangle_{E\otimes F}$

$= C_I \mu \frac{1}{i} \left\langle \psi, \sum_{\beta=0}^3 V^\beta \sum_{a=1}^3 \gamma C \left(G^a_{r\beta} \overrightarrow{\kappa}_a + G^a_{w\beta} \overrightarrow{\kappa}_{a+3} \right) \psi \right\rangle_{E\otimes F}$

$= C_I \mu \frac{1}{i} \sum_{\beta=0}^3 V^\beta \sum_{a=1}^3 \left(G^a_{r\beta} \langle \psi, \gamma C(\overrightarrow{\kappa}_a) \psi \rangle_{E\otimes F} + G^a_{w\beta} \langle \psi, \gamma C(\overrightarrow{\kappa}_{a+3}) \psi \rangle_{E\otimes F} \right)$

$= 4C_I \mu \frac{1}{i} \sum_{\beta=0}^3 V^\beta \left\langle \upsilon \left(\langle \psi, \gamma C(\overrightarrow{\kappa}_a) \psi \rangle_{E\otimes F}, - \langle \psi, \gamma C(\overrightarrow{\kappa}_{a+3}) \psi \rangle_{E\otimes F} \right), G_\beta \right\rangle_{Cl}$

$= 32 C_G \sum_{\beta=0}^3 \left\langle J^\beta_G, G_\beta \right\rangle_{Cl}$

For the EM field :

$\frac{1}{i} C_I \mu \left\langle \psi, i\psi \widehat{\grave{A}} \right\rangle = 8 C_{EM} \sum_{\beta=0}^3 J^\beta_{EM} \grave{A}_\beta$

For the other fields :

$\frac{1}{i} \mu C_I \left\langle \psi, [\psi] \left[\widehat{\grave{A}} \right] \right\rangle = C_I \mu \left\langle \psi, [\psi] \sum_{\alpha=0}^3 \sum_{a=1}^m V^\alpha \grave{A}^a_\alpha \frac{1}{i} [\theta_a] \right\rangle$

$= C_I \mu \sum_{\alpha=0}^3 \sum_{a=1}^m \grave{A}^a_\alpha \left\langle \psi, [\psi] \sum_{\alpha=0}^3 \sum_{a=1}^m V^\alpha \frac{1}{i} [\theta_a] \right\rangle$

$= 8 C_A \sum_{\beta=0}^3 \left\langle J^\beta_A, \grave{A}_\beta \right\rangle_{T_1 U}$

$$C_I \mu \operatorname{Im} \langle \psi, \nabla_V \psi \rangle = C_I \mu k^t \left([C(r)]^t \left([D(r)] \frac{dr}{dt} + \frac{1}{2} j(w) \frac{dw}{dt} \right) \right) \quad (7.22)$$

$$+ 8 \sum_{\beta=0}^3 \left(4C_G \left\langle J^\beta_G, G_\beta \right\rangle_{Cl} + C_A \left\langle J^\beta_A, \grave{A}_\beta \right\rangle_{T_1 U} + C_{EM} J^\beta_{EM} \grave{A}_\beta \right) \quad (7.23)$$

This expression of the variation of energy of the particle, using the currents, is more familiar. The first term is the kinetic energy of the particle, and the others represent the action of the fields, through the coupling of the potential, in its usual meaning, with a current. For instance for the EM field $J_{EM} = \frac{C_I}{8C_A}\mu qV$ with the electric charge q. What is significant is that the same occurs with the gravitational field.

7.3.2 Computation of the currents

The currents for the particles

It depends on the Lie algebra. We will consider only the gravitational field, which is the most interesting.

$J_G \in T_1Spin(3,1) \otimes TM$ so it is expressed as : $\upsilon(X,Y) \otimes V$.

$\langle \psi, \gamma C(\overrightarrow{\kappa}_a)\psi\rangle_E = \langle \gamma C(\sigma)\psi_0, \gamma C(\overrightarrow{\kappa}_a)\gamma C(\sigma)\psi_0\rangle_E$

$= \left\langle \psi_0, \gamma C(\sigma)^{-1}\gamma C(\overrightarrow{\kappa}_a)\gamma C(\sigma)\psi_0\right\rangle$

$= \left\langle \psi_0, \gamma C\left(\sigma^{-1}\cdot\overrightarrow{\kappa}_a\cdot\sigma\right)\psi_0\right\rangle$

$= \langle \psi_0, \gamma C(\mathbf{Ad}_{\sigma^{-1}}\overrightarrow{\kappa}_a)\psi_0\rangle$

$= \left\langle \psi_0, \gamma C\left(\mathbf{Ad}_{\sigma_r^{-1}}\mathbf{Ad}_{\sigma_w^{-1}}\overrightarrow{\kappa}_a\right)\psi_0\right\rangle$

We have met several times these quantities (see Total connection).

$\mathbf{Ad}_{\sigma_r^{-1}}\mathbf{Ad}_{\sigma_w^{-1}}\overrightarrow{\kappa}_a = \upsilon(X_a,Y_a) = \sum_{b=1}^3 X_a^b\overrightarrow{\kappa}_b + Y_a^b\overrightarrow{\kappa}_{b+3}$

$\left\langle \psi_0, \gamma C\left(\mathbf{Ad}_{\sigma_r^{-1}}\mathbf{Ad}_{\sigma_w^{-1}}\overrightarrow{\kappa}_a\right)\psi_0\right\rangle$

$= \langle \psi_0, \gamma C(\upsilon(X_a,Y_a))\psi_0\rangle$

$\gamma C(\upsilon(X_a,Y_a)) = \frac{1}{2}\begin{bmatrix}\sigma(Y_a - iX_a) & 0\\ 0 & -\sigma(Y_a+iX_a)\end{bmatrix}$

$\gamma C(\upsilon(X_a,Y_a))[\psi_0] = \frac{1}{2}\begin{bmatrix}\sigma(Y_a - iX_a)\psi_R\\ -\epsilon i\sigma(Y_a+iX_a)\psi_R\end{bmatrix}$

$\langle \psi_0, \gamma C(\upsilon(X_a,Y_a))\psi_0\rangle$

$= \frac{1}{2}Tr\left(\psi_R^*(-\epsilon\sigma(Y_a+iX_a)\psi_R) - \epsilon i\psi_R^*(i\sigma(Y_a-iX_a)\psi_R)\right)$

$= -i\epsilon\sum_{b=1}^3 X_a^b Tr(\psi_R^*\sigma_b\psi_R) = ik^tX$

with the inertial tensor k.

X_a, Y_a are the components, in $T_1Spin(3,1)$ of $\mathbf{Ad}_{\sigma_r^{-1}}\mathbf{Ad}_{\sigma_w^{-1}}\overrightarrow{\kappa}_a$.

a = 1,2,3 :

$X_a = \sum_{b=1}^3 [C^tA\varepsilon_a]^b\overrightarrow{\kappa}_b = \sum_{b=1}^3 [C^tA]_a^b\overrightarrow{\kappa}_b$

$\langle \psi_0, \gamma C(\upsilon(X_a,Y_a))\psi_0\rangle = ik^tX_\alpha = i\sum_{b=1}^3 k^b[C^tA]_a^b = ik^t[C^tA]_a$

$J_G = \frac{C_I}{4C_G}\mu\sum_{a=1}^3 [k^tC^tA]_a V\otimes\overrightarrow{\kappa}^a$

a = 4,5, 6 :

$X_a = \sum_{b=1}^3 [C^tB\varepsilon_a]^b\overrightarrow{\kappa}_b = \sum_{b=1}^3 [C^tB]_a^b\overrightarrow{\kappa}_b$

$\langle \psi_0, \gamma C(\upsilon(X_a,Y_a))\psi_0\rangle = i\sum_{b=1}^3 X_a^bk_b = i\sum_{b=1}^3 k^b[C^tB]_a^b$

$J_G = \frac{C_I}{4C_G}\mu\sum_{a=1}^3 [k^tC^tB]_a V\otimes\overrightarrow{\kappa}^{a+3}$

Representing the vectors X_r, X_w in $\upsilon(X_r,X_w)$ by column matrices we have the simple expressions :

$$J_G = \frac{C_I}{8C_G}\mu\upsilon\left([A(w(t))][C(r(t))]k, -[B(w(t))][C(r(t))]k\right)\otimes V$$

with $[A(w(t))]^t = [A(w(t))], [B(w(t))]^t = -[B(w(t))]$

For a model with individual particles :
$J_{Gp} = \frac{C_I}{8C_G}\upsilon\left(\left[A\left(w_p\left(t\right)\right)\right]\left[C\left(r_p\left(t\right)\right)\right]k_p, -\left[B\left(w_p\left(t\right)\right)\right]\left[C\left(r_p\left(t\right)\right)\right]k_p\right)\otimes V_p$
The force exercised by the gravitational field on the particle is :
$\sum_{\beta=0}^3 32C_G\left\langle J_G^\beta, G_\beta\right\rangle_{Cl} = 32C_G\left\langle J_G, \sum_{\beta=0}^3 V^\beta G_\beta\right\rangle_{Cl} = 32C_G\left\langle J_G, \widehat{G}\right\rangle_{Cl}$
$= C_I\left(\left[\widehat{G}_r\right]\left[A\left(w\left(t\right)\right)\right]\left[C\left(r\left(t\right)\right)\right]k + \left[\widehat{G}_w\right]\left[B\left(w\left(t\right)\right)\right]\left[C\left(r\left(t\right)\right)\right]k\right)$
With
$w \simeq \left(1+\frac{1}{8}\frac{v^2}{c^2}\right)\frac{v}{c}$
$a_w \simeq 1+\frac{1}{8}\frac{v^2}{c^2}$
$J_G \simeq \frac{C_I}{8C_G}\upsilon\left(\left[C\left(r\left(t\right)\right)\right]k, -j\left(w\left(t\right)\right)\left[C\left(r\left(t\right)\right)\right]k\right)\otimes V$
so that usually the current : $J_G \simeq \frac{C_I}{8C_G}\upsilon\left(\left[C\left(r\left(t\right)\right)\right]k, 0\right)\otimes V \in T_1Spin\left(3\right)\otimes TM$ and for $r=0 : J_G \simeq \frac{C_I}{8C_G}\upsilon\left(k,0\right)\otimes V$. The translational motion (represented by w) has a non null effect, but which is usually very weak. As well as the component G_w.

The intensity of the coupling between the gravitational field (represented by the potential) and the particle, can be assessed through the scalar product $\left\langle J_G, J_G\right\rangle$, which can be computed with the scalar product on $T_1Spin\left(3,1\right)$.
$\left\langle \upsilon\left(\left[A\left(w_p\right)\right]^t\left[C\left(r_p\right)\right]k_p, \left[B\left(w_p\right)\right]^t\left[C\left(r_p\right)\right]k_p\right), \upsilon\left(\left[A\left(w_p\right)\right]^t\left[C\left(r_p\right)\right]k_p, \left[B\left(w_p\right)\right]^t\left[C\left(r_p\right)\right]k_p\right)\right\rangle_{Cl} =$
$= \frac{1}{4}\left(k^tC^tAA^tCk - k^tC^tB^tBCk\right)$
$= \frac{1}{4}k^tC^t\left(A^2+B^2\right)Ck$
$= \frac{1}{4}k^tC^tCk = \frac{1}{4}k^tk$
using the identities :
$A = A^t, B^t = -B$
$A^2+B^2 = I; AB = BA$
$CC^t = C^tC = I_3$

$$\left\langle \upsilon\left(\left[k^tC^tA\right], \left[k^tC^tB\right]\right), \upsilon\left(\left[k^tC^tA\right], \left[k^tC^tB\right]\right)\right\rangle_{Cl} = \frac{1}{4}k^tk \tag{7.24}$$

The currents for the fields

For the EM field the bracket is null, so $\phi_{EM} = 0$.

Gravitational field
$\phi_G^a = \sum_\beta\left[\mathcal{F}^{\alpha\beta}, G_\beta\right]^a\partial\xi_\alpha$
$= \sum_{\beta=1}^3\left[\mathcal{F}_G^{0\beta}, G_\beta\right]^a\partial\xi_0$
$+\left[\mathcal{F}_G^{10}, G_0\right]^a\partial\xi_1 + \left[\mathcal{F}_G^{12}, G_2\right]^a\partial\xi_1 + \left[\mathcal{F}_G^{13}, G_3\right]^a\partial\xi_1$
$+\left[\mathcal{F}_G^{20}, G_0\right]^a\partial\xi_2 + \left[\mathcal{F}_G^{21}, G_1\right]^a\partial\xi_2 + \left[\mathcal{F}_G^{23}, G_3\right]^a\partial\xi_2$
$+\left[\mathcal{F}_G^{30}, G_0\right]^a\partial\xi_3 + \left[\mathcal{F}_G^{31}, G_1\right]^a\partial\xi_3 + \left[\mathcal{F}_G^{32}, G_2\right]^a\partial\xi_3$
$\phi_G^0 = \frac{1}{\det P'}\sum_{\beta=1}^3\upsilon\left(j\left(\left[*\mathcal{F}_r^w\right]_\beta\right)G_{r\beta} - j\left(\left[*\mathcal{F}_w^w\right]_\beta\right)G_{w\beta}, j\left(\left[*\mathcal{F}_w^w\right]_\beta\right)G_{r\beta} + j\left(\left[*\mathcal{F}_r^w\right]_\beta\right)G_{w\beta}\right)$
$\phi_G^{a1} = -\left[\mathcal{F}^{01}, G_0\right]^a - \left[\mathcal{F}^{21}, G_2\right]^a + \left[\mathcal{F}^{13}, G_3\right]^a$
$\phi_G^1 = \frac{1}{\det P'}\{-\upsilon\left(j\left(\left[*\mathcal{F}_r^w\right]_1\right)G_{0r} - j\left(\left[*\mathcal{F}_w^w\right]_1\right)G_{0w}, j\left(\left[*\mathcal{F}_w^w\right]_1\right)G_{0r} + j\left(\left[*\mathcal{F}_r^w\right]_1\right)G_{0w}\right)$
$-\upsilon\left(j\left(\left[*\mathcal{F}_r^r\right]_3\right)G_{2r} - j\left(\left[*\mathcal{F}_w^r\right]_3\right)G_{2w}, j\left(\left[*\mathcal{F}_w^r\right]_3\right)G_{2r} + j\left(\left[*\mathcal{F}_r^r\right]_3\right)G_{2w}\right)$
$+\upsilon\left(j\left(\left[*\mathcal{F}_r^r\right]_2\right)G_{3r} - j\left(\left[*\mathcal{F}_w^r\right]_2\right)G_{3w}, j\left(\left[*\mathcal{F}_w^r\right]_2\right)G_{3r} + j\left(\left[*\mathcal{F}_r^r\right]_2\right)G_{3w}\right)\}$
$\phi_G^{a2} = -\left[\mathcal{F}^{02}, G_0\right]^a + \left[\mathcal{F}^{21}, G_1\right]^a - \left[\mathcal{F}^{32}, G_3\right]^a$
$\phi_G^2 = \frac{1}{\det P'}\{-\upsilon\left(j\left(\left[*\mathcal{F}_r^w\right]_2\right)G_{0r} - j\left(\left[*\mathcal{F}_w^w\right]_2\right)G_{0w}, j\left(\left[*\mathcal{F}_w^w\right]_2\right)G_{0r} + j\left(\left[*\mathcal{F}_r^w\right]_2\right)G_{0w}\right)$

$+\upsilon\left(j\left(\left[*\mathcal{F}_r^r\right]_3\right)G_{1r}-j\left(\left[*\mathcal{F}_w^r\right]_3\right)G_{1w},j\left(\left[*\mathcal{F}_w^r\right]_3\right)G_{1r}+j\left(\left[*\mathcal{F}_r^r\right]_3\right)G_{1w}\right)$
$-\upsilon\left(j\left(\left[*\mathcal{F}_r^r\right]_1\right)G_{3r}-j\left(\left[*\mathcal{F}_w^r\right]_1\right)G_{3w},j\left(\left[*\mathcal{F}_w^r\right]_1\right)G_{3r}+j\left(\left[*\mathcal{F}_r^r\right]_1\right)G_{3w}\right)\}$
$\phi_G^{a3}=-\left[\mathcal{F}^{03},G_0\right]^a-\left[\mathcal{F}^{13},G_1\right]^a+\left[\mathcal{F}^{32},G_2\right]^a$
$\phi_G^3=\frac{1}{\det P'}\{-\upsilon\left(j\left(\left[*\mathcal{F}_r^w\right]_3\right)G_{0r}-j\left(\left[*\mathcal{F}_w^w\right]_3\right)G_{0w},j\left(\left[*\mathcal{F}_w^w\right]_3\right)G_{0r}+j\left(\left[*\mathcal{F}_r^w\right]_3\right)G_{0w}\right)$
$-\upsilon\left(j\left(\left[*\mathcal{F}_r^r\right]_2\right)G_{1r}-j\left(\left[*\mathcal{F}_w^r\right]_2\right)G_{1w},j\left(\left[*\mathcal{F}_w^r\right]_2\right)G_{1r}+j\left(\left[*\mathcal{F}_r^r\right]_2\right)G_{1w}\right)$
$+\upsilon\left(j\left(\left[*\mathcal{F}_r^r\right]_1\right)G_{2r}-j\left(\left[*\mathcal{F}_w^r\right]_1\right)G_{2w},j\left(\left[*\mathcal{F}_w^r\right]_1\right)G_{2r}+j\left(\left[*\mathcal{F}_r^r\right]_1\right)G_{2w}\right)\}$

Other fields

$\phi_A^{a0}=\frac{1}{\det P'}\sum_{\beta=1}^3\left[\left[*\mathcal{F}_A^w\right]_\beta,\grave{A}_\beta\right]^a$

$\phi_A^{a1}=\frac{1}{\det P'}\left(-\left[\left[*\mathcal{F}_A^w\right]_1,\grave{A}_0\right]-\left[\left[*\mathcal{F}_A^r\right]_3,\grave{A}_2\right]+\left[\left[*\mathcal{F}_A^r\right]_2,\grave{A}_3\right]\right)^a$

$\phi_A^{a2}=\frac{1}{\det P'}\left(-\left[\left[*\mathcal{F}_A^w\right]_2,\grave{A}_0\right]+\left[\left[*\mathcal{F}_A^r\right]_3,\grave{A}_1\right]-\left[\left[*\mathcal{F}_A^r\right]_1,\grave{A}_3\right]^a\right)^a$

$\phi_A^{a3}=\frac{1}{\det P'}\left(-\left[\left[*\mathcal{F}_A^w\right]_3,\grave{A}_0\right]-\left[\left[*\mathcal{F}_A^r\right]_2,\grave{A}_1\right]+\left[\left[*\mathcal{F}_A^r\right]_1,\grave{A}_2\right]\right)^a$

7.3.3 Main theorem

The quantities : $\frac{1}{\det P'}\sum_\beta\partial_\beta\left(\mathcal{F}_G^{a\alpha\beta}\det P'\right),\frac{1}{\det P'}\sum_\beta\partial_\beta\left(\mathcal{F}_A^{a\alpha\beta}\det P'\right)$ on the right hand side of the equations are defined everywhere, do not depend on the presence of particles, and have a geometric interpretation. For any 2 form $\mathcal{F}$ the coefficients $\mathcal{F}^{a\alpha\beta}\det P'$ are the components of the Hodge dual :

$*\mathcal{F}^r=-\left(\mathcal{F}^{01}d\xi^3\wedge d\xi^2+\mathcal{F}^{02}d\xi^1\wedge d\xi^3+\mathcal{F}^{03}d\xi^2\wedge d\xi^1\right)\det P'$
$*\mathcal{F}^w=-\left(\mathcal{F}^{32}d\xi^0\wedge d\xi^1+\mathcal{F}^{13}d\xi^0\wedge d\xi^2+\mathcal{F}^{21}d\xi^0\wedge d\xi^3\right)\det P'$

The exterior differential $d(*\mathcal{F})$ is a 3 form.

$d(*\mathcal{F}^r)=\partial_0\left(\mathcal{F}^{01}\det P'\right)d\xi^0\wedge d\xi^2\wedge d\xi^3-\partial_0\left(\mathcal{F}^{02}\det P'\right)d\xi^0\wedge d\xi^1\wedge d\xi^3+\partial_0\left(\mathcal{F}^{03}\det P'\right)d\xi^0\wedge d\xi^1\wedge d\xi^2$
$+\left(\partial_1\left(\mathcal{F}^{01}\det P'\right)+\partial_2\left(\mathcal{F}^{02}\det P'\right)+\partial_3\left(\mathcal{F}^{03}\det P'\right)\right)d\xi^1\wedge d\xi^2\wedge d\xi^3$
$d(*\mathcal{F}^w)=\left(\partial_1\left(\mathcal{F}^{13}\det P'\right)-\partial_2\left(\mathcal{F}^{32}\det P'\right)\right)d\xi^0\wedge d\xi^1\wedge d\xi^2+\left(\partial_1\left(\mathcal{F}^{21}\det P'\right)+\partial_3\left(\mathcal{F}^{32}\det P'\right)\right)d\xi^0\wedge d\xi^1\wedge d\xi^3$
$+\left(\partial_2\left(\mathcal{F}^{21}\det P'\right)-\partial_3\left(\mathcal{F}^{13}\det P'\right)\right)d\xi^0\wedge d\xi^2\wedge d\xi^3$
$d*\mathcal{F}=\left(\partial_0\left(\mathcal{F}^{01}\det P'\right)+\partial_1\left(\mathcal{F}^{01}\det P'\right)+\partial_2\left(\mathcal{F}^{02}\det P'\right)+\partial_3\left(\mathcal{F}^{03}\det P'\right)\right)d\xi^1\wedge d\xi^2\wedge d\xi^3$
$+\left(\partial_1\left(\mathcal{F}^{21}\det P'\right)+\partial_3\left(\mathcal{F}^{32}\det P'\right)-\partial_0\left(\mathcal{F}^{02}\det P'\right)\right)d\xi^0\wedge d\xi^1\wedge d\xi^3$
$+\left(\partial_0\left(\mathcal{F}^{03}\det P'\right)+\left(\partial_1\left(\mathcal{F}^{13}\det P'\right)-\partial_2\left(\mathcal{F}^{32}\det P'\right)\right)\right)d\xi^0\wedge d\xi^1\wedge d\xi^2$
$+\left(\partial_0\left(\mathcal{F}^{01}\det P'\right)+\left(\partial_2\left(\mathcal{F}^{21}\det P'\right)-\partial_3\left(\mathcal{F}^{13}\det P'\right)\right)\right)d\xi^0\wedge d\xi^2\wedge d\xi^3$

$$d(*\mathcal{F})=\sum_{\alpha=0}^3(-1)^\alpha\left(\sum_{\beta=0}^3\partial_\beta\left(\mathcal{F}^{\alpha\beta}\det P'\right)\right)d\xi^0\wedge...\widehat{d\xi^\alpha}...\wedge d\xi^3 \tag{7.25}$$

where $\widehat{}$ means that the vector is skipped [1].

The inner product of the currents, which are vector fields, with the 4 form ϖ_4 are 3 forms, which read :

$\varpi_4\left(\phi_G^a\right)=i_{\phi_G^a}\varpi_4=\sum_{\alpha=0}^3(-1)^\alpha\left(\sum_\beta\left[\mathcal{F}_G^{\alpha\beta},G_\beta\right]^a\right)(\det P')\,d\xi^0\wedge...\widehat{d\xi^\alpha}...\wedge d\xi^3$
$\varpi_4\left(J_G^a\right)=i_{J_G^a}\varpi_4=\sum_{\alpha=0}^3(-1)^\alpha J_G^{a\alpha}(\det P')\,d\xi^0\wedge...\widehat{d\xi^\alpha}...\wedge d\xi^3$

[1] Beware. The exponent is α and not $\alpha-1$ because the vectors are labelled 0,1,2,3 and not 1,2,3,4. A legacy of decennium of notation.

$\varpi_4\left(\phi_A^a\right)=i_{\phi_A^a}\varpi_4=\sum_{\alpha=0}^3(-1)^\alpha\left(\sum_\beta\left[\mathcal{F}_A^{\alpha\beta},\grave{A}_\beta\right]^a\right)(\det P')\,d\xi^0\wedge...\widehat{d\xi^\alpha}...\wedge d\xi^3$
$\varpi_4\left(J_A^a\right)=i_{J_A^a}\varpi_4=\sum_{\alpha=0}^3(-1)^\alpha J_A^{a\alpha}(\det P')\,d\xi^0\wedge...\widehat{d\xi^\alpha}...\wedge d\xi^3$
$\varpi_4\left(J_{EM}\right)=i_{J_{EM}}\varpi_4=\sum_{\alpha=0}^3(-1)^\alpha J_{EM}^{a\alpha}(\det P')\,d\xi^0\wedge...\widehat{d\xi^\alpha}...\wedge d\xi^3$
So the equations can be written :
$\forall a=1...6: i_{\phi_G^a}\varpi_4-i_{J_G^a}\varpi_4=\frac{1}{2}d*\mathcal{F}_G^a$
$\forall a=1...m: i_{\phi_A^a}\varpi_4-i_{J_A^a}\varpi_4=\frac{1}{2}d*\mathcal{F}_A^a$
$-i_{J_{EM}}\varpi_4=\frac{1}{2}d*\mathcal{F}_{EM}$
$i_{B_G}\varpi_4, i_{J_G}\varpi_4, i_{B_A}\varpi_4, i_{J_A}\varpi_4, i_{J_{EM}}\varpi_4$ can be interpreted as the densities of the currents $\phi_G, J_G, \phi_A, J_A, J_{EM}$

We can also express the currents with the corresponding 1 form J^*,ϕ^*, by raising the indexes with g, and proceeding to the computations (Maths.1613) :
$J=\sum_\alpha J^\alpha\partial\xi_\alpha\to J^*=\sum_{\lambda\alpha}g_{\alpha\lambda}J^\lambda d\xi^\alpha=\sum_\alpha J_\alpha^*d\xi^\alpha$
$J^*\to *J^*=\sum_{\alpha,\beta=0}^3(-1)^\alpha g^{\alpha\beta}J_\beta^*(\det P')\,d\xi^0\wedge...\widehat{d\xi^\alpha}...\wedge d\xi^3$
$=\sum_{\alpha,\beta=0}^3(-1)^\alpha g^{\alpha\beta}g_{\beta\lambda}J^\lambda(\det P')\,d\xi^0\wedge...\widehat{d\xi^\alpha}...\wedge d\xi^3$
$=\sum_{\alpha=0}^3(-1)^\alpha J^\alpha(\det P')\,d\xi^0\wedge...\widehat{d\xi^\alpha}...\wedge d\xi^3$
$i_{J_G}\varpi_4=*J^*$
Thus the equations read :
$\forall a=1...6: *\phi_G^a-*J_G^{*a}=\frac{1}{2}d*\mathcal{F}_G^a$
$\forall a=1...m: *\phi_A^a-*J_A^{*a}=\frac{1}{2}d*\mathcal{F}_A^a$
$-*J_{RM}^*=\frac{1}{2}d*\mathcal{F}_{EM}$
We have the identities :
$**J_G^*=-J_G^*, **\phi_G^*=-\phi_G^*,$
$**J_A^*=-J_A^*, **\phi_A^*=-\phi_A^*,$
$**J_{EM}^*=-J_{EM}^*,$
$*d\circ(*\mathcal{F}_G)=\delta\mathcal{F}_G$
$*d\circ(*\mathcal{F}_A)=\delta\mathcal{F}_A$
$*d\circ(*\mathcal{F}_{EM})=\delta\mathcal{F}_{EM}$
where $-*d\circ(*\mathcal{F})=\delta\mathcal{F}$ is the **codifferential** (Maths.32.3.1).
The equations read equivalently :
$\phi_G^*-J_G^*=\frac{1}{2}\delta(\mathcal{F}_G)$
$\phi_A^*-J_A^*=\frac{1}{2}\delta(\mathcal{F}_A)$
$-J_{EM}^*=\frac{1}{2}\delta(\mathcal{F}_{EM})$
For the EM field : $J_{EM}^*=-\frac{1}{2}\delta\mathcal{F}_{EM}$ is the geometric expression of the second Maxwell equations in GR.

The codifferential reduces the order of a form by one. It is in some way the inverse operator of the exterior differential d. The codifferential is the adjoint of the exterior differential with respect to the scalar product of forms on TM (Maths.2498). For any 1-form λ on TM :
$\forall\lambda\in\Lambda_1TM: G_1(\lambda,\delta(\mathcal{F}))=G_2(d\lambda,\mathcal{F})$
So :
$\phi_G^*-J_G^*=\frac{1}{2}\delta(\mathcal{F}_G)\Rightarrow G_1(\lambda,\phi_G^*)=G_1(\lambda,J_G^*)+\frac{1}{2}G_1(\lambda,\delta(\mathcal{F}_G))=G_1(\lambda,J_G^*)+\frac{1}{2}G_2(d\lambda,\mathcal{F}_G)$
$J_A^*-\phi_A^*=\frac{1}{2}\delta(\mathcal{F}_A)\Rightarrow G_1(\lambda,\phi_A^*)=G_1(\lambda,J_A^*)+\frac{1}{2}G_1(\lambda,\delta(\mathcal{F}_A))=G_1(\lambda,J_A^*)+\frac{1}{2}G_2(d\lambda,\mathcal{F}_A)$
thus if $d\lambda=0$:
$G_1(\lambda,J_G^*)=G_1(\lambda,\phi_G^*)$
$G_1(\lambda,J_A^*)=G_1(\lambda,\phi_A^*)$
$G_1(\lambda,J_G^*)\varpi_4=G_1(\lambda,\phi_G^*)\varpi_4\Leftrightarrow *\lambda\wedge J_G^*=*\lambda\wedge\phi_G^*$
$G_1(\lambda,J_A^*)=G_1(\lambda,\phi_A^*)\Leftrightarrow *\lambda\wedge J_A^*=*\lambda\wedge\phi_A^*$

Take : $\lambda = df$ with any function $f \in C_1(M;\mathbb{R})$:
$*df = \sum_{\alpha,\beta=0}^{3} (-1)^{\alpha} g^{\alpha\beta} \partial_\beta f (\det P') \, d\xi^0 \wedge ...\widehat{d\xi^\alpha} ... \wedge d\xi^3$
$\left(\sum_\alpha J^*_\alpha d\xi^\alpha\right) \wedge \left(\sum_{\alpha,\beta=0}^{3} (-1)^{\alpha+1} g^{\alpha\beta} \partial_\beta f (\det P') \, d\xi^0 \wedge ...\widehat{d\xi^\alpha} ... \wedge d\xi^3\right)$
$= \left(\sum_\alpha \phi^*_\alpha d\xi^\alpha\right) \wedge \left(\sum_{\alpha,\beta=0}^{3} (-1)^{\alpha+1} g^{\alpha\beta} \partial_\beta f (\det P') \, d\xi^0 \wedge ...\widehat{d\xi^\alpha} ... \wedge d\xi^3\right)$
$\sum_{\alpha\beta} J^*_\alpha g^{\alpha\beta} \partial_\beta f = \sum_{\alpha\beta} \phi^*_\alpha g^{\alpha\beta} \partial_\beta f$
$\sum_\beta J^\beta \partial_\beta f = \sum_\beta \phi^\beta \partial_\beta f \Leftrightarrow f'(m) J = f'(m) \phi$
Take $f(m) = \xi^\alpha$ with $\alpha = 0, ...3$:
$f'(m) J = J_\alpha = f'(m) \phi = \phi_\alpha$
So:
$J_A = \phi_A; \ J_G = \phi_G$
As we had :
$\phi^*_G - J^*_G = \frac{1}{2}\delta(\mathcal{F}_G)$
$\phi^*_A - J^*_A = \frac{1}{2}\delta(\mathcal{F}_A)$
$\delta\mathcal{F}_A = 0; \ \delta\mathcal{F}_G = 0$
$*d \circ (*\mathcal{F}) = \delta\mathcal{F} = 0 \Rightarrow * * d \circ (*\mathcal{F}) = 0 = -d \circ (*\mathcal{F})$
$d * \mathcal{F}_A = 0; \ d * \mathcal{F}_G = 0$

The Laplacian is the differential operator : $\Delta = -(d\delta + \delta d)$ (Maths.2500), which does not change the order of a form.

Thus :
$\Delta\mathcal{F}_G = -(d\delta + \delta d)\mathcal{F}_G = -\delta d\mathcal{F}_G$
$\Delta\mathcal{F}_A = -(d\delta + \delta d)\mathcal{F}_A = -\delta d\mathcal{F}_A$
To sum up :

Theorem 109 *For the EM field*

$$\phi_{EM} = 0; J^*_{EM} = -\frac{1}{2}\delta\mathcal{F}_{EM} \tag{7.26}$$

$\frac{C_I}{4C_{EM}} \mu \langle\psi_0, \psi_0\rangle V^* = -\frac{1}{\det P'} \sum_\beta \partial_\beta \left(\mathcal{F}^{\alpha\beta}_{EM} \det P'\right)$
For the other fields :

$$\begin{bmatrix} J_A = \phi_A & J_G = \phi_G \\ d(*\mathcal{F}_A) = 0 & d(*\mathcal{F}_G) = 0 \\ \Delta\mathcal{F}_A = -\delta d\mathcal{F}_A & \Delta\mathcal{F}_G = -\delta d\mathcal{F}_G \end{bmatrix} \tag{7.27}$$

These equations come from the variation of the field, the state of particles being constant. They show that particles are the source of the fields. Usually they are called "equation of motion" but this name is inaccurate : the field is a free variable in their proof. However the model holds even when a known field is imposed, then the variable is the total field, internal + external, and if the field generated by the particles can be neglected then the equations above provide, as a first approximation, the motion of the particles, including their rotation.

The equations for the fields split, in two sets which have different physical meanings.

The strength of the field is the key variable for the propagation of the field, that is in the vacuum.

The potentials, $G, \grave{A}$ are the key variables for the interactions with particles.

The potential are deduced from the connection. The strength, which is a derivative of a connection, can be computed from the potential, but the converse is not true. This is a classic issue : if $\mathcal{F}_G$ is the strength of the potential G then $G+H$ will provide the same strength $\mathcal{F}_G$ if $dH + \sum_{\alpha,\beta} \upsilon(H_\alpha, H_\beta) = 0$. For instance take $h = h_0 \exp \tau \left(\sum_{\alpha=0}^3 \xi^\alpha X_\alpha \right) \in C(\Omega \times \mathbb{R}; Spin(3,1))$ with $X_\alpha \in T_1 Spin(3,1)$ fixed vectors such that $[X_\alpha, X_\beta] = 0$ then $H = (\partial_\alpha h|_{t=0})$ meets the condition above. In Electrodynamics this issue is solved by imposing additional constraints to the potential - the "gauge freedom".

The codifferential equation $\delta\mathcal{F} = 0$ holds everywhere. It suffices to determine $\mathcal{F}$ in the vacuum, with given initial conditions on the border.

The current equations $J^*_{EM} = \frac{1}{2}\delta\mathcal{F}_{EM}, \phi_A = J_A, \phi_G = J_G$ hold also everywhere. They are the genuine equations of motion, as they provide the force which is exercised by the field onto the particles (the equivalent of the Lorentz equations for the EM field). In this interaction the potential also is changed, and actually the current equations provide the value of the potential, from $\mathcal{F}$. So in a model which considers all interactions there is actually no "gauge freedom" in the usual meaning.

But there is an important difference between the two models which have been studied. In a physical model with individual particles the vacuum exists almost everywhere, $\mathcal{F}$ is defined by its propagation and $\delta\mathcal{F} = 0$, $\mathcal{F}$ is the dominant variable, then the smoothness assumptions are sufficient to define the potentials. In a model with a continuous distribution of matter there is no propagation of the field to speak of, and $\mathcal{F}$ is defined from the potentials with the additional condition $\delta\mathcal{F} = 0$ and the definition of $\mathcal{F}$ itself.

7.3.4 Codifferential Equation

The codifferential equation : $\delta\mathcal{F}_A = 0$; $\delta\mathcal{F}_G = 0$ holds at any point, for all fields but the EM field. It reads :

$\delta\mathcal{F}_G = 0 \Leftrightarrow d(*\mathcal{F}_G) = 0$

$d(*\mathcal{F}_G) = \sum_{a=1}^6 \sum_{\alpha=0}^3 (-1)^\alpha \left(\sum_{\beta=0}^3 \partial_\beta \left(\mathcal{F}_G^{a\alpha\beta} \det P' \right) \right) d\xi^0 \wedge ...\widehat{d\xi^\alpha}... \wedge d\xi^3 \otimes \overrightarrow{\kappa}_a$

$\Leftrightarrow \alpha = 0...3, a = 1..6 : \sum_{\beta=0}^3 \partial_\beta \left(\mathcal{F}_G^{a\alpha\beta} \det P' \right) = 0$

For the EM field the equation is the same in the vacuum : $\frac{1}{\det P'} \sum_\beta \partial_\beta \left(\mathcal{F}_{EM}^{\alpha\beta} \det P' \right) = 0$, so the results presented below hold for the EM in the vacuum.

We see now what can be deduced from this equation with the example of the gravitational field.

One form

$*\mathcal{F}_G$ is a closed form, one can extend the Poincaré's lemna for each component $*\mathcal{F}_G^a$: there is a one form $\mathcal{K} \in *_1(M; T_1 Spin(3,1))$ such that :

$$* \mathcal{F}_G = d\mathcal{K} \tag{7.28}$$

Any 2 form $\mathcal{F} \in \Lambda_2(M;\mathbb{R})$ can be expressed in the orthonormal basis $(\varepsilon^i)_{i=0}^3$:

$\mathcal{F} = \sum_{\alpha\beta} \mathcal{F}_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta = \sum_{ij} F_{ij} \varepsilon^i \wedge \varepsilon^j$ with $\varepsilon^i = \sum_{\alpha=0}^3 P'^i_\alpha d\xi^\alpha$

$[\mathcal{F}^w] = P'^0_0 [F^w][Q'] - ([F^w][P']_0)[P']^0 + [F^r] j([P']_0)[P']$

$[\mathcal{F}^r] = [F^w][Q'] j\left([P']^0\right) + (\det Q')[F^r][Q]^t$

with the obvious notations $[F^r], [F^w]$.
In the standard chart :
$[\mathcal{F}^w] = [F^w][Q']$
$[\mathcal{F}^r] = (\det Q')[F^r][Q]^t$
Then the Hodge dual reads :
$[*\mathcal{F}^r] = [\mathcal{F}^w]\left[g_3^{-1}\right]\det Q' = [F^w][Q']\left[g_3^{-1}\right]\det Q' = [F^w][Q'][Q][Q]^t\det Q' = [F^w][Q]^t\det Q'$
$[*\mathcal{F}^w] = -(\det Q)[\mathcal{F}^r][g_3] = -(\det Q)(\det Q')[F^r][Q]^t[g_3] = -[F^r][Q]^t[Q']^t[Q'] = -[F^r][Q']$
and the differential $d\mathcal{K}$ of a one form $\mathcal{K}$
$[d\mathcal{K}^r] = (\det Q')[dK^r][Q]^t$
$[d\mathcal{K}^w] = [dK^w][Q']$
with obvious notations $[d\mathcal{K}^r]_{1\times 3}, [d\mathcal{K}^w]_{1\times 3}$
The computation can be done for each component $*\mathcal{F}_G^a, a = 1...6$ .Thus :
$*\mathcal{F}_G = d\mathcal{K}$
$\Leftrightarrow$
$(\det Q')[d\mathcal{K}^r]^a[Q]^t = [F^w]^a[Q]^t\det Q'$
$[d\mathcal{K}^w]^a[Q'] = -[F^r]^a[Q']$
$[d\mathcal{K}^r]^a = [F^w]^a; [d\mathcal{K}^w]^a = -[F^r]^a$
and in matrix notation
$[d\mathcal{K}_r^r]_{3\times 3}, [d\mathcal{K}_w^r]_{3\times 3}, [d\mathcal{K}_r^w]_{3\times 3}, [d\mathcal{K}_w^w]_{3\times 3}$

$$\begin{bmatrix} [\mathcal{F}_r^w] = [d\mathcal{K}_r^r][Q'] \\ [\mathcal{F}_r^r] = -(\det Q')[d\mathcal{K}_r^w][Q]^t \\ [\mathcal{F}_w^w] = [d\mathcal{K}_w^r][Q'] \\ [\mathcal{F}_w^r] = -(\det Q')[d\mathcal{K}_w^w][Q]^t \end{bmatrix} \tag{7.29}$$

Chern Identity

The Chern identity :
$Tr\left([\mathcal{F}_r^r]^t[\mathcal{F}_r^w] - [\mathcal{F}_w^r]^t[\mathcal{F}_w^w]\right) = 0$
reads :
$Tr\left(-(\det Q')[Q][d\mathcal{K}_r^w]^t[d\mathcal{K}_r^r][Q'] + (\det Q')[Q][d\mathcal{K}_w^w]^t[d\mathcal{K}_w^r][Q']\right)$
$= (\det Q')\,Tr\left(-[Q]\left([d\mathcal{K}_r^w]^t[d\mathcal{K}_r^r] + [d\mathcal{K}_w^w]^t[d\mathcal{K}_w^r]\right)[Q']\right)$
$= (\det Q')\,Tr\left(-[d\mathcal{K}_r^w]^t[d\mathcal{K}_r^r] + [d\mathcal{K}_w^w]^t[d\mathcal{K}_w^r]\right)$
and sums up to
$Tr\left(-[d\mathcal{K}_r^w]^t[d\mathcal{K}_r^r] + [d\mathcal{K}_w^w]^t[d\mathcal{K}_w^r]\right) = 0$
$\sum_{p,a=1}^3\left([d\mathcal{K}_r^w]_a^p[d\mathcal{K}_r^r]_a^p - [d\mathcal{K}_w^w]_a^p[d\mathcal{K}_w^r]_a^p\right) = 0$
It is not difficult to check that for any differential of a one form :
$\sum_{p=1}^3\left([d\mathcal{K}^w]^p[d\mathcal{K}^r]^p\right) = 0$
is always met, so the Chern identity is always met.

PDE

Then, using the identities provided by $d(d\mathcal{K})=0$:
$\partial_1 [d\mathcal{K}^w]^a_3 - \partial_3 [d\mathcal{K}^w]^a_1 = \partial_0 [d\mathcal{K}^r]^a_2$
$\partial_3 [d\mathcal{K}^w]^a_2 - \partial_2 [d\mathcal{K}^w]^a_3 = \partial_0 [d\mathcal{K}^r]^a_1$
$\partial_2 [d\mathcal{K}^w]^a_1 - \partial_1 [d\mathcal{K}^w]^a_2 = \partial_0 [d\mathcal{K}^r]^a_3$
$\partial_3 [d\mathcal{K}^r]^a_3 + \partial_2 [d\mathcal{K}^r]^a_2 + \partial_1 [d\mathcal{K}^r]^a_1 = 0$
the codifferential equation is equivalent to the 24 PDE in $Q, \mathcal{F}_G$:

$$\begin{bmatrix} \partial_0 \left(\left([\mathcal{F}^w_r][Q]\right)_3\right) = -\partial_2 \left(\left([\mathcal{F}^r_r][Q']^t\right)_1 \det Q\right) + \partial_1 \left(\left([\mathcal{F}^r_r][Q']^t\right)_2 \det Q\right) \\ \partial_0 \left(\left([\mathcal{F}^w_r][Q]\right)_3\right) = -\partial_2 \left(\left([\mathcal{F}^r_r][Q']^t\right)_1 \det Q\right) + \partial_1 \left(\left([\mathcal{F}^r_r][Q']^t\right)_2 \det Q\right) \\ \partial_0 \left(\left([\mathcal{F}^w_r][Q]\right)_2\right) = -\partial_1 \left(\left([\mathcal{F}^r_r][Q']^t\right)_3 \det Q\right) + \partial_3 \left(\left([\mathcal{F}^r_r][Q']^t\right)_1 \det Q\right) \\ \partial_0 \left(\left([\mathcal{F}^w_r][Q]\right)_1\right) = -\partial_3 \left(\left([\mathcal{F}^r_r][Q']^t\right)_2 \det Q\right) + \partial_2 \left(\left([\mathcal{F}^r_r][Q']^t\right)_3 \det Q\right) \\ \partial_1 \left(\left([\mathcal{F}^w_r][Q]\right)_1\right) + \partial_2 \left(\left([\mathcal{F}^w_r][Q]\right)_2\right) + \partial_3 \left(\left([\mathcal{F}^w_r][Q]\right)_3\right) = 0 \end{bmatrix} \tag{7.30}$$

and similar equations for $[\mathcal{F}^w_w], [\mathcal{F}^r_w]$.

The orthonormal basis is defined up to an orthonormal transformation. We have represented this by the statement :
$\forall x \in \Omega_3(0), t \in [0,T] : [Q(\varphi_o(t,x))] = [Q(\varphi_o(0,x))][C_q(\varphi_o(t,x))]$ where $[C_q(\varphi_o(t,x))]$ and $[Q(\varphi_o(0,x))]$ is the known initial basis.
$\mathcal{F}_G$ has 36 degrees of freedom, so usually, even with the knowledge of $[C_q(\varphi_o(t,x))]$ these equations are not sufficient to compute $\mathcal{F}_G$.

Usually some symmetries are assumed in a model. They can easily be represented through a global change of gauge : $\mathcal{F}$ is valued in the adjoint bundle, and a physical local symmetry for a field implies that the components of $\mathcal{F}$ keep the same value in two gauges related by the symmetry. For instance a symmetry by spatial rotation σ_r implies :
$\mathcal{F}_G = \mathbf{Ad}_{\sigma_r} \mathcal{F}_G$
$[\mathcal{F}_{r\alpha\beta}(m)] = [C(r)][\mathcal{F}_{r\alpha\beta}(m)]$
$[\mathcal{F}_{w\alpha\beta}(m)] = [C(r)][\mathcal{F}_{w\alpha\beta}(m)]$
$[C(r)] = \left[1 + a_r j(r) + \frac{1}{2} j(r) j(r)\right] \in SO(3)$ so the only acceptable symmetries are axial symmetries, and
$[\mathcal{F}_{r\alpha\beta}(m)]^a = \lambda_{\alpha\beta}(m) r^a$
$[\mathcal{F}_{w\alpha\beta}(m)]^a = \mu_{\alpha\beta}(m) r^a$
$\mathcal{F}_G$ is a decomposable tensor : $\mathcal{F}_G = \widehat{\mathcal{F}} \otimes X$ with $X = \upsilon(r,r)$.
The decomposability of the tensor is also an assumption which is legitimate in the vacuum.

Decomposable tensor

If $\mathcal{F}_G$ can be considered as a decomposable tensor : $\mathcal{F}_G = \widehat{\mathcal{F}} \otimes X$:
$[\mathcal{F}^r_r] = [X_r]_{3\times1} \left[\widehat{\mathcal{F}}^r_G\right]_{1\times3} ; [\mathcal{F}^w_r] = [X_r]_{m\times1} \left[\widehat{\mathcal{F}}^w_G\right]_{1\times3}$
$[\mathcal{F}^r_w] = [X_w]_{3\times1} \left[\widehat{\mathcal{F}}^r_G\right]_{1\times3} ; [\mathcal{F}^w_w] = [X_w]_{3\times1} \left[\widehat{\mathcal{F}}^w_G\right]_{1\times3}$

then the PDE above have an obvious solution.

One one hand :

$$d * \widehat{\mathcal{F}} = 0 \Leftrightarrow *\widehat{\mathcal{F}} = d\widehat{\mathcal{K}} \tag{7.31}$$

with $\widehat{\mathcal{K}} \in \Lambda_1(M;\mathbb{R})$
$\left[\widehat{\mathcal{F}}^w\right] = \left[d\widehat{\mathcal{K}}^r\right][Q'] ; \left[\widehat{\mathcal{F}}^r\right] = -(\det Q')\left[d\widehat{\mathcal{K}}^w\right][Q]^t$
$\left[d\widehat{\mathcal{K}}^r\right] = \left[\widehat{\mathcal{F}}^w\right][Q] ; \left[d\widehat{\mathcal{K}}^w\right] = -(\det Q)\left[\widehat{\mathcal{F}}^r\right][Q']^t$
which gives the 4 scalar PDE :

$$\left[\begin{array}{l} \partial_1\left((\det Q)\left[\widehat{\mathcal{F}}^r\right][Q']^t\right)_3 - \partial_3\left((\det Q)\left[\widehat{\mathcal{F}}^r\right][Q']^t\right)_1 = -\partial_0\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_2 \\ \partial_3\left((\det Q)\left[\widehat{\mathcal{F}}^r\right][Q']^t\right)_2 - \partial_2\left((\det Q)\left[\widehat{\mathcal{F}}^r\right][Q']^t\right)_3 = -\partial_0\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_1 \\ \partial_2\left((\det Q)\left[\widehat{\mathcal{F}}^r\right][Q']^t\right)_1 - \partial_1\left((\det Q)\left[\widehat{\mathcal{F}}^r\right][Q']^t\right)_2 = -\partial_0\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_3 \\ \partial_3\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_3 + \partial_2\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_2 + \partial_1\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_1 = 0 \end{array}\right] \tag{7.32}$$

On the other hand :
$[\partial_0 X_r]\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_3 = -[\partial_2 X_r]\left(\left[\widehat{\mathcal{F}}^r\right][Q']^t\right)_1 \det Q + [\partial_1 X_r]\left(\left[\widehat{\mathcal{F}}^r\right][Q']^t\right)_2 \det Q$
$[\partial_0 X_r]\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_2 = -[\partial_1 X_r]\left(\left[\widehat{\mathcal{F}}^r\right][Q']^t\right)_3 \det Q + [\partial_3 X_r]\left(\left[\widehat{\mathcal{F}}^r\right][Q']^t\right)_1 \det Q$
$[\partial_0 X_r]\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_1 = -[\partial_3 X_r]\left(\left[\widehat{\mathcal{F}}^r\right][Q']^t\right)_2 \det Q + [\partial_2 X_r]\left(\left[\widehat{\mathcal{F}}^r\right][Q']^t\right)_3 \det Q$
$[\partial_1 X_r]\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_1 + [\partial_2 X_r]\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_2 + [\partial_3 X_r]\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_3 = 0$
$\Leftrightarrow$
with $[\partial X_r]_{3\times 3} = [\partial_\beta X^a]^{a=1,2,3}_{\beta=1,2,3}, [\partial_0 X_r]_{3\times 1} = [\partial_0 X^a]^{a=1,2,3}$
$[\partial_0 X_r]\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_3 = \left([\partial X_r] j\left([Q']\left[\widehat{\mathcal{F}}^r\right]^t\right)\right)_3 \det Q$
$[\partial_0 X_r]\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_2 = \left([\partial X_r] j\left([Q']\left[\widehat{\mathcal{F}}^r\right]^t\right)\right)_2 \det Q$
$[\partial_0 X_r]\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_1 = \left([\partial X_r] j\left([Q']\left[\widehat{\mathcal{F}}^r\right]^t\right)\right)_1 \det Q$
$[\partial X_r]_1\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_1 + [\partial X_r]_2\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_2 + [\partial X_r]_3\left(\left[\widehat{\mathcal{F}}^w\right][Q]\right)_3 = 0$
The last equation reads :
$[\partial X_r][Q]^t\left[\widehat{\mathcal{F}}^w\right]^t = 0$
and the 3 others about $[X]$:
$[\partial_0 X_r]\left[\widehat{\mathcal{F}}^w\right][Q]$
$= [\partial X_r] j\left([Q']\left[\widehat{\mathcal{F}}^r\right]^t\right)\det Q$
$= [\partial X_r]\left([Q']^t\right)^{-1} j\left(\widehat{\mathcal{F}}^r\right)([Q'])^{-1}\det(Q')\det Q$
$= [\partial X_r][Q]^t j\left(\widehat{\mathcal{F}}^r\right)[Q]$
with $j(Mx) = \left(M^{-1}\right)^t j(x) M^{-1}\det M$

$[\partial_0 X_r]\left[\widehat{\mathcal{F}}^w\right]=[\partial X_r][Q]^t j\left(\widehat{\mathcal{F}}^r\right)$

Using $d\widehat{\mathcal{K}}$:

$[\partial_0 X_r]\left[d\widehat{\mathcal{K}}^r\right][Q']$

$=-(\det Q')[\partial X_r][Q]^t j\left(\left[d\widehat{\mathcal{K}}^w\right][Q]^t\right)$

$=-(\det Q')[\partial X_r][Q]^t j\left([Q]\left[d\widehat{\mathcal{K}}^w\right]^t\right)$

$=-(\det Q')[\partial X_r][Q]^t[Q']^t j\left(\left[d\widehat{\mathcal{K}}^w\right]^t\right)[Q']\det Q$

$=-[\partial X_r] j\left(\left[d\widehat{\mathcal{K}}^w\right]^t\right)[Q']$

$[\partial X_r][Q]^t[Q']^t\left[d\widehat{\mathcal{K}}^r\right]^t=0$

$[\partial X_r]\left[d\widehat{\mathcal{K}}^r\right]^t=0$

We have similar equations for $[X_w]$. Overall we have 18 linear PDE for X, which has 6 degrees of freedom, with parameters $Q,\widehat{\mathcal{K}}$.

$$\left[\begin{array}{l}[\partial X_r]\left[d\widehat{\mathcal{K}}^r\right]^t=0\\ [\partial X_w]\left[d\widehat{\mathcal{K}}^r\right]^t=0\\ [\partial_0 X_r]\left[d\widehat{\mathcal{K}}^r\right]+[\partial X_r] j\left(\left[d\widehat{\mathcal{K}}^w\right]^t\right)=0\\ [\partial_0 X_w]\left[d\widehat{\mathcal{K}}^r\right]+[\partial X_w] j\left(\left[d\widehat{\mathcal{K}}^w\right]^t\right)=0\end{array}\right] \tag{7.33}$$

The current equations lead then, in the vacuum, to $[X_r]=\lambda[X_w]$ with a scalar constant λ.

Moreover, in the absence of particles $\pounds_V\widehat{\mathcal{F}}=0$ which implies that $d\widehat{\mathcal{K}}$ satisfies the equations :

$\sum_{\lambda=0}^3 -V^\alpha\partial_\alpha\left(-(\det Q')\left[d\widehat{\mathcal{K}}^w\right][Q]^t\right)=-(\det Q')\left[d\widehat{\mathcal{K}}^w\right][Q]^t\left([\partial v]^t-(div(v))I_3\right)-\left[d\widehat{\mathcal{K}}^r\right][Q'] j\left(\left[\partial V^0\right]\right)$

$\sum_{\lambda=0}^3 V^\alpha\partial_\alpha\left(\left[d\widehat{\mathcal{K}}^r\right][Q']\right)=-(\det Q')\left[d\widehat{\mathcal{K}}^w\right][Q]^t j(\partial_0 V)-\left[d\widehat{\mathcal{K}}^r\right][Q']\left(\partial_0 V^0+[\partial v]\right)$

Conclusion

In the second model, which is the most physical, the vacuum exists almost everywhere, there is an initial field coming from sources outside of the system, and $\mathcal{F}$ is the dominant variable. The gravitational field usually changes slowly with the location, so the assumption about the decomposability of $\mathcal{F}$ should be acceptable. Then the PDE equations above enable to compute the field $\mathcal{F}$ if the initial conditions and $[Q]$ are known. Moreover, as they hold for each field, they provide constraints for the tetrad itself. For the EM field the equation is the same in the vacuum, and actually $\mathcal{F}_{EM}=\widehat{\mathcal{F}}$ leads, in the prospect of unification, to suspect that, at least for the EM and gravitational fields, which have the similar characteristics, $\widehat{\mathcal{F}}$ is common to both.

In the second model there is, inside the system, no propagation to speak of, the leading variable is the potential, the first set of differential equations still holds, but not the assumption about the decomposability. The field is defined through the potential, and the equations above define the tetrad. And the EM field is usually null inside a conductive material.

7.3.5 Currents Equations

They read :

$J_A = \phi_A; J_G = \phi_G$

$J^*_{EM} = -\frac{1}{2}\delta\mathcal{F}_{EM}$

They represent the interactions fields /particles. For the particles the currents are clearly defined by their state (spinor, charge and motion), for the fields by the potentials ($\mathcal{F}_{EM} = d\grave{A}$). They can be seen either as providing the value of the fields created by the sources, or as giving the value of the force exercised by the fields onto the particles (the equivalent of the Lorentz equation). But the mechanisms and their representations are different in the 2 models.

In a physical model, with individual particles, there is a causal structure : with respect to the observer there are an incoming and an outgoing field. The incoming field is the field in the vacuum, $\mathcal{F}$ is either known or computed as above. The outgoing field results from the interaction, which changes the state of the particle (its motion) and the connection $\mathbf{G}, \grave{\mathbf{A}}$, hence the value of the potential. The strength $\mathcal{F}$ is the derivative of the connection, and reflects this change of the connection. In continuous models, as we have here, the connection is continuously differentiable and $\mathcal{F}$ is continuous : the incoming and outgoing field are equal, and define the change in the potential (and the motion through $\widehat{G}, \widehat{\grave{A}}$). Thus the current equations can be interpreted as if $\mathcal{F}$ was given. As the current equations hold also in the vacuum, one can conclude that this is $\mathcal{F}$ which is the true commanding variable : its value is either given (there is an external source) or computed as seen above, and it impacts the value of the potential and through it the motion, and this latter adjustment results both from the definition of $\mathcal{F}$ and the currents equations

However this mechanism is formal : it provides just a way to deal with the current equations and, as in all models based on the principle of least action, the PDE do not reflect a pointwise adjustment of physical quantities, but the conditions that the maps representing the variables must meet over all the area considered. Notably in the equations of the previous subsection the tetrad is a parameter, which is the key to a general adjustment.

In a model with a density μ of particles the causal structure is blurred. In the equilibrium the variables are smoothed out and simultaneously defined : the assumption about the decomposability of $\mathcal{F}$ does not hold any more, so the key variable is the potential from which is computed the value of $\mathcal{F}$. But in this picture, as noticed before, there is no propagation mechanism.

In this subsection we will see how to express, in a more efficient way, the current equations, for the gravitational field. According to the remarks above the value of $\mathcal{F}$ (and more precisely of its Hodge dual) will be considered as the parameter from which the value of the potential is deduced.

The equations

The currents related to the particles read :

- for a continuous matter field :

$J_G(\varphi_o(t,x)) = \frac{C_I}{8C_G}\mu v \left([A(w(t))]^t [C(r(t))] k, [B(w(t))]^t [C(r(t))] k \right) \otimes V$

for an individual particle :

$J_G(\varphi_o(t,x_p(t))) = \frac{C_I}{8C_G}\upsilon\left([A(w_p(t))]^t[C(r_p(t))]k,[B(w_p(t))]^t[C(r_p(t))]k\right)\otimes V_p$
with $V = \overrightarrow{v} + c\varepsilon_0(m) = c\left(\varepsilon_0 + \epsilon\frac{a_w}{2a_w^2-1}\sum_{i=1}^3 w_i\varepsilon_i\right) = c\left(\partial t + \epsilon\frac{a_w}{2a_w^2-1}\sum_{\alpha=1}^3\sum_{p=1}^3 w_p Q_p^\alpha\partial\xi_\alpha\right)$
We will denote in this section :
$\beta = 1,2,3 : v^\beta = \epsilon\frac{a_w}{2a_w^2-1}\sum_{\beta=1}^3\sum_{a=1}^3 w_a Q_a^\beta$
$q = \frac{C_I}{8C_G}\mu$ or $q = \frac{cC_I}{8C_G}$ which is assumed to vary with $\varphi_o(t,x)$.
thus :
$J_G^0(\varphi_o(t,x)) = q\upsilon\left([A(w(t))]^t[C(r(t))]k,[B(w(t))]^t[C(r(t))]k\right) = q\upsilon([A][C]k,-[B][C]k)$
$\beta = 1,2,3 : J_G^\beta(\varphi_o(t,x)) = v^\beta q\upsilon\left([A(w(t))]^t[C(r(t))]k,[B(w(t))]^t[C(r(t))]k\right) = v^\beta q\upsilon([A][C]k,-[B][C]k)$

Using the matrix notation :
the 3×3 matrices :
$[G_r]_{3\times3} = \left[G_{r\beta}^a\right]_{\beta=1,2,3}^{a=1,2,3}; [G_w]_{3\times3} = \left[G_{w\beta}^a\right]_{\beta=1,2,3}^{a=1,2,3}$
the 3×1 column matrices :
$[G_{r0}]_{3\times1} = [G_{r0}^a]^{a=1,2,3}; [G_{w0}]_{3\times1} = [G_{w0}^a]^{a=1,2,3}$
the 9×3 matrices :

$$[J_r^r] = \begin{bmatrix} j([*\mathcal{F}_r^r]_1) \\ j([*\mathcal{F}_r^r]_2) \\ j([*\mathcal{F}_r^r]_3)\end{bmatrix} : [J_w^r] = \begin{bmatrix} j([*\mathcal{F}_w^r]_1) \\ j([*\mathcal{F}_w^r]_2) \\ j([*\mathcal{F}_w^r]_3)\end{bmatrix}$$

$$[J_r^w] = \begin{bmatrix} j([*\mathcal{F}_r^w]_1) \\ j([*\mathcal{F}_r^w]_2) \\ j([*\mathcal{F}_r^w]_3)\end{bmatrix}; [J_w^w] = \begin{bmatrix} j([*\mathcal{F}_w^w]_1) \\ j([*\mathcal{F}_w^w]_2) \\ j([*\mathcal{F}_w^w]_3)\end{bmatrix}$$

with : $[J_r^r]^t = -\left[\begin{array}{ccc} j([*\mathcal{F}_r^r]_1) & j([*\mathcal{F}_r^r]_1) & j([*\mathcal{F}_r^r]_1)\end{array}\right]_{3\times9}, \ldots$
the 9×9 matrices :

$$[M_r^r] = \begin{bmatrix} 0 & -j([*\mathcal{F}_r^r]_3) & j([*\mathcal{F}_r^r]_2) \\ j([*\mathcal{F}_r^r]_3) & 0 & -j([*\mathcal{F}_r^r]_1) \\ -j([*\mathcal{F}_r^r]_2) & j([*\mathcal{F}_r^r]_1) & 0\end{bmatrix}_{9\times9}$$

and similar $[M_w^r], [M_r^w], [M_w^w]$

The equations $\phi_G^0 = J_G^0$
$\sum_{\beta=1}^3 j\left([*\mathcal{F}_r^w]_\beta\right)G_{r\beta} - j\left([*\mathcal{F}_w^w]_\beta\right)G_{w\beta} = q[A][C]k\det P'$
$\sum_{\beta=1}^3 j\left([*\mathcal{F}_w^w]_\beta\right)G_{r\beta} + j\left([*\mathcal{F}_r^w]_\beta\right)G_{w\beta} = -q[B][C]k\det P'$
read, in matricial form :

$$[J_r^w]^t\begin{bmatrix}[G_r]_1\\ [G_r]_2\\ [G_r]_3\end{bmatrix} - [J_w^w]^t\begin{bmatrix}[G_w]_1\\ [G_w]_2\\ [G_w]_3\end{bmatrix} = -q[A][C]k\det P'$$

$$[J_w^w]^t\begin{bmatrix}[G_r]_1\\ [G_r]_2\\ [G_r]_3\end{bmatrix} + [J_r^w]^t\begin{bmatrix}[G_w]_1\\ [G_w]_2\\ [G_w]_3\end{bmatrix} = q[B][C]k\det P'$$

For $\alpha = 1,2,3$ the equations $\phi_r^\alpha = J_r^\alpha$ read in matricial form :

$$[M_r^r]\begin{bmatrix}[G_r]_1\\ [G_r]_2\\ [G_r]_3\end{bmatrix} - [M_w^r]\begin{bmatrix}[G_w]_1\\ [G_w]_2\\ [G_w]_3\end{bmatrix} - [J_r^w][G_{r0}] + [J_w^w][G_{w0}] = \begin{bmatrix} v^1[A][C]k \\ v^2[A][C]k \\ v^3[A][C]k\end{bmatrix} q\det P'$$

For $\alpha = 1,2,3$ the equations $\phi_w^\alpha = J_w^\alpha$ read in matricial form :

$$[M_w^r]\begin{bmatrix}[G_r]_1\\ [G_r]_2\\ [G_r]_3\end{bmatrix}+[M_r^r]\begin{bmatrix}[G_w]_1\\ [G_w]_2\\ [G_w]_3\end{bmatrix}-[J_w^w][G_{r0}]-[J_r^w][G_{w0}]=-\begin{bmatrix}v^1[B][C]k\\ v^2[B][C]k\\ v^3[B][C]k\end{bmatrix}q\det P'$$

That we can sum up by :

$$\begin{bmatrix}[J_r^w]^t & -[J_w^w]^t\\ [J_w^w]^t & [J_r^w]^t\end{bmatrix}_{6\times 18}[G]_{18\times 1}=\begin{bmatrix}-[A][C]k\\ [B][C]k\end{bmatrix}_{6\times 1}q\det P'$$

$$\begin{bmatrix}[M_r^r] & -[M_w^r]\\ [M_w^r] & [M_r^r]\end{bmatrix}_{18\times 18}[G]_{18\times 1}+\begin{bmatrix}[J_r^w] & -[J_w^w]\\ [J_w^w] & [J_r^w]\end{bmatrix}_{18\times 6}\begin{bmatrix}[G_{r0}]\\ [G_{w0}]\end{bmatrix}_{6\times 1}=\begin{bmatrix}v\otimes[A][C]k\\ -v\otimes[B][C]k\end{bmatrix}q\det P'$$

with

$$[G]=\begin{bmatrix}[G_r]_1\\ [G_r]_2\\ [G_r]_3\\ [G_w]_1\\ [G_w]_2\\ [G_w]_3\end{bmatrix};$$

$$\begin{bmatrix}v^1[A][C]k\\ v^2[A][C]k\\ v^3[A][C]k\\ -v^1[B][C]k\\ -v^2[B][C]k\\ -v^3[B][C]k\end{bmatrix}_{18\times 1}=\begin{bmatrix}v\otimes[A][C]k\\ -v\otimes[B][C]k\end{bmatrix}$$ ($\otimes$ is the Kronecker's product of matrices)

Solutions

These 24 linear equations for the 24 components of G are quite intimidating, but the system can be solved explicitly, with parameters $\mathcal{F}$, r, w.

The matrices like $[M_r^r]$ are invertible if $\det[M_r^r]=\det[*\mathcal{F}_r^r]\neq 0$ (see Formulas) :

$$[M_r^r]^{-1}=\frac{1}{2\det[*\mathcal{F}_r^r]}\left([J_r^r][J_r^r]^t-\begin{bmatrix}[*\mathcal{F}_r^r]_1^t[*\mathcal{F}_r^r]_1 & [*\mathcal{F}_r^r]_2^t[*\mathcal{F}_r^r]_1 & [*\mathcal{F}_r^r]_3^t[*\mathcal{F}_r^r]_1\\ [*\mathcal{F}_r^r]_1^t[*\mathcal{F}_r^r]_2 & [*\mathcal{F}_r^r]_2^t[*\mathcal{F}_r^r]_2 & [*\mathcal{F}_r^r]_3^t[*\mathcal{F}_r^r]_2\\ [*\mathcal{F}_r^r]_1^t[*\mathcal{F}_r^r]_3 & [*\mathcal{F}_r^r]_2^t[*\mathcal{F}_r^r]_3 & [*\mathcal{F}_r^r]_3^t[*\mathcal{F}_r^r]_3\end{bmatrix}\right)$$

and similarly for $[M_w^r]$

Matrices like :

$M=\begin{bmatrix}A & -B\\ B & A\end{bmatrix}$ where A, B are invertible matrices,

are invertible if $R=4I+\left(A^{-1}B-B^{-1}A\right)^2$ is invertible (see Annex) :

$$\begin{bmatrix}A & -B\\ B & A\end{bmatrix}^{-1}=\begin{bmatrix}R^{-1}\left(A^{-1}+B^{-1}AB^{-1}\right) & R^{-1}\left(B^{-1}+A^{-1}BA^{-1}\right)\\ -R^{-1}\left(B^{-1}+A^{-1}BA^{-1}\right) & R^{-1}\left(A^{-1}+B^{-1}AB^{-1}\right)\end{bmatrix}$$

Thus $M=\begin{bmatrix}[M_r^r] & -[M_w^r]\\ [M_w^r] & [M_r^r]\end{bmatrix}$ is invertible if :

$\det[\mathcal{F}_r^r], \det[\mathcal{F}_w^r]\neq 0$

$\det\left(4I+\left([M_r^r]^{-1}[M_w^r]-[M_w^r]^{-1}[M_r^r]\right)^2\right)\neq 0$

Then the equations read :

$[G] = [M]^{-1}\left(-\begin{bmatrix} [J_r^w] & -[J_w^w] \\ [J_w^w] & [J_r^w] \end{bmatrix}\begin{bmatrix} [G_{r0}] \\ [G_{w0}] \end{bmatrix} + \begin{bmatrix} v\otimes[A][C]k \\ -v\otimes[B][C]k \end{bmatrix} q\det P'\right)$
and G_0 is computed by :
$\begin{bmatrix} [J_r^w]^t & -[J_w^w]^t \\ [J_w^w]^t & [J_r^w]^t \end{bmatrix}[M]^{-1}\begin{bmatrix} [J_r^w] & -[J_w^w] \\ [J_w^w] & [J_r^w] \end{bmatrix}\begin{bmatrix} [G_{r0}] \\ [G_{w0}] \end{bmatrix}$
$= \left(\begin{bmatrix} [J_r^w]^t & -[J_w^w]^t \\ [J_w^w]^t & [J_r^w]^t \end{bmatrix}[M]^{-1}\begin{bmatrix} v\otimes[A][C]k \\ -v\otimes[B][C]k \end{bmatrix} + \begin{bmatrix} [A][C]k \\ -[B][C]k \end{bmatrix}\right) q\det P'$

Moreover $\phi^0 = J_G^0$
$\Rightarrow \langle\phi^0,\phi^0\rangle_{Cl} = \langle J_G^0, J_G^0\rangle_{Cl} = \frac{1}{4}(k^tC^tAA^tCk - k^tC^tB^tBCk)(q\det P')^2 = \frac{1}{4}(k^tk)(q\det P')^2$
$[G]^t\begin{bmatrix} -\left([J_r^w][J_r^w]^t - [J_w^w][J_w^w]^t\right) & \left([J_r^w][J_w^w]^t + [J_w^w][J_r^w]^t\right) \\ \left([J_r^w][J_w^w]^t + [J_w^w][J_r^w]^t\right) & \left([J_r^w][J_r^w]^t - [J_w^w][J_w^w]^t\right) \end{bmatrix}[G] = -(k^tk)(q\det P')^2$

Decomposable tensor

If the tensor $\mathcal{F}$ is decomposable :
$[*\mathcal{F}_r^r] = [X_r]\left[d\widehat{\mathcal{K}}^r\right]; [*\mathcal{F}_r^w] = [X_r]\left[d\widehat{\mathcal{K}}^w\right]$
$[*\mathcal{F}_w^r] = [X_w]\left[d\widehat{\mathcal{K}}^r\right]; [*\mathcal{F}_w^w] = [X_w]\left[d\widehat{\mathcal{K}}^w\right]$
The equations $\phi_G^0 = J_G^0$ read :
$[j((X_w)G_r + j(X_r)G_w)]\left[d\widehat{\mathcal{K}}^w\right]^t = -q[B][C]k\det P'$
For $\alpha = 1,2,3$ the equations $\phi_r^\alpha = J_r^\alpha$ read :
$-\left\{(j(X_r)[G_r] - j(X_w)[G_w])j\left(d\widehat{\mathcal{K}}^r\right)\right\}_\alpha - \left\{(j(X_r)G_{0r} - j(X_w)G_{0w})\left[d\widehat{\mathcal{K}}^w\right]\right\}_\alpha = \{[A][C]kv^t\}_\alpha q\det P'$
For $\alpha = 1,2,3$ the equations $\phi_w^\alpha = J_w^\alpha$ read :
$-\left\{(j(X_w)G_r + j(X_r)G_w)j\left(d\widehat{\mathcal{K}}^r\right)\right\}_\alpha - \left\{(j(X_w)G_{0r} + j(X_r)G_{0w})\left[d\widehat{\mathcal{K}}^w\right]\right\}_\alpha = -\{[B][C]kv^t\}_\alpha q\det P'$
Which sums up to :
$[j(X_r)G_r - j(X_w)G_w]\left[d\widehat{\mathcal{K}}^w\right]^t = [A][C]k(q\det P')$
$[j(X_w)G_r + j(X_r)G_w]\left[d\widehat{\mathcal{K}}^w\right]^t = -[B][C]k(q\det P')$
$(j(X_r)[G_r] - j(X_w)[G_w])j\left(d\widehat{\mathcal{K}}^r\right) + (j(X_r)[G_{0r}] - j(X_w)[G_{0w}])\left[d\widehat{\mathcal{K}}^w\right] = -([A][C]k)v^t(q\det P')$
$(j(X_w)[G_r] + j(X_r)[G_w])j\left(d\widehat{\mathcal{K}}^r\right) + (j(X_w)[G_{0r}] + j(X_r)[G_{0w}])\left[d\widehat{\mathcal{K}}^w\right] = [B][C]kv^t(q\det P')$
By product with $[X_r]^t, [X_w]^t$
$[X_r]^t j(X_w)\left([G_w]j\left(d\widehat{\mathcal{K}}^r\right) + [G_{0w}]\left[d\widehat{\mathcal{K}}^w\right]\right) = [X_r]^t[A][C]kv^t(q\det P')$
$[X_r]^t j(X_w)\left([G_r]j\left(d\widehat{\mathcal{K}}^r\right) + [G_{0w}]\left[d\widehat{\mathcal{K}}^w\right]\right) = [X_r]^t[B][C]kv^t(q\det P')$
$[X_r]^t j(X_w)\left([G_r]j\left(d\widehat{\mathcal{K}}^r\right) + [G_{0r}]\left[d\widehat{\mathcal{K}}^w\right]\right) = [X_w]^t[A][C]kv^t(q\det P')$
$[X_r]^t j(X_w)\left([G_w]j\left(d\widehat{\mathcal{K}}^r\right) + [G_{0w}]\left[d\widehat{\mathcal{K}}^w\right]\right) = -[X_w]^t[B][C]kv^t(q\det P')$
The equations hold in the vacuum, which leads to $[X_r] = \lambda[X_w]$ and the equations of the previous subsection lead then to $\lambda = Ct \in \mathbb{R}$.
And then the equations sum up to :
$(1-\lambda^2)j(X_w)[G_w]\left[d\widehat{\mathcal{K}}^w\right]^t = 0$

$$\left(1+\lambda^{2}\right)j\left(X_{w}\right)\left[G_{r}\right]\left[d\widehat{\mathcal{K}}^{w}\right]^{t}=0$$
$$\left(1-\lambda^{2}\right)j\left(X_{w}\right)\left\{\left[G_{w}\right]j\left(d\widehat{\mathcal{K}}^{r}\right)+\left[G_{0w}\right]\left[d\widehat{\mathcal{K}}^{w}\right]\right\}=0$$
$$\left(\lambda^{2}+1\right)j\left(X_{w}\right)\left\{\left[G_{r}\right]j\left(d\widehat{\mathcal{K}}^{r}\right)+\left[G_{0r}\right]\left[d\widehat{\mathcal{K}}^{w}\right]\right\}=0$$
$$\Rightarrow$$
$$\left[G_{r}\right]=\mu\left[G_{w}\right],\mu\in C_{\infty}\left(M;\mathbb{R}\right)$$
$$\upsilon\left(d\left(\mu G_{w\alpha\beta}\right),dG_{w\alpha\beta}\right)=\upsilon\left(\lambda X_{w}^{a}d\left(\widehat{\mathcal{K}}_{\alpha\beta}\right),X_{w}^{a}d\left(\widehat{\mathcal{K}}_{\alpha\beta}\right)\right)$$
$$j\left(X_{w}\right)\left[G_{w}\right]\left[d\widehat{\mathcal{K}}^{w}\right]^{t}=0$$
$$j\left(X_{w}\right)\left[G_{w}\right]j\left(d\widehat{\mathcal{K}}^{r}\right)=0$$

So, with the codifferential and current equations one can expect to compute the strength of the field and the potential. But we have still the tetrad as parameter. The tetrad equation should provide the necessary adjustment

7.4 ENERGY AND MOMENTUM OF THE SYSTEM

7.4.1 Energy of the system

Conservation of energy

The lagrangian is, up to constants C_A, C_G, C_I depending of the unities, the balance of energy between the components of the system.

The conservation of the energy of the system requires : $div\left(\varepsilon_0 L\right) = 0$ (see Lagrangian), that is :

$\sum_{\alpha=0}^{3} \frac{d}{d\xi^\alpha}\left(\varepsilon_0^\alpha L \det P'\right) = \frac{d}{d\xi^0}\left(L \det P'\right)$

$= \frac{d}{dt}\left(L \det P'\right) = \det P' \frac{dL}{dt} + L \frac{d}{dt}\left(\det P'\right)$

$= \det P' \frac{dL}{dt} - L \sum_{i\alpha} P_\alpha'^i \frac{dP_\alpha^i}{dt}\left(\det P'\right)$

that is :

$\frac{dL}{dt} = L \sum_{i\alpha} P_\alpha'^i \frac{dP_\alpha^i}{dt}$

But the tetrad equation reads

$\frac{dL}{dP_i^\alpha} - LP_\alpha'^i = 0$

and on shell the condition sums up to the identity :

$\frac{dL}{dt} = \sum_{i\alpha} LP_\alpha'^i \frac{dP_\alpha^i}{dt} = \sum_{i\alpha} \frac{dL}{dP_i^\alpha} \frac{dP_\alpha^i}{dt}$

The tetrad equation implies the conservation of energy, and the result holds for any lagrangian in the tetrad formalism.

So this equation has a special significance :

- it expresses, in the more general setting, a general principle which goes beyond the Principle of Least Action,
- it encompasses all the system, and its physical objects (particles and fields),
- as seen in the second model, it can be derived by the use of functional derivatives, and does not require all the smoothness conditions imposed by the Lagrange equations,

Energy of the fields

The energy density is for the gravitational field :

$\langle \mathcal{F}_G, \mathcal{F}_G \rangle$

$= \sum_{a=1}^{3} \sum_{\{\alpha\beta\}} \mathcal{F}_r^{a\alpha\beta} \mathcal{F}_{r\alpha\beta}^a - \mathcal{F}_w^{a\alpha\beta} \mathcal{F}_{w\alpha\beta}^a$

$= \sum_{a=1}^{3} G_2\left(\mathcal{F}_r^a, \mathcal{F}_r^a\right) - G_2\left(\mathcal{F}_w^a \mathcal{F}_w^a\right)$

$= \sum_{a=1}^{3} G_2\left(\mathcal{F}_r^a, dG_r^a\right) + 2G_2\left(\mathcal{F}_r^a, \sum_{\{\alpha,\beta\}=0}^{3}\left(j\left(G_{r\alpha}\right)G_{r\beta} - j\left(G_{w\alpha}\right)G_{w\beta}\right)^a d\xi^\alpha \wedge d\xi^\beta\right)$

$-G_2\left(\mathcal{F}_w^a dG_w^a\right) - 2G_2\left(\mathcal{F}_w^a, \sum_{\{\alpha,\beta\}=0}^{3}\left(j\left(G_{w\alpha}\right)G_{r\beta} + j\left(G_{r\alpha}\right)G_{w\beta}\right)^a d\xi^\alpha \wedge d\xi^\beta\right)$

The codifferential is the adjoint of the exterior differential (Maths.2498) :

$G_2\left(\mathcal{F}_r^a, dG_r^a\right) = G_1\left(\delta\mathcal{F}_r^a, G_r^a\right)$

$G_2\left(\mathcal{F}_w^a, dG_w^a\right) = G_1\left(\delta\mathcal{F}_w^a, G_w^a\right)$

and on shell : $\delta\mathcal{F}_G = 0$

$\langle \mathcal{F}_G, \mathcal{F}_G \rangle$

$= 2G_2\left(\mathcal{F}_r^a, \sum_{\{\alpha,\beta\}=0}^{3}\left(j\left(G_{r\alpha}\right)G_{r\beta} - j\left(G_{w\alpha}\right)G_{w\beta}\right)^a d\xi^\alpha \wedge d\xi^\beta\right)$

$-2G_2\left(\mathcal{F}_w^a, \sum_{\{\alpha,\beta\}=0}^{3}\left(j\left(G_{w\alpha}\right)G_{r\beta} + j\left(G_{r\alpha}\right)G_{w\beta}\right)^a d\xi^\alpha \wedge d\xi^\beta\right)$

$= 2\left\langle \mathcal{F}_G, \sum_{a=1}^{6} \sum_{\{\alpha,\beta\}=0}^{3} \left[G_\alpha, G_\beta\right]^a d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\kappa}_a \right\rangle_G$

We have proven earlier that :

$\langle\langle X,[Y,Z]\rangle\rangle_G=\frac{1}{2}\sum_{a=1}^{3}\sum_{\alpha,\beta=0}^{3}X_r^{a\alpha\beta}\left[Y_\alpha,Z_\beta\right]_r^a-X_w^{a\alpha\beta}\left[Y_\alpha,Z_\beta\right]_w^a=2\sum_{\alpha,\beta=0}^{3}\left\langle\left[X^{\alpha\beta},Y_\alpha\right],Z_\beta\right\rangle_{Cl}$
$\left\langle\mathcal{F}_G,\sum_{a=1}^{6}\sum_{\{\alpha,\beta\}=0}^{3}\left[G_\alpha,G_\beta\right]^a d\xi^\alpha\wedge d\xi^\beta\otimes\overrightarrow{\kappa}_a\right\rangle$
$=2\sum_{\alpha,\beta=0}^{3}\left\langle\left[\mathcal{F}_G^{\alpha\beta},G_\alpha\right],G_\beta\right\rangle_{Cl}$
$=2\sum_{\alpha,\beta=0}^{3}\left\langle\left[\mathcal{F}_G^{\beta\alpha},G_\beta\right],G_\alpha\right\rangle_{Cl}$
$=-2\sum_{\beta=0}^{3}\left\langle\sum_{\alpha=0}^{3}\left[\mathcal{F}_G^{\alpha\beta},G_\beta\right],G_\alpha\right\rangle_{Cl}=-2\sum_{\beta=0}^{3}\left\langle\phi_{G\beta},G_\beta\right\rangle_{Cl}$

$$\langle\mathcal{F}_G,\mathcal{F}_G\rangle=-4\sum_{\beta=0}^{3}\left\langle\phi_{G\beta},G_\beta\right\rangle_{Cl}\tag{7.34}$$

And we have similarly for the fields other than EM :
$\langle\mathcal{F}_A,\mathcal{F}_A\rangle$
$=\sum_{a=1}^{m}\sum_{\{\alpha\beta\}}\mathcal{F}_A^{a\alpha\beta}\mathcal{F}_{A\alpha\beta}^a$
$=\sum_{a=1}^{m}G_2\left(\mathcal{F}_A^a,\mathcal{F}_A^a\right)$
$=\sum_{a=1}^{m}G_2\left(\mathcal{F}_A^a,d\grave{A}^a\right)+2G_2\left(\mathcal{F}_A^a,\sum_{\{\alpha,\beta\}=0}^{3}\left[\grave{A}_\alpha,\grave{A}_\beta\right]^a d\xi^\alpha\wedge d\xi^\beta\right)$
$=\sum_{a=1}^{m}G_1\left(\delta\mathcal{F}_A^a,\grave{A}^a\right)+2G_2\left(\mathcal{F}_A^a,\sum_{\{\alpha,\beta\}=0}^{3}\left[\grave{A}_\alpha,\grave{A}_\beta\right]^a d\xi^\alpha\wedge d\xi^\beta\right)$
$=2\sum_{a=1}^{m}G_2\left(\mathcal{F}_A^a,\sum_{\{\alpha,\beta\}=0}^{3}\left[\grave{A}_\alpha,\grave{A}_\beta\right]^a d\xi^\alpha\wedge d\xi^\beta\right)$
$=2\sum_{a=1}^{m}\sum_{\{\alpha\beta\}}\mathcal{F}_A^{a\alpha\beta}\left[\grave{A}_\alpha,\grave{A}_\beta\right]^a$
$=2\sum_{\{\alpha\beta\}}\left\langle\mathcal{F}_A^{\alpha\beta},\left[\grave{A}_\alpha,\grave{A}_\beta\right]\right\rangle_{T_1U}$
$=2\sum_{\{\alpha\beta\}}\left\langle\left[\mathcal{F}_A^{\alpha\beta},\grave{A}_\alpha\right],\grave{A}_\beta\right\rangle_{T_1U}$
$=\sum_{\alpha\beta}\left\langle\left[\mathcal{F}_A^{\alpha\beta},\grave{A}_\alpha\right],\grave{A}_\beta\right\rangle_{T_1U}$
$=-\sum_{\alpha\beta}\left\langle\left[\mathcal{F}_A^{\beta\alpha},\grave{A}_\alpha\right],\grave{A}_\beta\right\rangle_{T_1U}$

$$\langle\mathcal{F}_A,\mathcal{F}_A\rangle=-\sum_{\beta=0}^{3}\left\langle\phi_{A\beta},\grave{A}_\beta\right\rangle_{T_1U}\tag{7.35}$$

For the EM field the currents are null.
$\langle\mathcal{F}_{EM},\mathcal{F}_{EM}\rangle=\sum_{\{\alpha\beta\}}\mathcal{F}_{EM}^{\alpha\beta}\mathcal{F}_{EM\alpha\beta}=G_2\left(\mathcal{F}_{EM},\mathcal{F}_{EM}\right)=G_2\left(\mathcal{F}_{EM},d\grave{A}\right)=G_1\left(\delta\mathcal{F}_{EM},\grave{A}\right)$

$$\langle\mathcal{F}_{EM},\mathcal{F}_{EM}\rangle\varpi_4=G_1\left(\delta\mathcal{F}_{EM},\grave{A}\right)\tag{7.36}$$

So that on shell :
$L_{Fields}=\left(\sum_{\alpha\beta}C_G\left(\sum_{a=1}^{3}\mathcal{F}_{G\alpha\beta}^a\mathcal{F}_G^{a\alpha\beta}-\sum_{a=4}^{6}\mathcal{F}_{G\alpha\beta}^a\mathcal{F}_G^{a\alpha\beta}\right)+C_A\sum_{a=1}^{m}\mathcal{F}_{A\alpha\beta}^a\mathcal{F}_A^{a\alpha\beta}+C_{EM}\mathcal{F}_{EM}^{\alpha\beta}\mathcal{F}_{EM\alpha\beta}\right)$
$=2\left\{C_G\langle\mathcal{F}_G,\mathcal{F}_G\rangle_G+C_A\langle\mathcal{F}_A,\mathcal{F}_A\rangle_A+C_{EM}\langle\mathcal{F}_{EM},\mathcal{F}_{EM}\rangle\right\}$
$=2\sum_{\beta=0}^{3}\left(-4C_G\left\langle\phi_{G\beta},G_\beta\right\rangle_{Cl}-C_A\left\langle\phi_{A\beta},\grave{A}_\beta\right\rangle_{T_1U}+C_{EM}G_1\left(\delta\mathcal{F}_{EM},\grave{A}\right)\right)$

Energy of the system on shell

We have seen previously that the energy of the particles can be expressed as :

$L_{Particles} = C_I \mu \operatorname{Im} \langle \psi, \nabla_V \psi \rangle$

$= C_I \mu k^t \left([C(r)]^t \left([D(r)] \frac{dr}{dt} + \frac{1}{2} j(w) \frac{dw}{dt} \right) \right)$

$+8 \sum_{\beta=0}^{3} \left(4C_G \left\langle J_G^\beta, G_\beta \right\rangle_{Cl} + C_A \left\langle J_A^\beta, \grave{A}_\beta \right\rangle_{T_1 U} + C_{EM} J_{EM}^\beta \grave{A}_\beta \right)$

And on shell :

$C_I \mu \operatorname{Im} \langle \psi, \nabla_V \psi \rangle = 0$

$C_I \mu k^t \left([C(r)]^t \left([D(r)] \frac{dr}{dt} + \frac{1}{2} j(w) \frac{dw}{dt} \right) \right)$

$= -8 \sum_{\beta=0}^{3} \left(4C_G \left\langle J_G^\beta, G_\beta \right\rangle_{Cl} + C_A \left\langle J_A^\beta, \grave{A}_\beta \right\rangle_{T_1 U} + C_{EM} J_{EM}^\beta \grave{A}_\beta \right)$

Then, on shell :

$L_{Shell} = L_{Fields} = 2 \left(\sum_{\beta=0}^{3} -4C_G \langle \phi_{G\beta}, G_\beta \rangle_{Cl} - C_A \left\langle \phi_{A\beta}, \grave{A}_\beta \right\rangle_{T_1 U} + C_{EM} G_1 \left(\delta \mathcal{F}_{EM}, \grave{A} \right) \right)$

and with :

$J_G = \phi_G; J_A = \phi_A; \frac{1}{2} \delta \mathcal{F}_{EM} = -J_{EM}$

$L_{Fields} = 2 \left(\sum_{\beta=0}^{3} -4C_G \langle \phi_{G\beta}, G_\beta \rangle_{Cl} - C_A \left\langle \phi_{A\beta}, \grave{A}_\beta \right\rangle_{T_1 U} + C_{EM} G_1 \left(\delta \mathcal{F}_{EM}, \grave{A} \right) \right)$

$= -2 \left(\sum_{\beta=0}^{3} 4C_G \langle \phi_{G\beta}, G_\beta \rangle_{Cl} + C_A \left\langle J_A, \grave{A}_\beta \right\rangle_{T_1 U} + C_{EM} G_1 \left(J_{EM}, \grave{A} \right) \right)$

$= \frac{1}{4} C_I \mu k^t \left([C(r)]^t \left([D(r)] \frac{dr}{dt} + \frac{1}{2} j(w) \frac{dw}{dt} \right) \right)$

or equivalently $\frac{1}{4}$ of the kinetic energy of the particles.

This holds for a continuous distribution of particles or for individual particles.

7.4.2 Tetrad equation

The equation reads :

$\forall \alpha, \beta = 0...3 : C_I \frac{1}{i} V^\beta \langle \psi_p, \nabla_\alpha \psi_p \rangle + 4 \sum_\gamma \{ C_G \sum_{a=1}^{3} \mathcal{F}_{r\alpha\gamma}^a \mathcal{F}_r^{a\beta\gamma} - \mathcal{F}_{w\alpha\gamma}^a \mathcal{F}_w^{a\beta\gamma} + C_A \sum_{a=1}^{m} \mathcal{F}_{A\alpha\gamma}^a \mathcal{F}_A^{a\beta\gamma} + C_{EM} \mathcal{F}_{EM\alpha\gamma} \mathcal{F}_{EM}^{\beta\gamma} \}$

$= 2\delta_\alpha^\beta \left(\sum_{\lambda\mu} \left(C_G \langle \mathcal{F}_G, \mathcal{F}_G \rangle + C_A \langle \mathcal{F}_A, \mathcal{F}_A \rangle + C_{EM} \langle \mathcal{F}_{EM}, \mathcal{F}_{EM} \rangle \right) \right)$

It is the only equation which shows a relation between the gravitational field and the other fields. It is roughly the equivalent of the Einstein equation, and it has the same drawbacks : in the vacuum it leads to a solution where $\langle \mathcal{F}, \mathcal{F} \rangle = 0$: the field does not propagate.

The terms related to the fields can be expressed in a more operational way.

Any scalar 2 form $\mathcal{F}$ can be written in matrix form. The 4×4 matrix $[\mathcal{F}_{\alpha\beta}]_{\beta=0..3}^{\alpha=0..3}$ can be written :

$$[\mathcal{F}]_{\beta=0..3}^{\alpha=0..3} = \begin{bmatrix} 0 & \mathcal{F}_{01} & \mathcal{F}_{02} & \mathcal{F}_{03} \\ -\mathcal{F}_{01} & 0 & -\mathcal{F}_{21} & \mathcal{F}_{13} \\ -\mathcal{F}_{02} & \mathcal{F}_{21} & 0 & -\mathcal{F}_{32} \\ -\mathcal{F}_{03} & -\mathcal{F}_{13} & \mathcal{F}_{32} & 0 \end{bmatrix} = \begin{bmatrix} 0 & [\mathcal{F}^w] \\ -[\mathcal{F}^w]^t & [j([\mathcal{F}^r])] \end{bmatrix}$$

A quantity such as : $\sum_{\lambda=0}^{3} \mathcal{F}_{\alpha\lambda} \mathcal{F}^{\beta\lambda}$ is then the coefficient (α, β) of the matrices product

$\sum_{\lambda=0}^{3} \mathcal{F}_{\alpha\lambda} \mathcal{F}^{\beta\lambda} = -\sum_{\lambda=0}^{3} \mathcal{F}_{\alpha\lambda} \mathcal{F}^{\lambda\beta} = \sum_{\lambda=0}^{3} [\mathcal{F}]_\lambda^\alpha [\mathcal{F}^*]_\beta^\lambda = -([\mathcal{F}][\mathcal{F}^*])_\beta^\alpha$

with the matrix : $[\mathcal{F}^*]_{\beta=0..3}^{\alpha=0..3} = \left[\mathcal{F}^{\alpha\beta} \right]$.

$\mathcal{F}^{\alpha\beta}$ are the coefficients of the Hodge dual, but they do not relate to the same indices :

$*\mathcal{F}^r = -\left(\mathcal{F}^{01} d\xi^3 \wedge d\xi^2 + \mathcal{F}^{02} d\xi^1 \wedge d\xi^3 + \mathcal{F}^{03} d\xi^2 \wedge d\xi^1\right) \det P'$
$*\mathcal{F}^w = -\left(\mathcal{F}^{32} d\xi^0 \wedge d\xi^1 + \mathcal{F}^{13} d\xi^0 \wedge d\xi^2 + \mathcal{F}^{21} d\xi^0 \wedge d\xi^3\right) \det P'$
so that :
$[\mathcal{F}^*] = \frac{1}{\det P'} \begin{bmatrix} 0 & -[*\mathcal{F}^r] \\ [*\mathcal{F}^r]^t & -[j([*\mathcal{F}^w])] \end{bmatrix}$
and in the standard chart :
$[*\mathcal{F}^r] = [\mathcal{F}^w]\left[g_3^{-1}\right] \det Q'$
$[*\mathcal{F}^w] = -[\mathcal{F}^r][g_3] \det Q$
$[\mathcal{F}^*] = \frac{1}{\det Q'} \begin{bmatrix} 0 & -[\mathcal{F}^w][g_3]^{-1} \det Q' \\ [g_3]^{-1}[\mathcal{F}^w]^t \det Q' & -j([\mathcal{F}^r][g_3]) \det Q \end{bmatrix}$
with the identity $j(Mx) = [M^t]^{-1} j(x) [M]^{-1} \det M$
$j([\mathcal{F}^r][g_3]) = [g_3]^{-1} j([\mathcal{F}^r]^t) [g_3]^{-1} (\det Q')^2$
$[\mathcal{F}^*] = \begin{bmatrix} 0 & -[\mathcal{F}^w][g_3]^{-1} \\ [g_3]^{-1}[\mathcal{F}^w]^t & -[g_3]^{-1} j(\mathcal{F}^r)[g_3]^{-1} \end{bmatrix}$
and

$$\left[\sum_{\lambda=0}^{3} \mathcal{F}_{\alpha\lambda}\mathcal{F}^{\beta\lambda}\right] = \begin{bmatrix} -[\mathcal{F}^w][g_3]^{-1}[\mathcal{F}^w]^t & [\mathcal{F}^w][g_3]^{-1} j(\mathcal{F}^r)[g_3]^{-1} \\ -j(\mathcal{F}^r)[g_3]^{-1}[\mathcal{F}^w]^t & \left(j(\mathcal{F}^r)[g_3]^{-1} j(\mathcal{F}^r) - [\mathcal{F}^w]^t[\mathcal{F}^w]\right)[g_3]^{-1} \end{bmatrix} \quad (7.37)$$

Then $4\left\langle \mathcal{F}_{G\alpha\gamma}, \mathcal{F}_G^{\beta\gamma}\right\rangle_{Cl} = \sum_{a=1}^3 \left([\mathcal{F}_r]^a ([\mathcal{F}_r^*]^a)^t\right)_\beta^\alpha - \left([\mathcal{F}_w]^a ([\mathcal{F}_w^*]^a)^t\right)_\beta^\alpha$
$\sum_{a=1}^3 \left([\mathcal{F}_r]^a ([\mathcal{F}_r^*]^a)^t\right)$
$= \sum_{a=1}^3 \begin{bmatrix} -[\mathcal{F}_r^w]^a [g_3]^{-1} ([\mathcal{F}_r^w]^a)^t & [\mathcal{F}_r^w]^a [g_3]^{-1} j([\mathcal{F}_r^r]^a)[g_3]^{-1} \\ -j([\mathcal{F}_r^r]^a)[g_3]^{-1}([\mathcal{F}_r^w]^a)^t & \left(j([\mathcal{F}_r^r]^a)[g_3]^{-1} j([\mathcal{F}_r^r]^a) - ([\mathcal{F}_r^w]^z)^t [\mathcal{F}_r^w]^a\right)[g_3]^{-1} \end{bmatrix}$
$\sum_{a=1}^3 [\mathcal{F}_r^w]^a [g_3]^{-1} ([\mathcal{F}_r^w]^a)^t = \sum_{a,p,q=1}^3 [\mathcal{F}_r^w]_p^a \left([g_3]^{-1}\right)_q^p [\mathcal{F}_r^w]_q^a = Tr[\mathcal{F}_r^w][g_3]^{-1}[\mathcal{F}_r^w]^t = Tr[\mathcal{F}_r^w]^t[\mathcal{F}_r^w][g_3]^{-1}$
$\alpha, \beta = 1, 2, 3$:
$\sum_{a=1}^3 \left([\mathcal{F}_r^w]^a [g_3]^{-1} j([\mathcal{F}_r^r]^a)[g_3]^{-1}\right)_\beta$
$= (\det Q)^2 \sum_{a=1}^3 \left([\mathcal{F}_r^w]^a j([\mathcal{F}_r^r]^a [g_3])\right)_\beta$
$= -(\det Q)^2 \sum_{a,b,c=1}^3 \epsilon(\beta, b, c)([\mathcal{F}_r^r]^a [g_3])_b ([\mathcal{F}_r^w]^a)_c$
$= (\det Q)^2 \sum_{b,c=1}^3 \epsilon(\beta, b, c)\left([\mathcal{F}_r^w]^t [\mathcal{F}_r^r][g_3]\right)_c^b$
$= (\det Q)^2 J\left([\mathcal{F}_r^w]^t [\mathcal{F}_r^r][g_3]\right)_\beta$
$\sum_{a=1}^3 [\mathcal{F}_r^w]^a [g_3]^{-1} j([\mathcal{F}_r^r]^a)[g_3]^{-1} = (\det Q)^2 J\left([\mathcal{F}_r^w]^t [\mathcal{F}_r^r][g_3]\right)$
with :
$M = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \to J(M) = \begin{bmatrix} M_{23} - M_{32} & M_{13} - M_{31} & M_{12} - M_{21} \end{bmatrix}$
$\sum_{a=1}^3 j([\mathcal{F}_r^r]^a)[g_3]^{-1}([\mathcal{F}_r^w]^a)^t$
$= \sum_{a=1}^3 [g_3][g_3]^{-1} j([\mathcal{F}_r^r]^a)[g_3]^{-1}([\mathcal{F}_r^w]^a)^t$
$= -\sum_{a=1}^3 [g_3]\left([\mathcal{F}_r^w]^a [g_3]^{-1} j([\mathcal{F}_r^r]^a)[g_3]^{-1}\right)^t$
$= -(\det Q)^2 [g_3]\left(J\left([\mathcal{F}_r^w]^t [\mathcal{F}_r^r][g_3]\right)\right)^t$

$\sum_{a=1}^{3}\left(\left([\mathcal{F}_r^w]^a\right)^t[\mathcal{F}_r^w]^a[g_3]^{-1}\right)^\alpha_\beta=\sum_{a,p=1}^{3}\left([\mathcal{F}_r^w]^a_\alpha[\mathcal{F}_r^w]^a_p\left([g_3]^{-1}\right)^p_\beta\right)$
$=\sum_{p=1}^{3}\left([\mathcal{F}_r^w]^t[\mathcal{F}_r^w]\right)^\alpha_p\left([g_3]^{-1}\right)^p_\beta=\left([\mathcal{F}_r^w]^t[\mathcal{F}_r^w][g_3]^{-1}\right)^\alpha_\beta$

$\sum_{a=1}^{3}\left(j\left([\mathcal{F}_r^r]^a\right)[g_3]^{-1}j(\mathcal{F}_r^r)[g_3]^{-1}\right)^\alpha_\beta$
$=\sum_{a=1}^{3}\left(j\left([\mathcal{F}_r^r]^a\right)j\left([\mathcal{F}_r^r]^a[g_3]\right)\right)^\alpha_\beta(\det Q)^2$
$=(\det Q)^2\sum_{a=1}^{3}\left(\left(\left([\mathcal{F}_r^r]^a[g_3]\right)^t\left([\mathcal{F}_r^r]^a\right)\right)-\left(\left([\mathcal{F}_r^r]^a[g_3]\right)\left([\mathcal{F}_r^r]^a\right)^t\right)I\right)^\alpha_\beta$
$=(\det Q)^2\sum_{a=1}^{3}\left(\left([\mathcal{F}_r^r]^a[g_3]\right)^t\left([\mathcal{F}_r^r]^a\right)\right)^\alpha_\beta-\delta^\alpha_\beta\left([\mathcal{F}_r^r]^a[g_3]\right)\left([\mathcal{F}_r^r]^a\right)^t$
$=(\det Q)^2\sum_{a,p,q=1}^{3}[g_3]^\alpha_p[\mathcal{F}_r^r]^a_p[\mathcal{F}_r^r]^a_\beta-\delta^\alpha_\beta[\mathcal{F}^r]^a_p[g_3]^p_q[\mathcal{F}_r^r]^a_q$
$=(\det Q)^2\sum_{a,p,q=1}^{3}\left([g_3][\mathcal{F}_r^r]^t[\mathcal{F}_r^r]\right)^\alpha_\beta-\delta^\alpha_\beta Tr\left([g_3][\mathcal{F}_r^r]^t[\mathcal{F}_e^r]\right)$
$\sum_{a=1}^{3}j\left([\mathcal{F}_r^r]^a\right)[g_3]^{-1}j(\mathcal{F}_r^r)[g_3]^{-1}=(\det Q)^2\left([g_3][\mathcal{F}_r^r]^t[\mathcal{F}_r^r]-Tr\left([g_3][\mathcal{F}_r^r]^t[\mathcal{F}_r^r]\right)\right)$
$\sum_{a=1}^{3}\left([\mathcal{F}_r]^a\left([\mathcal{F}_r^*]^a\right)^t\right)=$

$$\begin{bmatrix}-Tr[\mathcal{F}_r^w]^t[\mathcal{F}_r^w][g_3]^{-1} & (\det Q)^2J\left([\mathcal{F}_r^w]^t[\mathcal{F}_r^r][g_3]\right)\\ (\det Q)^2[g_3]\left(J\left([\mathcal{F}_r^w]^t[\mathcal{F}_r^r][g_3]\right)\right)^t & (\det Q)^2\left([g_3][\mathcal{F}_r^r]^t[\mathcal{F}_r^r]-Tr\left([g_3][\mathcal{F}_r^r]^t[\mathcal{F}_r^r]\right)\right)-[\mathcal{F}_r^w]^t[\mathcal{F}_r^w][g_3]^{-1}\end{bmatrix}$$

and similarly for the other quantities in the left hand side of the equation.

If the tensor $\mathcal{F}$ is decomposable each matricial quantity is the product of row and column matrices by a scalar :
$[\mathcal{F}_r^w]^a[g_3]^{-1}\left([\mathcal{F}_r^w]^a\right)^t=(X^a)^2\left[\widehat{\mathcal{F}}^w\right][g_3]^{-1}\left[\widehat{\mathcal{F}}_r^w\right]^t$
and
$\sum_{a=1}^{3}\mathcal{F}^a_{r\alpha\gamma}\mathcal{F}_r^{a\beta\gamma}-\mathcal{F}^a_{w\alpha\gamma}\mathcal{F}_w^{a\beta\gamma}$

$$=\left([X_r]^t[X_r]-[X_w][X_x]\right)\begin{bmatrix}-\left[\widehat{\mathcal{F}}^w\right][g_3]^{-1}\left[\widehat{\mathcal{F}}^w\right]^t & \left[\widehat{\mathcal{F}}^w\right][g_3]^{-1}j(\widehat{\mathcal{F}}^r)[g_3]^{-1}\\ -j\left(\widehat{\mathcal{F}}^r\right)[g_3]^{-1}\left[\widehat{\mathcal{F}}^w\right]^t & \left(j\left(\widehat{\mathcal{F}}^r\right)[g_3]^{-1}j(\mathcal{F}^r)-\left[\widehat{\mathcal{F}}^w\right]^t\left[\widehat{\mathcal{F}}^w\right]\right)[g_3]^{-1}\end{bmatrix}$$

The terms in the right hand side is just L_{Fields}
$\left[2\delta^\beta_\alpha\left(\sum_{\lambda\mu}\left(C_G\langle\mathcal{F}_G,\mathcal{F}_G\rangle+C_A\langle\mathcal{F}_A,\mathcal{F}_A\rangle+C_{EM}\langle\mathcal{F}_{EM},\mathcal{F}_{EM}\rangle\right)\right)\right]=L_{Fields}\times I_4$
with
$\langle\mathcal{F},\mathcal{F}\rangle_G=\frac{1}{\det P'}Tr\left(\left([*\mathcal{F}_r^w][\mathcal{F}_r^r]^t+[*\mathcal{F}_r^r][\mathcal{F}_r^w]^t\right)-\left([*\mathcal{F}_w^w][\mathcal{F}_w^r]^t+[*\mathcal{F}_w^r][\mathcal{F}_w^w]^t\right)\right)$
$\langle\mathcal{F},\mathcal{F}\rangle_A=\frac{1}{\det P'}Tr\left([*\mathcal{F}^w][\mathcal{F}^r]^t+[*\mathcal{F}^r][\mathcal{F}^w]^t\right)$
And if the tensor $\mathcal{F}$ is decomposable :
$\langle\mathcal{F},\mathcal{F}\rangle_G=\frac{1}{\det P'}\left([X_r]^t[X_r]-[X_w]^t[X_w]\right)\left(\left[*\widehat{\mathcal{F}}_G^w\right]\left[\widehat{\mathcal{F}}_G^r\right]^t+\left[*\widehat{\mathcal{F}}_G^r\right]\left[\widehat{\mathcal{F}}^w\right]^t\right)$
$\langle\mathcal{F}_A,\mathcal{F}_A\rangle_A=\frac{1}{\det P'}\left([X_A]^t[X_A]\right)\left(\left(\left[*\widehat{\mathcal{F}}_A^w\right]\left[\widehat{\mathcal{F}}_A^r\right]^t\right)+\left(\left[*\widehat{\mathcal{F}}_A^r\right]\left[\widehat{\mathcal{F}}_A^w\right]^t\right)\right)$

Tetrad equation for the gravitational field in the vacuum

To see what can be deduced from this equation we will focus on the gravitational field, in a system without any other field and no particles.

In the absence of particle, the equation sums up to :
$\forall \alpha, \beta = 0...3 : \sum_{\gamma=0}^{3} \sum_{a=1}^{3} \mathcal{F}_{r\alpha\gamma}^{a} \mathcal{F}_{r}^{a\beta\gamma} - \mathcal{F}_{w\alpha\gamma}^{a} \mathcal{F}_{w}^{a\beta\gamma}$
$= \frac{1}{2}\delta_{\alpha}^{\beta}\{C_G Tr\left((\det Q)^2\left([\mathcal{F}_r^r][g_3][\mathcal{F}_r^r]^t - [\mathcal{F}_w^r][g_3][\mathcal{F}_w^r]^t\right) + [\mathcal{F}_w^w]\left[g_3^{-1}\right][\mathcal{F}_w^w]^t - [\mathcal{F}_r^w]\left[g_3^{-1}\right][\mathcal{F}_r^w]^t\right)$
Using the matricial representation, we are left with 3 equations :
i) $Tr\left([\mathcal{F}_w^w]^t[\mathcal{F}_w^w] - Tr[\mathcal{F}_r^w]^t[\mathcal{F}_r^w]\right)[g_3]^{-1}$
$= \frac{1}{2}\{Tr\left((\det Q)^2\left([\mathcal{F}_r^r]^t[\mathcal{F}_r^r] - [\mathcal{F}_w^r]^t[\mathcal{F}_w^r]\right)[g_3] + \left([\mathcal{F}_w^w]^t[\mathcal{F}_w^w] - [\mathcal{F}_r^w]^t[\mathcal{F}_r^w]\right)\left[g_3^{-1}\right]\right)$
ii) $(\det Q)^2\left([g_3]\left([\mathcal{F}_r^r]^t[\mathcal{F}_r^r] - [\mathcal{F}_w^r]^t[\mathcal{F}_w^r]\right) + I_3 Tr[g_3]\left([\mathcal{F}_w^r]^t[\mathcal{F}_w^r] - [\mathcal{F}_r^r]^t[\mathcal{F}_r^r]\right)\right) + \left([\mathcal{F}_w^w]^t[\mathcal{F}_w^w] - [\mathcal{F}_r^w]^t[\mathcal{F}_r^w]\right)[$
$= \frac{1}{2}\{Tr\left((\det Q)^2\left([\mathcal{F}_r^r]^t[\mathcal{F}_r^r] - [\mathcal{F}_w^r]^t[\mathcal{F}_w^r]\right)[g_3] + \left([\mathcal{F}_w^w]^t[\mathcal{F}_w^w] - [\mathcal{F}_r^w]^t[\mathcal{F}_r^w]\right)\left[g_3^{-1}\right]\right)\}I_3$
iii) $J\left([\mathcal{F}_r^w]^t[\mathcal{F}_r^r][g_3]\right) - J\left([\mathcal{F}_w^w]^t[\mathcal{F}_w^r][g_3]\right) = 0$

The first equations gives :
$Tr\left(\left([\mathcal{F}_w^w]^t[\mathcal{F}_w^w] - [\mathcal{F}_r^w]^t[\mathcal{F}_r^w]\right)[g_3]^{-1}\right) = -(\det Q)^2 Tr\left([\mathcal{F}_w^r]^t[\mathcal{F}_w^r] - [\mathcal{F}_r^r]^t[\mathcal{F}_r^r]\right)[g_3]$
The second equation reads :
$-(\det Q)^2[g_3]\left([\mathcal{F}_w^r]^t[\mathcal{F}_w^r] - [\mathcal{F}_r^r]^t[\mathcal{F}_r^r]\right) + \left([\mathcal{F}_w^w]^t[\mathcal{F}_w^w] - [\mathcal{F}_r^w]^t[\mathcal{F}_r^w]\right)[g_3]^{-1}$
$= 2Tr\left(\left([\mathcal{F}_w^w]^t[\mathcal{F}_w^w] - [\mathcal{F}_r^w]^t[\mathcal{F}_r^w]\right)[g_3]^{-1}\right)I_3$
By taking the trace :
$-(\det Q)^2 Tr[g_3]\left([\mathcal{F}_w^r]^t[\mathcal{F}_w^r] - [\mathcal{F}_r^r]^t[\mathcal{F}_r^r]\right) + Tr\left([\mathcal{F}_w^w]^t[\mathcal{F}_w^w] - [\mathcal{F}_r^w]^t[\mathcal{F}_r^w]\right)[g_3]^{-1}$
$= 6Tr\left(\left([\mathcal{F}_w^w]^t[\mathcal{F}_w^w] - [\mathcal{F}_r^w]^t[\mathcal{F}_r^w]\right)[g_3]^{-1}\right)$
$Tr\left(\left([\mathcal{F}_w^w]^t[\mathcal{F}_w^w] - [\mathcal{F}_r^w]^t[\mathcal{F}_r^w]\right)[g_3]^{-1}\right) = 0$
$\Rightarrow$

$$Tr\left([\mathcal{F}_w^w]^t[\mathcal{F}_w^w] - [\mathcal{F}_r^w]^t[\mathcal{F}_r^w]\right)[g_3]^{-1} = Tr\left([\mathcal{F}_w^r]^t[\mathcal{F}_w^r] - [\mathcal{F}_r^r]^t[\mathcal{F}_r^r]\right)[g_3] = 0 \tag{7.38}$$

As a consequence $\langle \mathcal{F}_G, \mathcal{F}_G \rangle = 0$: the field is stationary.
And the second equation reads :

$$[\mathcal{F}_w^r]^t[\mathcal{F}_w^r] - [\mathcal{F}_r^r]^t[\mathcal{F}_r^r] = (\det Q')^2[g_3]^{-1}\left([\mathcal{F}_w^w]^t[\mathcal{F}_w^w] - [\mathcal{F}_r^w]^t[\mathcal{F}_r^w]\right)[g_3]^{-1} \tag{7.39}$$

The last equation reads :
$(\det Q)^2\left(J\left([\mathcal{F}_r^w]^t[\mathcal{F}_r^r][g_3]\right) - J\left([\mathcal{F}_w^w]^t[\mathcal{F}_w^r][g_3]\right)\right) = 0$
$J\left(\left([\mathcal{F}_r^w]^t[\mathcal{F}_r^r] - [\mathcal{F}_w^w]^t[\mathcal{F}_w^r]\right)[g_3]\right) = 0$
$\forall \beta = 1,2,3 : \sum_{b,c=1}^{3} \epsilon(\beta, b, c)\left(\left([\mathcal{F}_w^w]^t[\mathcal{F}_w^r] - [\mathcal{F}_r^w]^t[\mathcal{F}_r^r]\right)[g_3]\right)_c^b = 0$
which is equivalent to :
$\left([\mathcal{F}_w^w]^t[\mathcal{F}_w^r] - [\mathcal{F}_r^w]^t[\mathcal{F}_r^r]\right)[g_3] = \left(\left([\mathcal{F}_w^w]^t[\mathcal{F}_w^r] - [\mathcal{F}_r^w]^t[\mathcal{F}_r^r]\right)[g_3]\right)^t$

$$\left([\mathcal{F}_w^w]^t[\mathcal{F}_w^r] - [\mathcal{F}_r^w]^t[\mathcal{F}_r^r]\right) = [g_3]\left([\mathcal{F}_w^r]^t[\mathcal{F}_w^w] - [\mathcal{F}_r^r]^t[\mathcal{F}_r^w]\right)[g_3]^{-1} \tag{7.40}$$

The codifferential equation hold in all cases. Using the relations above the equations can be written with $d\mathcal{K}$ and we get :

$[d\mathcal{K}_w^w]^t [d\mathcal{K}_w^w] - [d\mathcal{K}_r^w]^t [d\mathcal{K}_r^w] = [d\mathcal{K}_w^r]^t [d\mathcal{K}_w^r] - [d\mathcal{K}_r^r]^t [d\mathcal{K}_r^r]$

and similar expressions which can be expressed as a scalar product in $T_1 Spin(3,1)$:

$p, q = 1, 2, 3$:

$\left\langle [d\mathcal{K}^w]_p , [d\mathcal{K}^w]_q \right\rangle_{Cl} = \left\langle [d\mathcal{K}^r]_p , [d\mathcal{K}^r]_q \right\rangle_{Cl}$

$\left\langle [d\mathcal{K}^w]_p , [d\mathcal{K}^r]_q \right\rangle_{Cl} = \left\langle [d\mathcal{K}^w]_q , [d\mathcal{K}^r]_p \right\rangle_{Cl}$

$\sum_{p=1}^3 \left\langle [d\mathcal{K}^w]_p , [d\mathcal{K}^w]_p \right\rangle_{Cl} = 0$

$\sum_{p=1}^3 \left\langle [d\mathcal{K}^r]_p , [d\mathcal{K}^r]_p \right\rangle_{Cl} = 0$

So the tetrad equation sums up to $\langle \mathcal{F}_G, \mathcal{F}_G \rangle = 0$ and a set of equations in the scalar products $[d\mathcal{K}^w]_p , [d\mathcal{K}^r]_q$.

The equations are always met if

$\mathcal{F}_G = \widehat{\mathcal{F}} \otimes X$

$\left[\widehat{\mathcal{F}}^w\right] = \left[d\widehat{\mathcal{K}}^r\right] [Q'] ; \left[\widehat{\mathcal{F}}^r\right] = -(\det Q') \left[d\widehat{\mathcal{K}}^w\right] [Q]^t$ with $\widehat{\mathcal{K}} \in \Lambda_1 (M; \mathbb{R})$

Then $\langle \mathcal{F}_G, \mathcal{F}_G \rangle = 0$ and $d\widehat{\mathcal{K}}$ is undefined in the tetrad equation, and as a consequence the tetrad also is not defined.

As expected such a model cannot represent the propagation. Which leads to the introduction of the Energy-Momentum tensor.

7.4.3 Energy-momentum tensor

The Energy-Momentum tensor is a crucial part of the Physics of RG. In the usual presentation of GR, the Einstein equation is proven using the Principle of Least action and the action :

$S = \int_\Omega \left(\mathbf{R} - \frac{8\pi G}{\sqrt{c}} T\right) \varpi_4$

where G is the gravitational constant and T the Energy-Momentum tensor.

The metric is the only variable, then the variational derivative of S gives (with $\det P' = \sqrt{-\det g}$) :

$\frac{\delta S}{\delta g} = \int_\Omega \left(Ric_{\alpha\beta} - \frac{1}{2}\mathbf{R} g_{\alpha\beta} - \left(\frac{\partial T}{\partial g^{\alpha\beta}} + \frac{1}{2} g_{\alpha\beta} T \right)\right) \varpi_4$

and the Einstein equation reads : $Ric_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta} = \frac{8\pi G}{\sqrt{c}} T_{\alpha\beta}$ with $T_{\alpha\beta} = \frac{\partial T}{\partial g^{\alpha\beta}} - \frac{1}{2} g_{\alpha\beta} T$

The energy-momentum tensor comes from classical mechanics. In GR there is no general formula to specify the energy-momentum tensor T, only phenomenological laws. The most usual are based on the behavior of dust clouds, including sometimes thermodynamic components.

We have not proceeded this way, but started from a general lagrangian. However it is useful to remind the genuine meaning of the Energy-Momentum tensor, and the previous models provide a good framework for this purpose.

Definition

The concept of equilibrium is at the core of the Principle of Least Action. So, for any tentative change of the values of the variables, beyond the point of equilibrium, the system reacts by showing resistance against the change : this is the inertia of the system. It is better understood with the functional derivatives.

Let us consider an action with integrals of the kind :

$\ell\left(z^i, z^i_\alpha\right) = \int_\Omega \mathcal{L}\left(z^i, z^i_\alpha, P^\alpha_i\right) \varpi_4$

$= \int L\left(z^i, z^i_\alpha, P^\alpha_i\right) (\det P') \, d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

The variational derivative is, when δZ comes from a section (see the proof of Maths.2601), that is when : $\delta z^i_\alpha = \partial_\alpha \delta z^i$:

$\frac{\delta \ell}{\delta z^i}\left(\delta z^i\right) = \int \left(\frac{\partial \mathcal{L}}{\partial z^i} - \sum_\alpha \frac{d}{d\xi^\alpha}\left(\frac{\partial \mathcal{L}}{\partial z^i_\alpha}\right)\right)\left(\delta z^i\right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$= \int \left(\frac{\partial L}{\partial z^i} \det P' - \sum_\alpha \frac{d}{d\xi^\alpha}\left(\frac{\partial L}{\partial z^i_\alpha} \det P'\right)\right)\left(\delta z^i\right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$\int \frac{\partial L}{\partial z^i}\delta z^i \varpi_4 = \frac{\delta \ell}{\delta z^i}\left(\delta z^i\right) + \int \left(\sum_\alpha \frac{d}{d\xi^\alpha}\left(\frac{\partial L}{\partial z^i_\alpha} \det P'\right)\right)\delta z^i d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$\frac{\delta \ell}{\delta P^\alpha_i}\left(\delta P^\alpha_i\right) = \int \left(\frac{\partial L}{\partial P^\alpha_i} \det P' - L P'^i_\alpha \det P'\right)\left(\delta P^\alpha_i\right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$= \int \left(\frac{\partial L}{\partial P^\alpha_i} - L P'^i_\alpha\right)\left(\delta P^\alpha_i\right) \varpi_4$

For a change $\delta Z = \left(\delta z^i, \delta P^\alpha_i\right)$ the change of the action is :

$\delta \ell =$

$\sum_i \frac{\delta \ell}{\delta z^i}\delta z^i + \sum_{i\alpha} \frac{\delta \ell}{\delta P^\alpha_i}\delta P^\alpha_i + \int \left(\sum_{i\alpha} \frac{d}{d\xi^\alpha}\left(\frac{\partial L}{\partial z^i_\alpha} \det P'\right)\right)\delta z^i d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

By definition, on shell the first part is null, thus on the neighborhood of the equilibrium :

$\delta \ell = \int \left(\sum_\alpha \frac{d}{d\xi^\alpha}\left(\frac{\partial L}{\partial z^i_\alpha} \det P'\right)\right)\delta z^i d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$\delta Y = \sum_i \delta z^i \sum_\alpha \frac{\partial L}{\partial z^i_\alpha}\partial \xi_\alpha$ is a vector field (see Lagrangian) and

$\sum_\alpha \frac{d}{d\xi^\alpha}\left(\sum_i \frac{\partial L}{\partial z^i_\alpha}\delta z^i \left(\det P'\right)\right) = \left(\det P'\right) div\left(\delta Y\right)$

$$\delta \ell = \int_\Omega div\left(\delta Y\right)\varpi_4 \tag{7.41}$$

The variables z^i are sections of a fiber bundle. If we consider a translation : $\delta v = \sum_{\beta=0}^3 v^\beta \delta \xi_\beta$:

$\delta z^i = \sum_\beta \frac{\partial z^i}{\partial \xi^\beta}\delta v^\beta = \sum_\beta z^\beta_i \delta v^\beta$

and :

$\delta Y\left(\delta v\right) = \sum_i \left(\sum_\beta z^\beta_i \delta v^\beta\right)\sum_\alpha \frac{\partial L}{\partial z^i_\alpha}\partial \xi_\alpha = \sum_{i\alpha\beta} \frac{\partial L}{\partial z^i_\alpha} z^\beta_i \delta v^\beta \partial \xi_\alpha$

The quantity :

$$T = \sum_{i\alpha\beta} \frac{\partial L}{\partial z^i_\alpha} z^\beta_i \partial \xi_\alpha \otimes d\xi^\beta \tag{7.42}$$

is a tensor, and $\delta Y = T\left(\delta v\right)$

$$\delta \ell = \int_\Omega div\left(T\left(\delta v\right)\right)\varpi_4 \tag{7.43}$$

T is called the **energy-momentum tensor.** Its name and definition, closely related to the lagrangian, come from fluid mechanics. The lagrangian has the meaning of a density of energy for the whole system, $div\left(T\left(\delta v\right)\right)$ has the meaning of a variation of this energy due to the action of 4 dimensional forces T acting along the direction δv. T can be interpreted as the resistance of the system against a change in the direction $\partial \xi_\alpha$, or equivalently as the inertia of the system, and it is opposite to the 4 dimensional force necessary to change the equilibrium.

The computation above is quite general, holds for any lagrangian, and in the neighborhood of an equilibrium, and not just when an equilibrium is met. And with the use of functional derivatives we are not limited to smooth variables (δv does not need to be resulting from an infinitesimal change). So the Energy-Momentum Tensor seems to be an interesting tool to study discontinuous processes.

Usually $\delta \ell \neq 0$ for a change δZ of the variables. However if the divergence of the vector $\delta Y = T\left(\delta v\right)$ is null then $\delta \ell = 0$. If such vectors δv exist they show privileged directions over which the system can be deformed without energy spent : they provide equivalent states of equilibrium.

The energy-momentum tensor in our framework

With the more general lagrangian $\mathcal{L} = L(\psi, \nabla_\alpha \psi_p, P_i^\alpha, \mathcal{F}_{G\alpha\beta}, \mathcal{F}_{A\alpha\beta}, V^\alpha) \det P'$ the tensor T reads :

$T = \sum_{\alpha\beta} \{ \sum_{ij} \frac{\partial L}{\partial \partial_\alpha \psi^{ij}} \partial_\beta \psi^{ij} + \sum_{a\gamma} \frac{\partial L}{\partial \partial_\alpha \grave{A}^a_\gamma} \partial_\beta \grave{A}^a_\gamma + \sum_{a\gamma} \frac{\partial L}{\partial \partial_\alpha G^a_\gamma} \partial_\beta G^a_\gamma \} \partial \xi_\alpha \otimes d\xi^\beta$

Notice that P_i^α, V^αdo not appear. With :

$\frac{dL}{d\partial \partial_\alpha \psi^{ij}} = \Pi^\alpha_{\nabla ij}$

$\frac{dL}{d\partial_\beta G^a_\alpha} = -2\Pi^{\alpha\beta}_{Ga}$

$\frac{dL}{d\partial_\beta \grave{A}^a_\alpha} = -2\Pi^{\alpha\beta}_{Aa}$

$$T = \sum_{\alpha\beta} \left(\sum_{ij} \Pi^\alpha_{\nabla ij} \partial_\beta \psi^{ij} - 2 \sum_{a\gamma} \Pi^{\alpha\gamma}_{Aa} \partial_\beta \grave{A}^a_\gamma + \Pi^{\alpha\gamma}_{Ga} \partial_\beta G^a_\gamma \right) \partial \xi_\alpha \otimes d\xi^\beta \tag{7.44}$$

So that, conversely, the momenta can be derived from the Energy-Momentum tensor, in a way which is usual in fluid mechanics : $\Pi^\alpha_{\nabla ij} = \frac{\partial T}{\partial \partial_\beta \psi^{ij}}, ...$

With the perturbative lagrangian in a model of the first type :

$\sum_{ij} \frac{\partial L}{\partial \partial_\alpha \psi^{ij}} \partial_\beta \psi^{ij} = C_I \mu \frac{1}{i} \sum_{ij} \left(\frac{\partial}{\partial \partial_\alpha \psi^{ij}} V^\gamma \langle \psi, \nabla_\gamma \psi \rangle \right) \partial_\beta \psi^{ij}$

$= C_I \mu \frac{1}{i} \sum_{ij} \frac{\partial}{\partial \partial_\alpha \psi^{ij}} \left(V^\gamma Tr \left([\psi^*] [\gamma_0] [\partial_\gamma \psi] \right) \right) \partial_\beta \psi^{ij}$

$= C_I \mu \frac{1}{i} \sum_{ij} \frac{\partial}{\partial \partial_\alpha \psi^{ij}} \left(V^\gamma \left(\sum_{pqr} ([\psi^*])^p_q [\gamma_0]^q_r [\partial_\gamma \psi]^r_p \right) \right) [\partial_\beta \psi]^i_j$

$= C_I \mu \frac{1}{i} V^\alpha Tr \left([\psi^*] [\gamma_0] [\partial_\beta \psi] \right)$

$= C_I \mu \frac{1}{i} V^\alpha \langle \psi, \partial_\beta \psi \rangle$

$\frac{\partial L}{\partial \partial_\beta G^a_{r\alpha}} = -4 C_G \mathcal{F}^{a\alpha\beta}_r$

$\frac{dL}{d\partial_\beta G^a_{w\alpha}} = 4 C_G \mathcal{F}^{a\alpha\beta}_w$

$\sum_{a\gamma} \frac{\partial L}{\partial \partial_\alpha G^a_\gamma} \partial_\beta G^a_r + \frac{dL}{d\partial_\beta G^a_{w\alpha}} \partial_\beta G^a_w$

$= 4 C_G \sum_{a\gamma} \mathcal{F}^{a\gamma\alpha}_w \partial_\beta G^a_w - \mathcal{F}^{a\gamma\alpha}_r \partial_\beta G^a_r$

$= 16 C_G \langle \mathcal{F}^{\alpha\gamma}_G, \partial_\beta G \rangle_{Cl}$

$\frac{dL}{d\partial_\beta \grave{A}^a_\alpha} = -4 C_A \mathcal{F}^{a\alpha\beta}_A$

$\sum_{a\gamma} \frac{\partial L}{\partial \partial_\alpha \grave{A}^a_\gamma} \partial_\beta \grave{A}^a_\gamma = -4 C_A \sum_{a\gamma} \mathcal{F}^{a\gamma\alpha}_A \partial_\beta \grave{A}^a_\gamma = 4 C_A \sum_{a\gamma} \mathcal{F}^{a\alpha\gamma}_A \partial_\beta \grave{A}^a_\gamma = 4 C_A \sum_\gamma \left\langle \mathcal{F}^{\alpha\gamma}_A, \partial_\beta \grave{A}_\gamma \right\rangle_{T_1 U}$

$$T^\alpha_\beta = C_I \mu \frac{1}{i} V^\alpha \langle \psi, \partial_\beta \psi \rangle + 4 \sum_{\gamma=0}^{3} 4 C_G \langle \mathcal{F}^{\alpha\gamma}_G, \partial_\beta G_\gamma \rangle_{Cl} + C_A \left\langle \mathcal{F}^{\alpha\gamma}_A, \partial_\beta \grave{A}_\gamma \right\rangle_{T_1 U} \tag{7.45}$$

So, with any lagrangian one can compute an explicit energy-momentum tensor. It is usually assumed that the energy momentum tensor is symmetric : $T^\alpha_\beta = T^\beta_\alpha$, which it should be because, in the Einstein equation, the Ricci tensor, with the Lévy-Civita connection, is symmetric. In our framework T is usually not symmetric.

Energy-momentum tensor and tetrad equation

This formula shows a strong similitude with the tetrad equation, which reads :

$\forall\alpha,\beta=0...3:C_I\frac{1}{i}V^\alpha\left\langle\psi,\nabla_\beta\psi\right\rangle+4\sum_\gamma\{C_G\sum_{a=1}^3\mathcal{F}_{r\beta\gamma}^a\mathcal{F}_r^{a\alpha\gamma}-\mathcal{F}_{w\beta\gamma}^a\mathcal{F}_w^{a\alpha\gamma}+C_A\sum_{a=1}^m\mathcal{F}_{A\beta\gamma}^a\mathcal{F}_A^{a\alpha\gamma}\}$

$=2\delta_\alpha^\beta\left(\sum_{\lambda\mu}\left(C_G\left\langle\mathcal{F}_G,\mathcal{F}_G\right\rangle+C_A\left\langle\mathcal{F}_A,\mathcal{F}_A\right\rangle\right)\right)$

As seen before (Currents - Definition) :

$C_I\frac{1}{i}V^\alpha\left\langle\psi,\nabla_\beta\psi\right\rangle=C_I\frac{1}{i}V^\alpha\left\langle\psi,\partial_\beta\psi\right\rangle+8\left(4C_G\left\langle J_G^\alpha,G_\beta\right\rangle_{Cl}+C_A\left\langle J_A^\alpha,\grave{A}_\beta\right\rangle_{T_1U}+C_{EM}J_{EM}^\alpha\grave{A}_\beta\right)$

$4\sum_{\gamma=0}^3\sum_{a=1}^3\mathcal{F}_{r\beta\gamma}^a\mathcal{F}_r^{a\alpha\gamma}-\mathcal{F}_{w\beta\gamma}^a\mathcal{F}_w^{a\alpha\gamma}=16\sum_{\gamma=0}^3\left\langle\mathcal{F}_{G\beta\gamma},\mathcal{F}_G^{\alpha\gamma}\right\rangle_{Cl}$

$=16\sum_{\gamma=0}^3\left\langle dG_{\beta\gamma},\mathcal{F}_G^{\alpha\gamma}\right\rangle_{Cl}+2\left\langle\left[G_\beta,G_\gamma\right],\mathcal{F}_G^{\alpha\gamma}\right\rangle_{Cl}$

$=16\sum_{\gamma=0}^3\left\langle\partial_\beta G_\gamma,\mathcal{F}_G^{\alpha\gamma}\right\rangle_{Cl}-\left\langle\partial_\gamma G_\beta,\mathcal{F}_G^{\alpha\gamma}\right\rangle_{Cl}+2\left\langle G_\alpha,\left[G_\gamma,\mathcal{F}_G^{\beta\gamma}\right]\right\rangle_{Cl}$

$=16\sum_{\gamma=0}^3\left\langle\partial_\beta G_\gamma,\mathcal{F}_G^{\alpha\gamma}\right\rangle_{Cl}-\left\langle\partial_\gamma G_\beta,\mathcal{F}_G^{\alpha\gamma}\right\rangle_{Cl}-2\left\langle G_\alpha,\left[\mathcal{F}_G^{\beta\gamma},G_\gamma\right]\right\rangle_{Cl}$

$=-32\left\langle G_\alpha,\phi_{G\beta}\right\rangle_{Cl}+16\sum_{\gamma=0}^3\left\langle\partial_\beta G_\gamma,\mathcal{F}_G^{\alpha\gamma}\right\rangle_{Cl}-\left\langle\partial_\gamma G_\beta,\mathcal{F}_G^{\alpha\gamma}\right\rangle_{Cl}$

$4\sum_\gamma\sum_{a=1}^m\mathcal{F}_{A\beta\gamma}^a\mathcal{F}_A^{a\alpha\gamma}=4\sum_\gamma\left\langle\mathcal{F}_{A\beta\gamma},\mathcal{F}_A^{\alpha\gamma}\right\rangle_{T_1U}$

$=4\sum_\gamma\left\langle d\grave{A}_{\beta\gamma},\mathcal{F}_A^{\alpha\gamma}\right\rangle_{T_1U}+2\left\langle\left[\grave{A}_\beta,\grave{A}_\gamma\right],\mathcal{F}_A^{\alpha\gamma}\right\rangle_{T_1U}$

$=4\sum_\gamma\left\langle d\grave{A}_{\beta\gamma},\mathcal{F}_A^{\alpha\gamma}\right\rangle_{T_1U}+2\left\langle\grave{A}_\beta,\left[\grave{A}_\gamma,\mathcal{F}_A^{\alpha\gamma}\right]\right\rangle_{T_1U}$

$=-8\left\langle\grave{A}_\beta,\phi_A\right\rangle_{T_1U}+4\sum_\gamma\left\langle\partial_\beta\grave{A}_\gamma,\mathcal{F}_A^{\alpha\gamma}\right\rangle_{T_1U}-\left\langle\partial_\gamma\grave{A}_\beta,\mathcal{F}_A^{\alpha\gamma}\right\rangle_{T_1U}$

$C_I\frac{1}{i}V^\alpha\left\langle\psi,\nabla_\beta\psi\right\rangle+4\sum_\gamma\{C_G\sum_{a=1}^3\mathcal{F}_{r\beta\gamma}^a\mathcal{F}_r^{a\alpha\gamma}-\mathcal{F}_{w\beta\gamma}^a\mathcal{F}_w^{a\alpha\gamma}+C_A\sum_{a=1}^m\mathcal{F}_{A\beta\gamma}^a\mathcal{F}_A^{a\alpha\gamma}\}$

$=C_I\frac{1}{i}V^\alpha\left\langle\psi,\partial_\beta\psi\right\rangle+32C_G\left\langle J_G^\alpha,G_\beta\right\rangle_{Cl}-32C_G\left\langle G_\alpha,\phi_{G\beta}\right\rangle_{Cl}+8C_A\left\langle J_A^\alpha,\grave{A}_\beta\right\rangle_{T_1U}-8C_A\left\langle\grave{A}_\beta,\phi_A^\alpha\right\rangle_{T_1U}$

$+16\sum_{\gamma=0}^3C_G\left\langle\partial_\beta G_\gamma,\mathcal{F}_G^{\alpha\gamma}\right\rangle_{Cl}-C_G\left\langle\partial_\gamma G_\beta,\mathcal{F}_G^{\alpha\gamma}\right\rangle_{Cl}+4\sum_\gamma C_A\left\langle\partial_\beta\grave{A}_\gamma,\mathcal{F}_A^{\alpha\gamma}\right\rangle_{T_1U}-C_A\left\langle\partial_\gamma\grave{A}_\beta,\mathcal{F}_A^{\alpha\gamma}\right\rangle_{T_1U}$

On shell : $J_G^\alpha=G_\alpha,J_A^\alpha=\phi_A^\alpha$ and the tetrad equation can be expressed as :

$$\forall\alpha,\beta=0...3:T_\beta^\alpha=4\sum_{\gamma=0}^3\left(4C_G\left\langle\partial_\gamma G_\beta,\mathcal{F}_G^{\alpha\gamma}\right\rangle_{Cl}+C_A\left\langle\partial_\gamma\grave{A}_\beta,\mathcal{F}_A^{\alpha\gamma}\right\rangle_{T_1U}\right)+\delta_\alpha^\beta L_{Shell}\qquad(7.46)$$

with :

$L_{Fields}=2\left\{C_G\left\langle\mathcal{F}_G,\mathcal{F}_G\right\rangle_G+C_A\left\langle\mathcal{F}_A,\mathcal{F}_A\right\rangle_A+C_{EM}\left\langle\mathcal{F}_{EM},\mathcal{F}_{EM}\right\rangle\right\}=L_{Shell}$

For a deformable solid which is not submitted to external fields the interactions induced by the motion of the particles can be seen as the forces resulting from the deformation. So one can state similarly that for any deformation $(\delta r,\delta w)$ the energy-momentum tensor provides, in a continuous deformation at equilibrium, the value of the induced fields, or equivalently the deformation induced by external fields.

The obstruction to the symmetry of T comes here from the first term and T would not be symmetric even with the Lévy-Civita connection.

The Einstein equation is the starting point for models in cosmology. However, as it comes from the application of the Principle pf Least Action, it leads, in the vacuum ($T=0$) to static universes. This is an issue that we have met several times. The usual solution is to add a cosmological constant Λ in the Einstein equation, which becomes, in the vacuum :

$Ric_{\alpha\beta}-\frac{1}{2}\mathbf{R}g_{\alpha\beta}=\frac{1}{2}\mathbf{\Lambda}g_{\alpha\beta}$

this sums up to state that there is an energy-momentum tensor for the vacuum, and it is usually assumed that Λ is not null.

Following a similar idea, we can propose an alternate tetrad equation in the vacuum.

7.4.4 Alternate tetrad equation

In the vacuum the main feature of the propagation of the field is the existence of a function F, specific to the field, which defines the front waves, the vector V, and is related to the energy of the interacting field, that is $\langle \mathcal{F}, \mathcal{F} \rangle_G$. So an alternate solution is to replace, in the right hand side of the tetrad equation, $\langle \mathcal{F}, \mathcal{F} \rangle_G$ by $\mathcal{E}_0 - F$.

More precisely, in the vacuum with only the gravitational field the tetrad equation reads :
$\forall \alpha, \beta = 0...3 : \sum_{a=1}^3 \mathcal{F}^a_{r\beta\gamma} \mathcal{F}^{a\alpha\gamma}_r - \mathcal{F}^a_{w\beta\gamma} \mathcal{F}^{a\alpha\gamma}_w = \frac{1}{2} \delta^\beta_\alpha \langle \mathcal{F}_G, \mathcal{F}_G \rangle$
and the alternate equation is the following, in the vacuum :

$$\forall \alpha, \beta = 0...3 : \sum_{\gamma=0}^3 \sum_{a=1}^3 \mathcal{F}^a_{r\alpha\gamma} \mathcal{F}^{a\beta\gamma}_r - \mathcal{F}^a_{w\alpha\gamma} \mathcal{F}^{a\beta\gamma}_w = \delta^\beta_\alpha \frac{1}{2} (\mathcal{E}_0 - F(m)) \tag{7.47}$$

i) With :
$\sum_{a=1}^3 \mathcal{F}^a_{r\alpha\gamma} \mathcal{F}^{a\beta\gamma}_r - \mathcal{F}^a_{w\alpha\gamma} \mathcal{F}^{a\beta\gamma}_w = \left([X_r]^t [X_r] - [X_w]^t [X_w] \right) [M]^\beta_\alpha$
$[M] = \begin{bmatrix} -\left[\widehat{\mathcal{F}}^w\right] [g_3]^{-1} \left[\widehat{\mathcal{F}}^w\right]^t & \left[\widehat{\mathcal{F}}^w\right] [g_3]^{-1} j(\widehat{\mathcal{F}}^r) [g_3]^{-1} \\ -j\left(\widehat{\mathcal{F}}^r\right) [g_3]^{-1} \left[\widehat{\mathcal{F}}^w\right]^t & \left(j\left(\widehat{\mathcal{F}}^r\right) [g_3]^{-1} j(\mathcal{F}^r) - \left[\widehat{\mathcal{F}}^w\right]^t \left[\widehat{\mathcal{F}}^w\right] \right) [g_3]^{-1} \end{bmatrix}$
$\left([X_r]^t [X_r] - [X_w]^t [X_x] \right) \widehat{F} = \frac{1}{2} (\mathcal{E}_0 - F)$
the equation reads in matricial form :
$[M] = \widehat{F}(m) I_4$
Using :
$\left[\widehat{\mathcal{F}}^w\right] = \left[d\widehat{\mathcal{K}}^r\right] [Q'] ; \left[\widehat{\mathcal{F}}^r\right] = -(\det Q') \left[d\widehat{\mathcal{K}}^w\right] [Q]^t$
$[M] = \begin{bmatrix} -\left[d\widehat{\mathcal{K}}^r\right] \left[d\widehat{\mathcal{K}}^r\right]^t & -\left[d\widehat{\mathcal{K}}^r\right] j(d\widehat{\mathcal{K}}^w) [Q]^t \\ [Q']^t j(d\widehat{\mathcal{K}}^w) \left[d\widehat{\mathcal{K}}^r\right] & [Q']^t \left(j(d\widehat{\mathcal{K}}^w) j(d\widehat{\mathcal{K}}^w) - \left[d\widehat{\mathcal{K}}^r\right]^t \left[d\widehat{\mathcal{K}}^r\right] \right) [Q]^t \end{bmatrix}$
i) Then a computation similar to the previous one gives :

$$d\widehat{\mathcal{K}}^r = \pm 2 d\widehat{\mathcal{K}}^w = \epsilon 2 d\widehat{\mathcal{K}}^w \tag{7.48}$$

$$\widehat{F} = -2 \left[d\widehat{\mathcal{K}}^w\right] \left[d\widehat{\mathcal{K}}^w\right]^t \tag{7.49}$$

As a consequence :

$$\left[\widehat{\mathcal{F}}^w\right] = -\epsilon 2 (\det Q) \left[\widehat{\mathcal{F}}^r\right] [g_3] \tag{7.50}$$

$$\widehat{F} = -2 (\det Q)^2 \left[\widehat{\mathcal{F}}^r\right] [g_3] \left[\widehat{\mathcal{F}}^r\right]^t \tag{7.51}$$

and the tensor $\mathcal{F}_G$ is defined by the matrix $[\mathcal{F}^r_r]$ and the fixed scalar $\lambda : [X_r] = \lambda [X_w]$
$[\mathcal{F}^w_r] = [X_r] \left[\widehat{\mathcal{F}}^w\right] = -\epsilon 2 (\det Q) [\mathcal{F}^r_r] [g_3]$
$[\mathcal{F}^w_w] = [X_w] \left[\widehat{\mathcal{F}}^w\right] = -\frac{1}{\lambda} \epsilon 2 (\det Q) [\mathcal{F}^r_r] [g_3]$

$[\mathcal{F}_w^r] = [X_w]\left[\widehat{\mathcal{F}}^r\right] = \frac{1}{\lambda}[\mathcal{F}_r^r]$
ii) As we had : $[X_r] = \lambda [X_w]$
$\langle \mathcal{F}, \mathcal{F}\rangle_G = (\mathcal{E}_0 - F) = 2\left([X_r]^t [X_r] - [X_w]^t [X_x]\right)\widehat{F}$

$$F(m) = \mathcal{E}_0 + 4\left(\lambda^2 - 1\right)(\det Q)^2 [X_w]^t [X_w]\left[\widehat{\mathcal{F}}^r\right][g_3]\left[\widehat{\mathcal{F}}^r\right]^t \tag{7.52}$$

$F(m) = \mathcal{E}_0 + 4\left(\lambda^2 - 1\right)\left([X_w]^t [X_w]\right)\left[d\widehat{\mathcal{K}}^w\right]\left[d\widehat{\mathcal{K}}^w\right]^t$
iii) From the codifferential equation we have for $[X]$:
$[\partial X_w]\left[d\widehat{\mathcal{K}}^r\right]^t = 0$
$[\partial_0 X_r]\left[d\widehat{\mathcal{K}}^r\right] + [\partial X_r]\, j\left(\left[d\widehat{\mathcal{K}}^w\right]^t\right) = 0$
read :
$[\partial X_w]\left[d\widehat{\mathcal{K}}^w\right]^t = 0$
$\epsilon 2[\partial_0 X_w]\left[d\widehat{\mathcal{K}}^w\right] + [\partial X_w]\, j\left(d\widehat{\mathcal{K}}^w\right) = 0$
By multiplication of the second by $\left[d\widehat{\mathcal{K}}^w\right]^t$:
$\epsilon 2[\partial_0 X_w]\left[d\widehat{\mathcal{K}}^w\right]\left[d\widehat{\mathcal{K}}^w\right]^t = 0 = -4\epsilon[\partial_0 X_w]\widehat{F} \Rightarrow [\partial_0 X_w] = [\partial_0 X_r] = 0$
and we are left with the equations
$[\partial X_w]\, j\left(d\widehat{\mathcal{K}}^w\right) = 0$
$[\partial X_w]\left[d\widehat{\mathcal{K}}^w\right]^t = 0$
which have for unique solution $[\partial X_w] = 0$. So $[X_r] = \lambda [X_w], [X_w]$ are constant, and determined by their initial values on the border of Ω.

$$[X_G] = Ct \tag{7.53}$$

iv) We had from the current equations :
$\upsilon\left(d\left(\mu G_{w\alpha\beta}\right), dG_{w\alpha\beta}\right) = \upsilon\left(\lambda X_w^a d\left(\widehat{\mathcal{K}}_{\alpha\beta}\right), X_w^a d\left(\widehat{\mathcal{K}}_{\alpha\beta}\right)\right)$
As a consequence :
$\upsilon\left(d\left(\mu G_{w\alpha\beta}\right), dG_{w\alpha\beta}\right) = \upsilon\left(\lambda d\left(X_w^a \widehat{\mathcal{K}}_{\alpha\beta}\right), d\left(X_w^a \widehat{\mathcal{K}}_{\alpha\beta}\right)\right) \Rightarrow \mu = \lambda = Ct$

$$\mathcal{F}_{G\alpha\beta} = \upsilon\left(\lambda d\left(X_w^a \widehat{\mathcal{K}}_{\alpha\beta}\right), d\left(X_w^a \widehat{\mathcal{K}}_{\alpha\beta}\right)\right) = \upsilon\left(\lambda dG_{w\alpha\beta}, dG_{w\alpha\beta}\right) \tag{7.54}$$

v) $\pounds_V \widehat{\mathcal{F}} = 0$ gives the equations for $\left[d\widehat{\mathcal{K}}^w\right], [Q], V$:
$\sum_{\alpha=0}^3 -V^\alpha \partial_\alpha\left(-(\det Q')\left[d\widehat{\mathcal{K}}^w\right][Q]^t\right) = -(\det Q')\left[d\widehat{\mathcal{K}}^w\right][Q]^t\left([\partial v]^t - (div\,(v))\, I_3\right) - \left[d\widehat{\mathcal{K}}^r\right][Q']\, j\left(\left[\partial V^0\right]\right)$
$\sum_{\alpha=0}^3 V^\alpha \partial_\alpha\left(\left[d\widehat{\mathcal{K}}^r\right][Q']\right) = -(\det Q')\left[d\widehat{\mathcal{K}}^w\right][Q]^t\, j\left(\partial_0 V\right) - \left[d\widehat{\mathcal{K}}^r\right][Q']\left(\partial_0 V^0 + [\partial v]\right)$
which read with $[U] = \sum_{\alpha=0}^3 V^\alpha [Q'][\partial_\alpha Q]$:
$\sum_{\alpha=0}^3 V^\alpha\left[\partial_\alpha d\widehat{\mathcal{K}}^w\right]$
$= -\left[d\widehat{\mathcal{K}}^w\right]\left(-Tr\,[U] + [U]^t - (div\,(v))\, I_3 + \left([Q'][\partial v][Q]\right)^t + \epsilon 2j\left(\left[\partial V^0\right][Q]\right)\right)$
$= -\left[d\widehat{\mathcal{K}}^w\right]\left(-[U] + \partial_0 V^0 + [Q'][\partial v][Q] + \frac{1}{2\epsilon} j\left([Q']\,\partial_0 V\right)\right)$
vi) The equation $divV = 0$ reads :
$\sum_{\alpha=0}^3 \frac{1}{\det Q'} V^\alpha \partial_\alpha \det Q' + \partial_\alpha V^\alpha = \sum_{\alpha=0}^3 V^\alpha Tr\left[\partial_\alpha Q'\right][Q] + \partial_\alpha V^\alpha = 0$

$\partial_0 V^0 + div(v) = \sum_{\alpha=0}^3 \partial_\alpha V^\alpha = -\sum_{\alpha=0}^3 V^\alpha Tr[\partial_\alpha Q'][Q] = \sum_{\alpha=0}^3 V^\alpha Tr[Q'][\partial_\alpha Q] = Tr[U]$
$div(v) = Tr[U] - \partial_0 V^0$
$\sum_{\alpha=0}^3 V^\alpha \left[\partial_\alpha d\widehat{\mathcal{K}}^w\right]$
$= -\left[d\widehat{\mathcal{K}}^w\right]\left(-2Tr[U] + [U]^t + \partial_0 V^0 + ([Q'][\partial v][Q])^t + \epsilon 2j\left(\left[\partial V^0\right][Q]\right)\right)$
$= -\left[d\widehat{\mathcal{K}}^w\right]\left(-[U] + \partial_0 V^0 + [Q'][\partial v][Q] + \frac{1}{2\epsilon} j([Q']\partial_0 V)\right)$
The equality of the two last expressions give :

$$\left[d\widehat{\mathcal{K}}^w\right]\left(-2Tr[U] + [U] + [U]^t + ([Q'][\partial v][Q])^t - [Q'][\partial v][Q] + \epsilon j\left(2\left[\partial V^0\right][Q] - \frac{1}{2}[Q'][\partial_0 V]\right)\right) = 0 \tag{7.55}$$

and :

$$\sum_{\alpha=0}^{3} V^\alpha \left[\partial_\alpha d\widehat{\mathcal{K}}^w\right] = -\left[d\widehat{\mathcal{K}}^w\right]\left(-[U] + \partial_0 V^0 + [Q'][\partial v][Q] + \frac{1}{2\epsilon} j([Q']\partial_0 V)\right) \tag{7.56}$$

vii) The equations, $V = gradF, \langle V, V\rangle = 0$ give :
$V = gradF = \sum_{\alpha,\beta=0}^3 g^{\alpha\beta}(\partial_\beta F)\partial\xi_\alpha$
$= -2\left(4\left(\lambda^2-1\right)\left([X_w]^t[X_w]\right)\right)\sum_{\alpha,\beta,\lambda=0}^3 g^{\alpha\beta}\left(\left[\partial_\beta d\widehat{\mathcal{K}}^w\right]\left[d\widehat{\mathcal{K}}^w\right]^t\right)\partial\xi_\alpha$
$\langle V, V\rangle = 0 = \sum_{\gamma=0}^3 g_{\gamma\alpha} V^\gamma \left(-2\left(4\left(\lambda^2-1\right)\left([X_w]^t[X_w]\right)\right)\sum_{\alpha,\beta=0}^3 g^{\alpha\beta}\left(\left[\partial_\beta d\widehat{\mathcal{K}}^w\right]\left[d\widehat{\mathcal{K}}^w\right]^t\right)\right)$
$\sum_{\gamma=0}^3 g_{\gamma\alpha} V^\gamma \sum_{\alpha,\beta=0}^3 g^{\alpha\beta}\left(\left[\partial_\beta d\widehat{\mathcal{K}}^w\right]\left[d\widehat{\mathcal{K}}^w\right]^t\right) = 0$
$\sum_{\beta=0}^3 V^\beta \left[\partial_\beta d\widehat{\mathcal{K}}^w\right]\left[d\widehat{\mathcal{K}}^w\right]^t = 0$
$\left(\sum_{\beta=0}^3 V^\beta \left[\partial_\beta d\widehat{\mathcal{K}}^w\right]\right)\left[d\widehat{\mathcal{K}}^w\right]^t = 0$
Using the two equations above :
$\sum_{\beta=0}^3 V^\beta \left[\partial_\beta d\widehat{\mathcal{K}}^w\right]\left[d\widehat{\mathcal{K}}^w\right]^t$
$= \left[d\widehat{\mathcal{K}}^w\right]\left(-2Tr[U] + [U]^t + \partial_0 V^0 + ([Q'][\partial v][Q])^t + \epsilon 2j\left(\left[\partial V^0\right][Q]\right)\right)\left[d\widehat{\mathcal{K}}^w\right]^t = 0$
$= \left[d\widehat{\mathcal{K}}^w\right]\left(-[U] + \partial_0 V^0 + [Q'][\partial v][Q] + \frac{1}{2\epsilon} j([Q']\partial_0 V)\right)\left[d\widehat{\mathcal{K}}^w\right]^t = 0$
$\left[d\widehat{\mathcal{K}}^w\right] j\left(\left[\partial V^0\right][Q]\right)\left[d\widehat{\mathcal{K}}^w\right]^t = \left[d\widehat{\mathcal{K}}^w\right] j([Q']\partial_0 V)\left[d\widehat{\mathcal{K}}^w\right]^t = 0$
Transpose the first equation :
$\left[d\widehat{\mathcal{K}}^w\right]\left(-2Tr[U] + [U] + \partial_0 V^0 + [Q'][\partial v][Q]\right)\left[d\widehat{\mathcal{K}}^w\right]^t = 0$
Substract to the second :

$$\left[d\widehat{\mathcal{K}}^w\right](Tr[U] - [U])\left[d\widehat{\mathcal{K}}^w\right]^t = 0 \tag{7.57}$$

If V varies slowly, at the first order :
$\sum_{\alpha=0}^3 V^\alpha \left[\partial_\alpha d\widehat{\mathcal{K}}^w\right] = \left[d\widehat{\mathcal{K}}^w\right][U]$

$\left[d\widehat{\mathcal{K}}^w\right]\left(-2Tr\left[U\right]+\left[U\right]+\left[U\right]^t\right)=0$
$\Rightarrow \left[d\widehat{\mathcal{K}}^w\right]\left[U\right]=2\left(Tr\left[U\right]\right)\left[d\widehat{\mathcal{K}}^w\right]-\left[d\widehat{\mathcal{K}}^w\right]\left[U\right]^t$
$\left[d\widehat{\mathcal{K}}^w\right]\left(Tr\left[U\right]-\left[U\right]\right)\left[d\widehat{\mathcal{K}}^w\right]^t=0$
$\Rightarrow \left[d\widehat{\mathcal{K}}^w\right]\left[U\right]\left[d\widehat{\mathcal{K}}^w\right]^t=\left(Tr\left[U\right]\right)\left[d\widehat{\mathcal{K}}^w\right]\left[d\widehat{\mathcal{K}}^w\right]^t$
If $\left[U\right]=\left[U\right]^t$
$\left[d\widehat{\mathcal{K}}^w\right]\left[U\right]=\left(Tr\left[U\right]\right)\left[d\widehat{\mathcal{K}}^w\right]$
$\sum_{\alpha=0}^{3}V^{\alpha}\left[\partial_{\alpha}d\widehat{\mathcal{K}}^w\right]=\left(Tr\left[U\right]\right)\left[d\widehat{\mathcal{K}}^w\right]=\left[\frac{d}{ds}d\widehat{\mathcal{K}}^w\right]$
Which gives $d\widehat{\mathcal{K}}$ on the integral curves of V (with the additional constraint that this is a differential).

The computation of the gravitational field, in the framework of GR, which is so far the only one that we know, is a difficult endeavour. Explicit solutions of the problem, using the traditional Lévy-Civita connection and the scalar curvature, are known only for spherical symmetric systems, rotating or not, or for specific cases (Schwarzschild, Reissner-Nordstrom, Kerr, Kerr-Newman). The computations given here are speculative, but show that, in a more general framework, it is possible to have a manageable model, which could enlarge the scope of solutions.

Chapter 8

DISCONTINUOUS PROCESSES

Continuous models address a large scope of problems. They represent ideal physical cases : no collision, no discontinuity, no change in the number or the characteristics of the particles. By analogy with fluid mechanics they represent steady flows. These limitations can be alleviated, by the introduction of densities or collision operators. And if an equilibrium is not necessarily the result of a continuous process, in the physical world, no process is totally discontinuous : the discontinuity appears as a singular event, between periods of equilibrium. The Principle of Least Action and continuous models hold for the conditions existing before and after the discontinuity. Meanwhile discontinuous models are focused on the transitions between equilibrium.

Many physical phenomena involve, at some step, processes which are discontinuous.

At our scale : collision or breaking of material bodies, shock-waves on fluids, change of phase,...

At the atomic scale : collision of molecules or particles, elastic (without loss of energy) or not, disintegration of a nucleus, spontaneous or following collisions, creation or annihilation of particles, change of relativist spin,...

If discontinuous processes are ubiquitous, they present an issue for the Physicists. There is no general method to deal with them. It is a fact that we have by far more convenient and powerful mathematical tools to deal with smooth variables than with discontinuous ones, even if, in the practical computation, one uses numerical (and discontinuous) methods. Whatever our personal preferences, it suffices to open any book on Physics to see that, as quickly as possible, one comes back to more comfortable differential equations. Models of discontinuous processes naturally rely on statistics and probability. This dichotomy has an important impact on the theories. The study of discontinuous process leads naturally to probabilist, non determinist models. At the atomic level they are prevalent, all the more so that most experiments are focused on them. And since all proceeds from the atomic level, this leads to a bias towards a discreet, probabilist, weltanschauung, which is obvious in many interpretations of QM. When one has a hammer, everything looks like a nail. But, for practical purpose, the border between continuous / discontinuous depends on the scale. Many discontinuous phenomena can be dealt with in continuous models if one accepts to neglect what happens at the basic level : this is at the foundation of Fluid Mechanics and Thermodynamics. We do not know what is the physical world, one can only try to find its most sensible and efficient representations, and must not be confused, taking our representations, or worse, our formalism, for the real world. In the Copenhagen interpretation of QM, it is assumed that there are two Physics, one which applies at the atomic scale, and another to the usual world. Actually the border should be between continuous and discontinuous processes, and this border depends on the scale considered. They require different types of representations, depending of the purpose or the problem, but the difference in the formalism is not the proof of a dichotomic world, and even less of a continuous or discontinuous world.

If we acknowledge the existence of discontinuities in solids or fluids, we should consider their existence in force fields. So one should accept the idea that fields are not necessarily represented by smooth maps, and find a way to represent discontinuities of the fields themselves. This is the main purpose of this chapter. We will see how to deal with discontinuities in fields, how they can be represented in the framework that we have used so far, and show that, actually, these discontinuities "look like" particles : bosons, the force carriers of the Standard Model, can be seen as discontinuities of the fields.

8.1 BOSONS

Wave propagations, wave packets, burst of fields, or solitons are continuous processes : they are solutions, sometimes very specific to initial conditions or to the nature of the medium in which the field propagates, of regular differential equations such as $\square A = 0$. The fields stay represented by smooth maps. Discontinuities are different : the maps are no longer smooth. At the macroscopic level we have shock waves.

This is also during their propagation that discontinuities appear for force fields, so the variable to consider is the derivative of the connection. However the analogy with Fluid Mechanics is deceptive : the key variable in the representation of a fluid is a vector field, which provides, through the integral curves, a natural folliation which represents waves of same phase in the propagation. For Force Fields the key variable is the connection, which is a form on a tangent bundle. And we have shown (in the subsection Propagation) that there are no hypersurfaces over which the connection would have the same value. So discontinuities in force fields do not appear as "shock waves", that is along hypersurfaces, but along curves. We have a special vector field V provided by the function F, along which, in an adapted chart, the strength $\mathcal{F}$ keeps the same components. So that discontinuities in the force fields will appear as discontinuities of the derivative of the connection, localized along V.

8.1.1 The mathematical representation of discontinuities of force fields

The idea is that the connection is continuous, differentiable except in some point, where there is a right and left derivative, along an integral curve of the vector field V. We first give a representation of such a discontinuity. We take the field in P_U and connection $\grave{\mathbf{A}}$ as example, but this holds also for the gravitational field.

Discontinuity of the derivative at a point

Theorem 110 *A discontinuity $\Delta\grave{\mathbf{A}}$ of the strength of the field along a vector field can be represented as a one form on TM valued in the adjoint bundle $\Delta\grave{\mathbf{A}} \in \Lambda_1 (TM; P_U [T_1U, Ad])$.*

1) The most general definition of the derivative of a tensor is through the Lie derivative. The Lie derivative of a tensor T on the tangent space of a manifold, along any vector field V (or a curve, by taking its tangent) is :

$\pounds_V T = \frac{d}{ds}\Phi_V (s, m)^* \mathcal{F} (m) |_{s=0}$

It can be extended to any tensor valued in a fixed vector space, and it holds for any manifold.

The connection $\grave{\mathbf{A}}$ is a projection from the tangent space TP_U to the vertical bundle VP_U. So this is a tensor on the tangent space to the manifold P_U. If the connection is principal it is defined by a unique map and the connection form $\widehat{\grave{A}}$ is the tensor on TP_U valued in the fixed space T_1U :

$\widehat{\grave{A}} (p) : T_pP_U \to T_1U :: \grave{\mathbf{A}} (p) (v_p) = \zeta \left(\widehat{\grave{A}} (p) (v_p)\right) (p)$

The Lie derivative is defined on TP_U, but we want to keep the link with the base M. We can lift a vector field (or a curve) from TM to TP_U by the horizontal lift :

$\chi_L : \mathfrak{X} (TM) \to \mathfrak{X} (HP_U) :: \chi_L (p) (V) = \varphi'_{Gm} (m, g) V (m) - \zeta \left(Ad_{g^{-1}} \grave{A} (m) V (m)\right) (p) \in H_pP$

and we denote $W = \chi_L (V) \in \mathfrak{X} (HP_U)$. Thus $\grave{\mathbf{A}} (p) (W) = 0$.

We need a section $p \in \mathfrak{X}(P_U)$ to go from M to $P_U : p(m) \in P_U$

The Lie derivative of the tensor $\widehat{\grave{A}}$ on TP_U along W is :

$\pounds_W \widehat{\grave{A}} = \frac{d}{ds}\Phi_W(s,p)^* \widehat{\grave{A}}(p)|_{s=0}$

where s is the parameter along the integral curve of W going through p. Using the section p the definition is then consistent and one can go from M to P_U.

The derivative above is defined as any derivative :

$\frac{d}{ds}\Phi_W(s,p)^* \widehat{\grave{A}}(p)|_{s=0} = \lim_{h\to 0}\Delta_R(h)$ with $\Delta_R(h) = \frac{1}{h}\left(\Phi_W(s+h,p)^* \widehat{\grave{A}}(p) - \Phi_W(s,p)^* \widehat{\grave{A}}(p)\right)$

which makes sense because the two terms belong to the same vector space at p.

If the derivative is continuous :

$\frac{d}{ds}\Phi_W(s,p)^* \widehat{\grave{A}}(p)|_{s=0} = \lim_{h\to 0}\Delta_L(h)$ with $\Delta_L(h) = \frac{1}{h}\left(\Phi_W(s,p)^* \widehat{\grave{A}}(p) - \Phi_W(s-h,p)^* \widehat{\grave{A}}(p)\right)$

But of course one can have a right and a left derivative which have a different value : we have a discontinuity, that we can denote

$$\left[\begin{array}{l} \Delta\widehat{\grave{A}}(p) = \lim_{h\to 0}\Delta_R(h) - \lim_{h\to 0}\Delta_L(h) \\ \Delta\grave{\mathbf{A}}(p) = \zeta\left(\Delta\widehat{\grave{A}}(p)\right)(p) \in V_p P_U \end{array}\right] \tag{8.1}$$

The Lie derivative $\pounds_W T$ of a tensor T is a tensor of same order and type as T. And similarly $\Delta\widehat{\grave{A}}(p) \in \Lambda_1(TP_U; T_1U)$.

2) The quantity $\Delta\grave{\mathbf{A}}(p)$ is a fundamental vector field, estimated at the point $p \in P_U$. In a change of gauge at the same point $m = \pi_U(p)$ by the right action ρ of $\chi(m) \in U$ its value change as :

$\Delta\grave{\mathbf{A}}(p) \to \widetilde{\Delta\grave{\mathbf{A}}(p)} = \rho'_p(p,\chi(m))\,\zeta\left(Ad_g\Delta\widehat{\grave{A}}(p)\right)(p)$

This is a change of gauge :

$\widetilde{\mathbf{p}} = \widetilde{\varphi}_U(m,1) = \varphi_U(m,\chi(m)) = \mathbf{p}\cdot\chi(m)$

$\zeta\left(\Delta\grave{A}(m)\right)(\widetilde{\mathbf{p}}) = \zeta\left(\Delta\grave{A}(m)\right)(\mathbf{p}\cdot\chi(m)) = \rho'_p(\mathbf{p},\chi(m))\,\zeta\left(Ad_{\chi(m)}\Delta\grave{A}((\mathbf{p}(m)))\right)(\widetilde{\mathbf{p}}(m))$

the measure changes as :

$\Delta\widehat{\grave{A}}(\mathbf{p}(m)) \to \widetilde{\Delta\widehat{\grave{A}}(\mathbf{p}(m))} = Ad_{\chi(m)}\Delta\grave{A}(\mathbf{p}(m))$

So $\Delta\grave{\mathbf{A}}(p)$ can be considered as a one form on TP_U valued in the adjoint vector bundle $P_U[T_1U, Ad]$. And using the pull back by the standard gauge $\mathbf{p}(m)^*\Delta\grave{\mathbf{A}} \in \Lambda_1(TM; P_U[T_1U, Ad])$

This result is important : the reason why a potential cannot explicitly be present in the lagrangian comes from its special rule in a change of gauge (see lagrangian), but this rule applies no longer to $\zeta\left(\Delta\widehat{\grave{A}}(\mathbf{p}(m))\right)(\mathbf{p}(m))$, which becomes a section of an associated vector bundle like the others.

3) The strength of the field $\mathcal{F}$ is a special derivative :

$\mathcal{F}_A(m) = -\mathbf{p}^*(m)\nabla_e\widehat{\grave{A}} = -\mathbf{p}^*(m)\chi^* d\widehat{\grave{A}}$

with the horizontal form :

$\chi(p) :: T_pP_U \to H_pP_U :: \chi(p)(v_p) = v_p - \grave{\mathbf{A}}(p)(v_p)$

The exterior differential $d\widehat{\grave{A}}$ of the form $\widehat{\grave{A}}$ valued in the fixed vector space T_1U is taken, through χ^* on horizontal vectors.

The Lie derivative and the exterior differential are related :

$\pounds_W\widehat{\grave{A}} = d\left(i_W\widehat{\grave{A}}\right) + i_W d\widehat{\grave{A}}$

but, because W is horizontal :

$i_W \widehat{\widehat{A}} = 0$

$\chi_* (W) = W$

$\Rightarrow \pounds_W \widehat{\widehat{A}} = i_W d\widehat{\widehat{A}} = i_W \chi^* d\widehat{\widehat{A}} = i_W \nabla_e \widehat{\widehat{A}}$

The result holds for any vector field $V \in \mathfrak{X}(TM)$ so one can write:

$\mathcal{F}_A (m) = -\mathbf{p}^* (m) \pounds \widehat{\widehat{A}}$

Along the vector field V, lifted on W :

$\Delta\mathcal{F}_A (V) = -\mathbf{p}^* (m) \pounds_W \widehat{\widehat{A}} = -\mathbf{p}^* (m) \Delta\grave{\mathbf{A}} (m)$

$\Delta\mathcal{F}_A (V), \mathbf{p}^* (m) \Delta\grave{\mathbf{A}} (m)$ are one form on M valued in the Lie algebra T_1U.

And a discontinuity of the exterior covariant derivative of the connection along the vector field V can be denoted :

$$\Delta\mathcal{F}_A (V) = -\mathbf{p}^* (m) \Delta\grave{\mathbf{A}} (m) \tag{8.2}$$

In the interactions, represented by continuous models, by construct the incoming $\mathcal{F}$ is equal to the outcoming $\mathcal{F}$: $\mathcal{F}$ is assumed to be continuous, which makes impossible to distinguish them properly and negate the causal structure. This can be corrected by acknowledging the possibility of a discontinuity.

For a photon $T_1U = \mathbb{R}$ and $\Delta\grave{\mathbf{A}} (m) \in \Lambda_1 (M; \mathbb{R})$.

Propagation of a discontinuity

The previous result holds for any vector field, at any point. By construct a discontinuity happens at a point. However a continuous 1 form $\Delta\grave{\mathbf{A}} \in \Lambda_1 (TM; T_1U)$ can be seen as defining a discontinuity of the connection at any point. A discontinuity of the field does not propagate on hypersurfaces but, in the picture above, it can be represented as 1 form $\Delta\grave{\mathbf{A}} \in \Lambda_1 (TM; T_1U)$ with support on an integral curve of the vector field V characteristic of the propagation of the field (the lift of a vector field as well as the Lie derivative can be defined along curves, using the tangent to the curve).

We have a picture similar to particles : an object living on a curve, and travelling on the curve with the parameter of the flow. Here the world line is an integral curve of the propagation of the field, and the proper time is the phase of the field. And we call **boson** such an object. The boson associated to the gravitational field is the graviton (which has never been observed). When only the gravitational and EM field are present the boson associated to the EM field is the photon (this is a composite boson when the weak and strong interactions are present).

State of a boson

Fundamental state of a boson

The quantization of $\mathcal{F}$ show that, along an integral curve we have : $\mathcal{F} (\Phi_V (s, n)) = \lambda Ad_{g(s,n)} \mathcal{F} (n)$ and the same reasoning leads to a similar formula : $\Delta\grave{\mathbf{A}} (\Phi_V (s, n)) = \lambda (s) Ad_{g(s,n)} \Delta\grave{\mathbf{A}} (n)$. For a particle its state ψ, is a geometric quantity which is preserved along its world line, its measure varying with the field, and we can expect to have something similar for bosons. The vector V is part of the definition of the boson, and the quantity :

$$B_0(s) = \sum_{\alpha=0}^{3} V^\alpha \Delta\mathcal{F}_A(V)(\Phi_V(s,n)) \in T_1U \tag{8.3}$$

does not depend on the chart, and is a section of $P_U[T_1U, Ad]$ with support the integral curve. We call it the fundamental state of the boson (equivalent to ψ_0). Of course its measure changes with the gauge.

Representation of a boson

And because the vector V is part of the definition of the boson, we represent a boson by the tensor :

$$B = B_0 \otimes V \in T_1U \otimes TM \tag{8.4}$$

A boson is represented like a current corresponding to a particle.

And similarly for the gravitons :

$$\Gamma = \Gamma_0 \otimes V \in T_1U \otimes TM \tag{8.5}$$

8.1.2 Properties of bosons

Energy of a boson

There is always a field of the same type as the boson, and the propagation of bosons is part of the propagation of the fields, that is of the interaction of the field with itself. So bosons interact with the underlying force field.

The main feature of the propagation is the existence of a function F such that :

$\langle \mathcal{F}, \mathcal{F} \rangle (m) = \mathcal{E}_0 - F(m)$

We have : $\Delta\mathcal{F}_A(V) = -\mathbf{p}^*(m)\,\Delta\grave{\mathbf{A}}(m)$ so the propagation of the boson will manifest itself as a discontinuity in $F(m)$, that is of the energy of the field. And we can define the energy of the boson by :

$$\mathcal{E}_B = \langle \mathcal{F}_A(V), \Delta\mathcal{F}_A(V) \rangle \tag{8.6}$$

$\mathcal{F}_A(V)$,$\Delta\mathcal{F}_A(V)$ are one form, valued in the adjoint bundle, and the scalar product is preserved in a change of gauge, so the energy does not depend of the gauge. And similarly in a change of chart : $\mathcal{E}_B$ is a scalar (in the unit used) .

Momentum of a boson

To any vector field V can be associated a section of the adjoint bundle $P_G[T_1Spin(3,1), \mathbf{Ad}]$ (see Theorem 79) :

$\upsilon(0,W) \in T_1Spin(3,1)$ such that $V = c(\varepsilon_0(m)) + \epsilon \sum_{a=1}^{3} W_a \varepsilon_a(m)$ where $\epsilon = +1$ if W is in the same direction as the spatial speed, and -1 if not, and : $[\upsilon(r,W), \varepsilon_0] = \upsilon(0,W)$ with $W^tW = 1$ if the field propagates at c.

So a boson can equivalently be represented by $B_0 \otimes \upsilon(0,W) \in T_1U \otimes T_1Spin(3,1)$ which is the momentum of the boson.

This representation assumes the choice of an observer ε_0 : this is the momentum with respect to an observer.

Bosons interact with particles

The interaction of a field and a particle is represented through the covariant derivative :

$[\nabla_\alpha \psi] = [\partial_\alpha \psi] + [\gamma C (G_\alpha)] [\psi] + [\psi] \left[\grave{A}_a\right]$

$G_\alpha \in T_1 Spin(3,1), \grave{A}_a \in T_1 U$

so we can define an interaction between a boson and a particle. However we must account for the fact that this interaction can occur only if the trajectories of the particle and of the boson cross : it occurs only at a point, and there is no continuous effect. The interaction results in a discontinuity of $[\nabla_\alpha \psi]$. which can be represented not by the derivative $\partial_\alpha \psi$ but by the momentum. The particle will not stay on the same world line, the interaction with the boson acknowledges this fact by the action of the one form $\mathcal{M}$ on B. And we must account for the discrepancy between the directions of the boson and of the particle.

Thus this the value $\mathcal{M}(V_p)$ of the momentum for the trajectory V_p which is impacted.

The momentum of the particle has the discontinuity :

$\delta \mathcal{M}(V_p) = ([\gamma C (\upsilon(0,W))] [\psi_0] + [\psi_0] [\varrho'(1)(B_0)])$

If only the EM field and gravity are present :

$\delta \mathcal{M}(V_p) = [\gamma C (\upsilon(0,W))] [S_0] + iq (B_0) [\gamma C (\sigma)] [S_0]$

Any boson has an impact on any particle, through its kinematic part $\upsilon(0,W)$: the trajectory of the particle changes.

A boson can interact with a charged particle of the same type. For photons this is the Compton effect.

The fundamental state ψ_0 does not change, but (σ, g) and then the world line change, and we have after the interaction :

$\vartheta\left(\sigma'^{-1}, g'^{-1}\right) \frac{d\psi'}{dt} = \mathcal{M}'\left(V_p'\right) = \mathcal{M}(V_p) + \delta \mathcal{M}(V)$

and because $\sigma_{In} + \delta\sigma$ does not usually belong to $Spin(3,1)$ the variation of (σ, g) is not necessarily continuous. For a spinor :

$S_{In} = \gamma C (\sigma_{In}) \psi_0 \to S_{Out} = \gamma C (\sigma_{Out}) \psi_0$

$\sigma_w = \sqrt{1 + \frac{1}{4} w^t w} + \upsilon(0,w) \to \sigma'_w = \sqrt{1 + \frac{1}{4}(w + \delta w)^t (w + \delta w)} + \upsilon(0, w + \delta w)$

In the scheme above we have assumed that the boson disappears after the transition. But a collision between a particle and a boson can have different outcomes. The boson can disappear, it can stay, in a modified state, or even the particle itself can change. After the interaction the system comes back to an equilibrium. The stationary states of particles are quantized, so the out state can take only some values, and the transition is not always possible. The rules in such transitions are : the conservation of the global momentum, the conservation of the energy, and the possible states of the particle.

Gravitons

However we have an issue for the graviton. Even if they have never be observed, it is useful to see how they could be represented in this picture. It would be by the tensorial product $T_1 Spin(3,1) \otimes T_1 Spin(3,1)$. To extend the structure of Lie algebra to the tensorial product of Lie algebras, the procedure goes first to the universal enveloping algebra (see Spinor fields). So a graviton can be represented as an element of the 2 graded universal enveloping algebra $U(T_1 Spin(3,1))$ of $T_1 Spin(3,1)$. Any representation of a group lifts as a representation of the universal algebra (Maths.1891), the map being $Ad \circ \imath$ where $\imath$ is the canonical injection $T_1 Spin(3,1) \to U(T_1 Spin(3,1))$:

The action is the product of the actions of Γ_0 and $\upsilon(0,w)$ (Maths.1891) :

$\delta\mathcal{M}(V_p) = [\gamma C(\Gamma_0)][\gamma C(\upsilon(0,w))][\psi_0] = [\gamma C(\Gamma_0 \cdot \upsilon(0,w))][\psi_0]$
but $\Gamma_0 \cdot \upsilon(0,w) \in T_1 Spin(3,1)$ iff :
$\Gamma_0 = \upsilon(r_0, w_0)$
$w^t w_0 = 0$
$w^t r_0 = 0$
$\Leftrightarrow w = \lambda j(r_0) w_0, \lambda \in \mathbb{R}$
then : $w^t w = 1 = \lambda^2 \left((r_0^t w_0)^2 - (r_0^t r_0)(w_0^t w_0) \right)$ which is impossible.
The only solution is that Γ is, up to a scalar K, the Casimir element of $U_2(T_1 Spin(3,1))$:
$\Omega = \sum_{i=4}^{6} (\kappa_i)^2 - \sum_{i=1}^{3} (\kappa_i)^2$

There would be a unique type of graviton. Then, because Ω is invariant by **Ad**, a graviton "looks the same" for any observer, and it acts by the scalar $\frac{3}{2}$ on the momentum : $\delta\mathcal{M}(V_p) = K\frac{3}{2}\mathcal{M}(V_p)$. The fact that gravitons have not been observed so far implies that K should be very small.

The Planck's law

Interaction of the EM field with solids

When an EM beam is directed to a solid body, several effects happen :

- a classic interaction between the charged particles and the field : the beam is redirected (this is the radar);
- a thermal effect : the beam transfers some momentum to the particles, increasing their kinetic energy and so the temperature of the solid;
- a pressure on the solid : due to the internal links between the particles part of the kinetic energy is transferred to the solid itself.

These two last effects are the consequence of the purely kinetic part of the photons, and they depend only on the quantity of energy of the beam.

- part of the kinetic energy absorbed by the particles is reemitted as radiance heat. The balance between emission and radiation depends on the frequency of the field. This is the black-body radiation.
- if the frequency is above a threshold, depending on the material, some electrons are ejected. Their kinetic energy depends on the frequency, but the number of electrons depend on the intensity of the beam. This is the photo-electric effect. The conversion of energy between the beam and the current collected is not 100 % : above the threshold there is only a probability, increasing with the intensity, that an electron is ejected.

The two last phenomena involve the charged part of the photon, and the frequency of the beam. They are at the core of the Planck's law. It is clear that they involve discontinuous processes : in a solid in equilibrium the particles are in a stationary state, which depends on discrete conditions. The transition from a state to another is a discontinuity, which manifests notably in their level of energy. So the exchange of energy with the field involves a discontinuity in the field itself, with respect to its propagation. The issue is to see how the frequency enters the picture.

Photons in the vacuum

Let us first look at the way one can acknowledge the existence of a photon in the vacuum. The propagation of a field is, by definition a continuous process, characterized by some function $F(m)$ related to the phase of the field. A boson manifests itself as a discontinuity in $F(m)$. In the Chapter on QM we have shown that a singularity in a variable measured on a sample of points is deemed detected if the signal to noise ratio $\frac{\rho_2(x)}{\rho(x)} > 1$. The quantity $\rho(x)$ is the *frequency* with which the value x is taken by the variable and $\rho_2(x)$ is the frequency with which the value x is taken by the discrepancy. If the variable is $\langle \mathcal{F}_A, \mathcal{F}_A \rangle$ then, at a given point x : $\langle \mathcal{F}_A, \mathcal{F}_A \rangle (\varphi_o(t,x)) = F(\varphi_o(t,x))$ and the discontinuity : :$\langle \mathcal{F}_A, \mathcal{F}_A \rangle (\varphi_o(t,x)) + \mathcal{E}_B$. For a monochromatic source $F(\varphi_o(t,x))$ varies with a frequency ν. Over a given period of time $\rho(F(m))$ is proportional to the frequency, and a photon is acknowledged only if $\mathcal{E}_B \geq \mathcal{E} \times \nu$. Any discrepancy $\mathcal{E}_B$ smaller than a threshold $\mathcal{E}\nu$ proportional to the frequency is discarded as belonging to the noise. And actually it is quite impossible to detect a photon with fields of large wave lengths. This is the basic interpretation of the Planck's law : $E = h\nu$.

Photons in interactions with particles

Now let us see the interaction of a photon and a charged particle. The classic action of a monochromatic field results in a vibration of the particles with a frequency ν equal or close to the frequency of the field. The change of the kinetic energy of the particle is proportional to :

$K = \frac{1}{2} M_p c^2 k_0^t \left([C(r)]^t \left([D(r)] \frac{dr}{dt} + \frac{1}{2} j(w) \frac{dw}{dt} \right) \right) = K_0 \nu$

Above a definite frequency the kinetic energy is greater than the link to the atom, and the electron is freed. To balance the energy transferred to the electron the field must adjust : there is a discontinuity of the same amount, which is thus proportional to ν.

The Planck's law, in spite of its ubiquity, holds only for photons. The interaction between other bosons and particle is modelled in the path integral formalism by a lagrangian similar to the one above.

Mass, Spin and charge of a boson

Mass

The mass at rest is defined, quite conventionally, for particles by : $M_p = \frac{1}{c}\sqrt{\epsilon \langle \psi_0, \psi_0 \rangle}$, that is from the fundamental state.

Similarly, for a boson :

$B = B_0 \otimes V \in T_1 U \otimes TM \rightarrow \langle B, B \rangle = \langle B_0, B_0 \rangle_{T_1 U} \langle V, V \rangle_{TM}$

so a boson associated to a field propagating at the speed of light should have a null mas, as can be expected. And conversely the fields of weak interactions should propagate at a speed $< c$ This is generally assumed to be linked to a short range.

Spin

As for particles, bosons have a relativist spin : if W is in the direction of the trajectory or the opposite, which takes the value +1 or -1.

The spin of a spinor is defined through the decomposition of $\sigma = \sigma_r \cdot \sigma_w$ which has no meaning for a boson. However one of the key property of the spin is its transformation law under $Spin(3)$: particles with an integrer spin shows a physical symmetry because their spin takes the same value under the action of $+s$ or $-s$ for $s \in Spin(3)$, meanwhile particles with a half integer spin take the opposite value. From this point of view the existence of such a symmetry can be checked for bosons : $\upsilon(0, W)$, computed from V transforms as :

$\mathbf{Ad}_s \upsilon(0, W) = \upsilon(0, [C(r)] W)$ with $[C(r)] = \left[1 + a_r j(r) + \frac{1}{2} j(r) j(r)\right]$

and $s = a_r + \upsilon(0, r), -s = -a_r + \upsilon(0, -r)$ are represented by the same matrix $[C]$ so $\pm s$ have the same action on W. And we can conclude that bosons have a spin 1.

Charge

The charge of a particle is, except for the EM field, defined by comparison with particles which have the same behavior with respect to the fields. So bosons have a charge with respect to the field from which they are issued and, because $B \in T_1U$, for each of them we as many "charges" as the dimension of U (8 gluons for $SU(3)$, 3 bosons for $SU(2)$ and 4 for $SU(2) \times U(1)$).

When they are considered alone all photons have the same behavior with the EM field, so they bear no charge.

Similarly there could be only one type of graviton.

8.2 DISCONTINUOUS MODELS

8.2.1 Discontinuous models in the classic theory of fields

There are many models to represent discontinuous behavior of particles. The oldest are the kinetic models. They usually derive from a hydrodynamic model (similar to the continuous models), but add a "collision operator" to represent elastic collisions between particles : because there is no loss of energy, the pure spatial momenta are conserved, which brings a fundamental relation between the variables before and after a collision. The models are based upon a distribution function $f(m,p)$ of particles of linear momentum p which shall follow a conservation law, using the collision operator. So the distribution of charges is itself given by a specific equation. Then the 4 dimensional action, with a lagrangian adapted to the fields considered, gives an equation relating the field and the distribution of charges. Usually the particles are assumed to have the same physical characteristics (mass and charge), which imposes an additional condition on the linear momentum : $\langle p,p\rangle = mc^2$. The frequency of collisions is related to a thermodynamic variable similar to temperature. Such models have been extensively studied with gravitational fields only (Boltzman systems), notably in Astrophysics, and the electromagnetic field for plasmas (Vlasov-Maxwell systems).

8.2.2 Quantum Theory of fields

There are several, complementary, methods to deal with discontinuous processes at the atomic scale. We will just give an overview of these methods and how they can be consistent with our picture.

Scattering

Whenever discontinuities in the fields are considered, bosons are involved, and they interact with particles as seen above. The basic rules are the conservation of energy and momentum, pointwise. The result should not depend on the gauge, which enables to choose and compare several ones, adapted to each of the objects participating in the collision. Some additional rules apply, more empirical than based on strict principles, depending on the problem, such as the conservation of charge, the conservation of the sum of weak isospin or of the number of baryons. The CPT conservation provides also a useful guide in predicting the outcome. Moreover the strong interaction and electromagnetic interaction seem to be invariant under the combined CP operation,

A discontinuous process occurs between periods during which an equilibrium is assumed to have been reached. So one considers transitions from one state of equilibrium to another, represented as a **scattering** process in experiments using beams of particles, with a population of incoming particles in "in" states, and outgoing particles in "out" states. In an equilibrium the Principle of Least Action applies and continuous models can be used to compute the possible "in" and "out" states :

For the "in" particles :

$t \in]-\infty, 0[: \psi(t) = \vartheta(\sigma(t), \varkappa(t))\, \psi_{0in}$

and for the "out" particles :

$t \in]0, +\infty[: \psi(t) = \vartheta(\sigma(t), \varkappa(t))\, \psi_{out}$

The conservation of energy and momentum is met at each location, for all the particles or bosons involved. And the rules are implemented locally. If a photon strikes a particle, whose fundamental state does not change, the variation will be imparted to the boson (this is the

Compton effect). Conversely if the state of the particle changes, following a collision with another particle for instance, the new equilibrium is reached with the emission of a boson (Black Body radiation). There is always an underlying field, at least the one created by the particle.

Lagrangian

In our picture bosons can be represented as one form valued in the adjoint algebra, they interact with particles through the same mechanisms as the potentials, and moreover they transform, in a change of gauge, in a normal way. So one can consider to introduce them in a lagrangian, as additional variables which come along the classic covariant derivative. This is actually the lagrangian of the standard model, but in our picture there is no need for a Higgs field to preserve the equivariance. One can consider either "bosonic fields", similar to "ferminonic fields", or isolated bosons on their trajectories, with the method of functional derivatives there is no problem to treat together integrals with different measures (however there is a problem for the measure along null trajectories) and discontinuities.

Path integrals

In a discontinuous process usually there can be several possible outcomes. The question is then to find which one will occur. This is the main purpose of the path integral theory. As many others in Quantum Physics, its idea comes from Statistical Mechanics, and was proposed notably by Wiener.

If the evolution of the system meets the criteria of the Theorem 26 (the variables are maps depending on time and valued in a normed vector space and the process is determinist) there is an operator $\Theta(t)$ such that : $X(t) = \Theta(t) X(0)$. When in addition the variables $X(t)$ and $X(t+\theta)$ represents the same state, $\Theta(t) = \exp t\Theta$ with a constant operator. The exponential of an operator on a Banach space is a well known object in Mathematics (Maths.1025), so the law of evolution is simple when $\Theta(t)$ is constant, which requires fairly strong conditions. However, because discontinuities are isolated points, at least at an elementary level, between the transitions points Θ can be considered as constant. Then we have a succession of laws :

$t \in [t_p, t_{p+1}[: X(t) = (\exp t\Theta_p) X(t_p)$

and :

$X(t) = (\exp(t - t_p)\Theta_p)(\exp(t_p - t_{p-1})\Theta_{p-1}) ... (\exp t_1\Theta_p) X(0)$

which are usually represented, starting from the derivative.

This is a generalization of the mathematical method to express the solution of the differential equation in $\mathbb{R}^m$: $\frac{dX}{dt} = \Theta(t) X(t)$:

$X(t) = \lim_{n\to\infty} \left(\prod_{p=0}^{n} \exp(t_{p+1} - t_p)\Theta(t_p) \right) X(0)$ (Maths.2570)

The Θ_p and the intermediary transition points are not known, but if we can attribute a probability to each transition, then we have a stochastic process (see Maths.11.4.4). The usual assumption is that the transitions are independent events, and the increment $(\Theta_{p+1} - \Theta_p)$ follow a fixed normal distribution law (a Wiener process). In this scheme all possible paths must be considered.

In QM the starting point is the Schrödinger equation, $i\hbar\frac{d\psi}{dt} = H\psi$, which has a similar meaning. However in a conventional QM interpretation there is no definite path (only the initial and the final states are considered) and furthermore, because of the singular role given to t, it seemed not compatible with Relativity. Dirac proposed the use of the lagrangian, and Feynman provided a full theory of path integrals, which is still one of the essential tools of QTF. The fundamental ideas, as expressed by Feynman, are that :

- to any physical event is associated a complex scalar ϕ, called an amplitude of probability,
- a physical process is represented by a path, in which several events occur successively,
- the amplitude of probability of a process along a path is the sum of the amplitude of probability of each event,
- the probability of occurrence of a process is the square of the module of the sum of the amplitudes of probability along any path which starts and ends as the initial and final states of the process (at least if there is no observation of any intermediate event).

The amplitude of probability of a given process is given by : $e^{\frac{i}{\hbar}S[z]}$ where $S[z]$ is the action, computed with the lagrangian :

$S[z] = \int_A^B L\left(z^i, z^i_\alpha ... z^i_{\alpha_1...\alpha_r}\right) dm$ evaluated from the r-jet extension of z. The total amplitude of probability to go from a state A to a state B is $\phi = \int e^{\frac{i}{\hbar}S[z]} Dz$ where Dz means that all the imaginable processes must be considered. Then the probability to go from A to B is $|\phi|^2$. So each path contributes equally to the amplitude of probability, but the probability itself is the square of the module of the second integral.

The QM wave function follows : $\psi(x,t) = \int_{-\infty}^{+\infty} \phi(x,t;\xi,\tau)\,\psi(\xi,\tau)\,d\xi d\tau$

If a process can be divided as : $A \to B \to C$ then

$\phi(A,C) = \phi(A,B)\,\phi(B,C)$ which is actually the idea of dividing the path in small time intervals.

It can be shown that, in the classical limit ($\hbar \to 0$) and certain conditions, the path integral is equivalent to the Principle of Least Action. With simplifications most of the usual results of QM can be retrieved.

Even if the literature emphasizes simple examples (such as the trajectory of a single particle), the path integral is used, with many variants, mainly to address the case of discontinuous processes in QTF, as this is the only general method known. It leads then to consider the multiple possibilities of collisions, emissions,... involving different kinds of particles or bosons, in paths called Feynman's diagrams.

The quantities which are involved are either force fields (gravitation is not considered), fermionic fields or bosonic fields. In the latter two cases a trajectory is computed as a path, but this is a bit awkward because if bosons follow a null curve the lagrangian cannot be defined on the same measure (for particles we have $\sqrt{|\langle V,V\rangle|}dt$ .and for boson dt).

It is clear that this formalism is grounded in the philosophical point of view that all physical processes are discreet and random. One can subscribe or not to this vision, but it leads to some strange explanations. For instance all the paths must be considered, even when they involve unphysical behaviors for the particles (the virtual particle are not supposed to follow the usual laws of physics). An explanation which is not necessary : we have eventually a variational calculus, so r-jets, in which the derivatives are independent variables, are the natural mathematical framework and we must consider all possible values for the variables, independently of their formal relations.

Beyond the simplest case, where it has little added value, the computation of path integrals is a dreadful mathematical endeavour. This is done essentially in a perturbative approach, where the lagrangian is simplified as we have done previously, so as to come back to quadratic expressions. The results are then developed in series of some scale constant. However it is full of mathematical inconsistencies, such as divergent integrals. The theory of path integrals is then essentially dedicated to find new computational methods or tricks, without few or no physical justification : renormalization, ghosts fields, Gladston bosons, Wick's rotation, BRST,...

Interacting micro-systems

The continuous model of type 1 was inspired by Fluid Mechanics, and the natural extension is Gas Mechanics, where a great number of particles interact together. Fortunately the framework used here fits well with the idea of interacting microsystems.

The single particle model does not provide the value of the fields out of the trajectory. However it shows that the interactions force fields / particles, in spite of being represented by maps which have supports of different dimension (a 4 manifold for force fields and a curve for fermions or bosons) lead to relations in which only the values at one point m matter : $\psi(m), \widehat{\grave{A}}(m) = \sum_{\alpha=0}^{3} \grave{A}_\alpha V^\alpha, \widehat{G}(m) = \sum_{\alpha=0}^{3} G_\alpha V^\alpha$. All these variables are geometric, so actually the location m does not matter : their value can be expressed in any system of frames or coordinates. And we have the same features for the bosons $B = B_0 \otimes \upsilon(0, w), \Gamma$. Moreover the state of the particle or the boson incorporates all the information about their motion, which is then defined without the need to exhibit components of vectors in a frame.

We can then consider micro systems, comprising one particle, the fields and possibly one boson, all located at the same point, and represented by the variables which enter the lagrangian : $\psi, \mathcal{F}_A, \mathcal{F}_G, \widehat{\grave{A}}, \widehat{G}, B, \Gamma$ and P in the GR context, albeit it is a bit of an over kill in these kinds of problems, the SR approximation suffices [1]. These variables are seen as force, fermionic or bosonic fields but, because their location does not matter, this is the value in the associated vector spaces which is considered (which is equivalent to take the gauge of the observer). If the conditions of the theorem 28 are met, then the state of a microsystem is represented in a Hilbert space H by a vector, which is the direct product of vectors representing each variable.

If we have a system comprised of N such microsystems, and if the bosons and fermions are of the same type, they have the same behavior, and are indistinguishable : we have a homogeneous system and we can apply the theorems 31 and 33. The interactions between the micro systems lead to the quantization of the states. This is done in several steps.

1. The states of the microsystems (encompassing all the variables $\psi, \mathcal{F}_A, \mathcal{F}_G, \widehat{\grave{A}}, \widehat{G}, B, \Gamma$ are associated to a Hilbert space H , and the states of the system are associated to the tensorial product $\otimes_{n=1}^{N} H$ of the Hilbert space H associated to each microsystem. An equilibrium of the system corresponds to a vector subspace $\mathbf{h}$ of $\otimes_{n=1}^{N} H$ which is defined by :

i) a class of conjugacy $\mathfrak{S}(\lambda)$ of the group of permutations $\mathfrak{S}(N)$

ii) p distinct vectors $(\widetilde{\varepsilon}_j)_{j=1}^{p}$ of a Hermitian basis of H which together define a vector space H_J

And $\mathbf{h}$ is then either $\odot_{n_1} H_J \otimes \odot_{n_2} H_J ... \otimes \odot_{n_p} H_J$ or $\wedge_{n_1} H_J \otimes \wedge_{n_2} H_J ... \otimes \wedge_{n_p} H_J$

The state Ψ of the system is then : $\Psi = \sum_{(i_1 ... i_n)} \Psi^{i_1..i_n} \widetilde{\varepsilon}_{i_1} \otimes .. \otimes \widetilde{\varepsilon}_{i_n}$ with an antisymmetric or a symmetric tensor.

2. We have global variables, which can be taken equivalently as the number of particles, or their charge, and the energy of the system. For each value of the global variables the state Ψ of the system belongs to one of the irreducible representations. The class of conjugacy λ and the vectors $(\widetilde{\varepsilon}_j)_{j=1}^{p}$ are fixed.

3. At the level of each microsystem, each vector $\widetilde{\varepsilon}_j \in H$ represents a definite state of a micro system, and the value of each variable of the micro-system is quantized. In a probabilist interpretation one can say that there are $(n_i)_{i=1}^{p}$ microsystems in the state $\widetilde{\varepsilon}_{j_i}$.But one cannot say with certainty what is the state of a given microsystem.

[1] Notice that P is a vector (ε_i) so a geometric quantity, and is not submitted a priori to a constraint because g is defined from P.

The quantization of each microsystem means that the vector ϕ representing its state in H belongs to a finite dimensional vector space :

$\phi = \sum_{(i_1,...,i_q)} \phi^{i_1...i_q} |e_{i_1}, e_{i_2}...e_{i_q} >$ where the vectors $|e_{i_1}, e_{i_2}...e_{i_q} >$ correspond to the $\widetilde{\varepsilon}_j$

The spin of the particle corresponds to one of the vectors e_j of the basis. The spin of a particle is always invariant by $Spin(3)$, the action of s and $-s$ give opposite results. If the spin number j is an integer then the particle has a specific, physical symmetry, and its spin is invariant by $SO(3)$. This property must be reflected in the states of the system.

If j is half an integer the representation of the system is by antisymmetric tensors to account for the antisymmetry by $Spin(3)$. As a consequence in each vector space $\Lambda_n H_J$ the components of the tensors, expressed in any basis, which correspond to the diagonal are null :

$\psi^{i_1..i_n} = 0$ for $i_1 = i_2 = .. = i_n$

The micro systems belonging to the same $\Lambda_n H_J$ must be in different states. This the Pauli's exclusion principle.

The particles whose spin number is half an integer are called fermions and are said to follow the Fermi-Dirac statistic.

The particles whose spin number is an integer are called bosons and are said to follow the Boose-Einstein statistic.

So the denominations fermions / bosons are here different from that we have used so far. All elementary particles are fermions, all discontinuities of the fields are bosons, but composite particles or atoms can be bosons if the spin number is an integer.

The exclusion principle does not apply to all the micro-systems. In a system there are usually different sets of microsystems, which corresponds to different subspaces $\Lambda_n H_J$ and therefore micro-systems belonging to different subspaces can have the same spin, however each of these subspaces is distinguished by other global variables, such the energy (for instance the electrons are organized in bands of valence in an atom).

Fock Spaces

In the model of interacting systems above, the number of particles is assumed to be constant. One removes this limitation by an extension of the tensorial product of Hilbert spaces. The space of representation of the system is then the tensorial algebra, called a Fock space : $\mathcal{F} = \oplus_{k=0}^{\infty} (\otimes_k H)$ where H stands for the Hilbert space of the microsystems (Maths.1209). k can be 0 so scalars can be vectors of the Fock spaces.

A vector Ψ of $\mathcal{F}_n = \oplus_{k=0}^{n} (\otimes_k H)$ is given by $n+1$ tensors :

$(\psi^m, \psi^m \in \otimes^m H, m = 0...n)$

The "ground state" is the vector $(1, 0, 0, ...)$ in the algebra.

Any operator on the Hilbert spaces can be extended to a linear continuous operator on the Fock space.

For each Fock space $\oplus_{k=1}^{\infty} (\otimes_k H)$ there is a number operator N, whose, dense, domain is :

$D(N) = \left\{ \psi^m \in \otimes_m H, \sum_{k \geq 0} m^2 \|\psi^m\|^2 < \infty \right\}$

$N(\Psi) = \left(0, \psi^1, 2\psi^2, ...m\psi^m...\right)$

N is self adjoint.

The annihilation operator cuts a tensor at its beginning :

$a_m : H \to \mathcal{L}(\otimes_m H; \otimes_{m-1} H) ::$

$a_m(\psi)(\psi_1 \otimes \psi_2 ... \otimes \psi_m) = \frac{1}{\sqrt{m}} \langle \psi, \psi_1 \rangle_H \psi_2 \otimes \psi_3 ... \otimes \psi_m$

The creation operator adds a vector to a tensor at its beginning :

$a_m^* : H \to \mathcal{L}(\otimes_m H; \otimes_{m+1} H) ::$

$a_m^*(\psi)(\psi_1 \otimes \psi_2 ... \otimes \psi_m) = \sqrt{m+1} \psi \otimes \psi_1 \otimes \psi_2 \otimes \psi_3 ... \otimes \psi_m$

a_m^* is the adjoint of a_m and a_m, a_m^* can be extended to the Fock space as a, a^*.

The physical meaning of these operators is clear from their names. They are the main tools to represent the variation of the number of particles.

The spaces of symmetric (called the Bose-Fock space) and antisymmetric (called the Fermi-Fock space) tensors in a Fock space have special properties. They are closed vector subspaces, so are themselves Hilbert spaces, with an adjusted scalar product. Any tensor of the Fock space can be projected on the Bose subspace (by P_+) or the Fermi space (by P_-) by symmetrization and antisymmetrization respectively, and P_+, P_- are orthogonal. The operator $\exp itN$ leaves both subspaces invariant. Any self-adjoint operator on the underlying Hilbert space has an essentially self adjoint prolongation on these subspaces (called its second quantification). However the creation and annihilation operators have extensions with specific commutation rules :

Canonical commutation rules (CCR) in the Bose space:

$[a_+(u), a_+(v)] = [a_+^*(u), a_+^*(v)] = 0$

$[a_+(u), a_+^*(v)] = \langle u, v\rangle 1$

Canonical anticommutation rules (CAR) in the Fermi space :

$\{a_+(u), a_+(v)\} = \{a_-^*(u), a_-^*(v)\} = 0$

$\{a_+(u), a_+^*(v)\} = \langle u, v\rangle 1$

where

$[X, Y] = X \circ Y - Y \circ X$

$\{X, Y\} = X \circ Y + Y \circ X$

These differences have important mathematical consequences. In the Fermi space the operators a_-, a_-^* have bounded (continuous) extensions. Any configuration of particles can be generated by the product of creation operators acting on the ground state. We have nothing equivalent for the bosons.

We will not pursue on these topics, which are exposed in many books, and requires more mathematical concepts.

Chapter 9

CONCLUSION

At the end of this book I hope that the reader has a better understanding of how Mathematical tools such as Group Representations, Clifford Algebras, Fiber Bundles, Connections can be used in Physics. They can be a bit abstract, but actually they are well suited, and quite efficient, to address the issues of modern Physics. I hope also to have brought some clarification on Quantum Mechanics, Relativity and gauge theories, as well as on concepts, such as the duality between particles and fields.

1. In the Second Chapter it has been proven that most of the axioms of QM come from the way models are expressed in Physics, and the following chapters have shown how the theorems can be used. The theorems state precise guidelines and requirements for their validity, and these requirements, albeit expressed as Mathematical conditions, lead to a deeper investigation of the physical meaning of the quantities which are used. For a property, the fact to be geometric is not a simple formality : it means that this is an entity which exists beyond the measures which can be made, and that these measures vary according to precise rules. The role of the observer in the process of measurement is clearly specified. The condition about Fréchet space, which seemed strange, takes all its importance in the need to look for norms on vector spaces. The relation between observables and statistical procedures has found a nice application to explain the Plank's law. There has been few examples of the use of observables, whose role is more central in models representing practical experiments, but their meaning should be clear.

2. Relativity, and particularly General Relativity, which is often seen as a difficult topic, can be understood if we accept to start from the beginning, from Geometry, the particularities of our Universe and if we accept to give up schemes and representations which have become too familiar, such as inertial frames. With the formalism of fiber bundles it is then easy to address very general topics without losing the mathematical rigor. We have given a consistent and operational definition of a deformable solid, which can be important in Astrophysics.

3. Clifford algebras are not new, but they appear really useful when one accepts fully the riches of their structure, without resorting to hybrid concepts such as quasi or axial vectors. With spinors the concept of matter fields becomes clear. In my opinion they are the only way to represent in a consistent and efficient manner the motion and the kinematics properties of material bodies in the GR context. So Spinors should be useful in Astrophysics, where gravitation is the only force involved and GR cannot be dismissed.

4. The use of connections to represent the force fields has become a standard in gauge theories. The strict usage of fiber bundles and spinors enables to put the gravitational field in the same framework, and it appears clearly that the traditional method based on the metric and the Levi-Civita connection imposes useless complications and miss some features which can be physically

important, such that the decomposition in transversal and spatial components. Propagation of force fields is a widely used concept, but to which too little theoretical work has been devoted. The model which has been introduced is based on the concept itself, and the assumption that fields are physical entities. It leads to important results which are consistent with classic models, and are the starting point to a representation of bosons.

5. The two models presented were essentially an example of how the theory of Lagrangians can be used practically. And we have given a strong mathematical backing to the functional derivatives calculus, based of an original theory of distributions on vector bundles. They are the starting point for the concepts of currents and energy-momentum tensor. Important theorems have been proven, and we have provided guidelines which can be used to find operational solutions of the difficult problem of the gravitational field.

6. The idea of bosons as discontinuities in the fields is more speculative. But it seemed necessary to complement the concept of matter fields, which is quite clear, with an interpretation of the corpuscular nature of the fields. The presentation leaves some gaps, which are due to our limited knowledge of the propagation of weak and strong interactions, beyond the Higgs mechanism, and the non observation of gravitons.

There are some new results in this book : QM, deformable solids in RG, spinors, motion of material bodies with the gravitational field, bosons. They are worth to be extended, by filling the gaps, or simply using the methods which have been introduced. For instance many other theorems could be proven in QM, a true Mechanics of deformable solids could be built, with the addition of Thermodynamics concepts, the representation of bosons could be more firmly grounded by the consideration of the known properties of all bosons. But from my point of view the one which is worth of the most efforts is gravitation. This is the most common and weakest of all force fields, but we are still unable to use it or to understand it properly. The representation of the gravitational field by connections on one hand, and of the gravitational charges by spinors on the other hand, shows striking similarities with the EM field : indeed they are the only fields which have an infinite range, the EM charge can be incorporated in the gravitational charge, and their propagation equations are similar. This similitude has been remarked by many authors, Heaviside, Negut, Jefimenko, Tajmar, de Matos,...and it has been developed in a full Theory, which has sometimes be opposed to GR. We find here that these similitudes exist in the frame of a GR theory which allows for a more general connection and the use of the Riemann tensor, so it seems more promising to explore this avenue than to fight against GR. The gravitational field shows, in all its aspects, two components. The "magnetic" component can be assimilated to the usual gravity : this is the one which acts in the 3 dimensional space. The "electric" component acts in the time dimension, and it seems logical to give it a cosmological interpretation : it would be the engine which moves matter on its world line. Both components have opposite effects, and there is no compelling reason that it should always be attractive.

This new look on the relation between the gravitational and the EM fields leads also to reconsider the "Great Unification". The Standard Model has not been the starting point to the unification of all force fields. It has brought the EM field with the weak and strong interactions with which it shares very few characteristics, meanwhile it has been unable to incorporate the gravitational field which seems close to the EM field, and all that at the price of the invention of a 5th force. For theoretical as well as practical purpose the right path seems to consider the forces which manifest at long range together, and to find a more specific framework for the nuclear forces. This seems a strange conclusion for a book which puts the gauge theories at the front. But fiber bundles, connection and gauge theories have their place in Physics as efficient tools, not as the embodiment of a Physical Theory. The fact that they can be used at any scale, and for practical studies, should suffice to support their interest.

QM and Relativity have deeply transformed the way we do Physics.

We were used to an eternal, flat, infinite Universe (an idea which is, after all, not so obvious). With Relativity we had to accept that we could represent the Universe as a four dimensional, curved, structure, which integrates the time. Beyond the change of mathematical formalism, Relativity has also put limits to our capability to know the Universe. We are allowed to model it as we want, with an infinite extension, in space and time, but the only Universe that is accessible to our measures and experiments is specific to each observer : we have as many windows on the Reality that there are observers. We can dream the whole world, we can put in our models variables which are related to the past or the future, as if they were there, but the world that I can perceive is the world that I see from my window, and my neighbor uses another window. I can imagine what is beyond my window, but to get a comprehensive picture I need to patch together different visions.

With QM we have realized that we can model the reality, whatever the scale, with mathematical objects, but these objects exist only in the abstract world of Mathematics, they are some idealization that we use because they are efficient in our computations, but we can access reality only with cruder objects, finite samples and statistic estimations. The discrepancy between the measures, necessarily circumstantial and probabilist, and the real world does not mean that the real world is discreet and proceeds according to random behaviors, only that we have to acknowledge the difference between a representation and the reality. And conversely it does not preclude the use of the models, as long as we are aware of their specific place : it is not because we cannot measure simultaneously location and speed that their concepts are void.

The Copenhagen interpretation of QM states the existence of 2 Physics, one which holds at the atomic level, and another at our scale. Actually the way we can use Mathematics to represent and model the physical world leads to distinguish continuous and discontinuous processes. The distinction holds at any scale, but the scale also matters, because discontinuous processes can be simplified and represented as continuous, if we accept to neglect part of the phenomena.

Contrary to many, I am a realist, I believe that there is a unique real world outside, it can be understood, it is not ruled by strange and erratic behaviors. But modern Physics, in a mischievous turn, has imposed the need to reintroduce the individual in Science, in the guise of the observer, and the discrepancy between imagination, which enables us to see the whole as if it was there, and the limited possibility to keep it in check. The genuine feature of the human brain is that it can conceive of things that do not exist, that will never occur as we dreamed them. This is precious and Science would be impossible without it. To impart to reality our limitations or to limit our ambitions to what we can check are equally wrong. Actually the only way for a Scientist to keep his sanity in front of all the possible explanations which are provided is that to remember that there is one world : the one in which he lives.

jc.dutailly@free.fr

Chapter 10

BIBLIOGRAPHY

R.D.Anderson *Some open questions in infinite dimensional topology* Proceeding of the 3d Prague symposium Praha (1972)

H.Andréasson *The Einstein-Vlasov systems / Kinetic theory* Living Review in Relativity 14 (2011),4

H.Araki *Mathematical theory of quantum fields* Oxford Science Publications (2000)

N.Ashby *Relativity in the global positioning system* Living Review in Relativity 6,(2003),1

J.C. Baez, M.Stay *Physics, Topology, Logic and Computation: A Rosetta Stone* arXiv 0903.0340 (2009)

H.Baumgartel *Operator Algebraic methods in Quantum Fields* Berlin Akademie verl. (1995)

M.Le Bellac *Physique quantique* CNRS (2003)

A.Berlinet, C.Thomas-Agnan *Reproducing kernel, Hilbert spaces in probability and statistics* Springer (2004)

A.Bird *Philosophy of science* Rootledge (1998)

J.D.Bjorken,S.D.Drell *Relativistic quantum fields* Mc Graw Hill (1965)

N.Bogolubov,A.A.Logunov, A.I.Ossak, I.T.Todorov *General principles of quantum fields theory* Kluwer (1990)

O.Bratelli, D.W.Robinson *Operators algebras and quantum statistical mechanics* Springer (2002)

Y.Choquet-Buhat, N.Noutcheguem *Système de Yang-Mills Vlasov en jauge temporelle* Annales de l'IHP section A tome 55 N°3 (1991)

B.Coecke, E.O.Paquette *Categories for the practising physicist* arXiv:0905-3010v1 [quant-ph] (16 may 2009)

B.Coecke *New Structures for Physics* Lecture Notes in Physics vol. 813, Springer, Berlin, 2011

R.Coquereaux *Clifford algebra, spinors and fundamental interactions : twenty years after* arXiv:math-ph/0509040v1 (16 sep 2005)

B.d'Espagnat *Reality and the physicist* Cambridge University Press (1989)

P.A.M.Dirac *The principles of Quantum Mechanics* Oxford Science Publications (1958)

J.C.Dutailly *Mathematics for theoretical physics* arXiv:1209-5665v2 [math-ph] (4 feb 2014)

J.C.Dutailly *Estimation of the probability of transition between phases* CNRS (http://hal.archives-ouvertes.fr/hal-01075940, 20 october 2014)

R.P.Feynman, A.R.Hibbs *Quantum Mechanics and Path Integrals* Dover (2005)

J.Finne *Asymptotic study of canonical correlation analysis: from matrix and analytic approach to operator and tensor approach* SORT 27 (2) July-December 2003, 165-174

Francis C.E.H *A construction of full QED using finite dimensional Hilbert space* EJTP 10 N°28 (2013)

Francis C.E.H *The Hilbert space of conditional clauses* arXiv:1205-4607 (2013)

G.Giachetta,L.Mangiarotti, G.Sardanashvily *Advanced classical field theory* World Scientific (2009)

Tepper L.Gill, G.R.Pantsulaia, W.W.Zachary *Constructive analysis in infinitely many variables* arXiv 1206-1764v2 [math-FA] (26 june 2012)

M.Guidry *Gauge fields theories* Wiley (1991)

H.Halvorson *Algebraic quantum fields theory* arXiv:math-ph/0602036v1 14 feb 2006

N.R. Hansen *Non-parametric likelihood based estimation of linear filters for point processes* arXiv:1304-0503v3 [stat:CO] (12 feb 2014)

D.W.Henderson *Infinite dimensional manifolds are open subsets of Hilbert spaces* (1969) Internet paper

S.Hetherington *Epistemology : the key thinkers* Continuum (2012)

S.S.Horuzhy *Introduction to algebraic quantum field theory* Riedel (1989)

J.M.Jauch *Foundations of Quantum Mechanics* AddisonWesley (1968)

E.T.Jaynes *Where do we stand on maximum entropy ?* Paper for a MIT conference (1978)

O.Jefimenko *Causality, Electromagnetic Induction, and Gravitation* Electret Scientific, Star City, (2000)

Sir M.Kendall, A.Stuart *The advanced theory of statistics* Charles Griffin & Co (1977)

S.C.Kleene *Mathematical logic* Wiley & Sons (1967)

A.W.Knapp *Lie groups beyond an introduction* Birkhäuser (2005)

I.Kolar, P.Michor, J.Slovak *Natural operations in differential geometry* Springer-Verlag (1991)

S.Kopeikin, B.Mashloon *Gravitomagnetic effects in the propagation of electromagnetic waves in variable gravitational fields of arbitrary moving and spinning bodies* arXiv / gr-qc / 0110101 v2 (25 oct 2001)

F.Laloë *Comprenons-nous vraiment la mécanique quantique ?* CNRS Editions (2011)

E.H.Lieb, M.Loss *Analysis* American Mathematical Society (2000)

A.Garrett Lisi *An explicit embedding of gravity and the standard model* arXiv: gr-qc 1006498v1 (25 june 2010)

A.Garrett Lisi *An exceptionnaly simple theory of everything* arXiv:0711.0770v1 [hep-th] (6 nov 2007)

A.Garrett Lisi *Clifford bundle formulation of BF gravity generalized to the standard model* arXiv:gr-qc 0511120v2 (21 nov 2005)

R.Haag *Local quantum physics* 2nd Ed.Springer (1991)

S.Mac Lane *Categories for the working mathematician* Spinger (1997)

G.Mackey *The mathematical fundations of Quantum Mechanics* W.A.Benjamin (1963)

M.A.Naimark *Linear representations of the Lorentz group* Pergamon (1964)

E.Negut *On intrinsic properties of relativistic motions* Revue Roumaine des Sciences Techniques (1990).

J.von Neumann *Mathematical Foundations of Quantum Mechanics*, Beyer, R. T., trans., Princeton Univ. Press. (1996 edition)

L.Nielsen *A Maxwell Analog Gravitation Theory* Niels Bohr Institute, Copenhagen, Gamma No. 9 (1972).

Y.Ohnuki *Unitary representations of the Poincaré group and relativistic wave equations* World Scientific (1988)

R.Omnès *The interpretation of quantum mechanics* Princeton (1994)

J. E. Palomar Tarancon *Conceptual systems, conceptual convergence theory and and algebras of analogies*

P.J.E.Peebles *Principles of physical cosmology* Princeton series in Physics (1993)

H.Poincaré *La valeur de la science* Flammarion (rep.2014)

H.Poincaré *La Science et l'hypothèse* Flammarion (rep.2014)

E.Poisson *An introduction to the Lorentz-Dirac equation* arXiv:gr-qc/ 9912045v1 (10 dec 1999)

S.Pokorski *Gauge fields theories* Cambridge (2000)

K.Popper *Quantum theory and the schism in physics* Routledge (1982)

K.Popper *The logic of scientific discovery* Rootledge (1959)

T.C.Quinn *Axiomatic approach to the radiation reaction of scalar point particles in curved space-time* arXiv:gr-qc/0005030v1/ (10 may 2000)

G.Röpke *Non equilibrium statistical physics* Wiley-vch (2013)

I.Schnaid *Wave function perturbations propagation in multi particles system* arXiv 1307.2510v1 [physics-gen.ph] (9 july 2013)

T.Schücker *Forces from Conne's geometry* arXiv:hep/th/0111236v3 (26 march 2007)

Mesgun Sehbatu *The standard model and beyond* (1992)

A.Smola, A.Gretton, L.Song, B.Schölkop *A Hilbert Space Embedding for Distributions*

M.Tajmar, C.J. de Matos *Coupling of Gravitation and Electromagnetism in the Weak Field Approximation* arXiv:gr-qc/0003011 (2003)

H.Torunczyk *Characterizing Hilbert spaces topology* Fundamental mathematica (1981)

J.Turri *Epistemology : a guide* Wiley (2014)

A. Vourdas *The complete Heyting algebra of subsystems and contextuality* arXiv:1310-3604v1 [quant-ph] (14 oct 2013)

R.M.Wald *General Relativity* Univ.Chicago (1984)

S.Weinberg *The quantum theory of fields* Cambridge University Press (1995)

S.Weinberg *Dreams of a Final Theory* Pantheon Books (1992)

H.Weyl *The theory of groups and quantum mechanics* Dover (1931 / 1950)

Appendix A

ANNEX

A.1 CLIFFORD ALGEBRAS

This annex gives proofs of some results presented in the core of the paper.

A.1.1 Products in the Clifford algebra

Many results are consequences of the computation of products in the Clifford algebra. The computations are straightforward but the results precious. In the following $\langle \varepsilon_0, \varepsilon_0 \rangle = -1$ with the signature (3,1) and +1 with the signature (1,3). The operator j is reminded in the Formulas at the end of this Annex.

Product $\upsilon(r,w) \cdot \upsilon(r',w')$

$\upsilon(r,w) = \frac{1}{2}\left(w^1\varepsilon_0 \cdot \varepsilon_1 + w^2\varepsilon_0 \cdot \varepsilon_2 + w^3\varepsilon_0 \cdot \varepsilon_3 + r^3\varepsilon_2 \cdot \varepsilon_1 + r^2\varepsilon_1 \cdot \varepsilon_3 + r^1\varepsilon_3 \cdot \varepsilon_2\right)$

$\upsilon(r',w') = \frac{1}{2}\left(w'^1\varepsilon_0 \cdot \varepsilon_1 + w'^2\varepsilon_0 \cdot \varepsilon_2 + w'^3\varepsilon_0 \cdot \varepsilon_3 + r'^3\varepsilon_2 \cdot \varepsilon_1 + r'^2\varepsilon_1 \cdot \varepsilon_3 + r'^1\varepsilon_3 \cdot \varepsilon_2\right)$

With signature (3,1) :

$\upsilon(r,w) \cdot \upsilon(r',w') = \frac{1}{4}\left(w^t w' - r^t r'\right) + \frac{1}{2}\upsilon\left(j(r)r' - j(w)w', j(w)r' + j(r)w'\right) - \frac{1}{4}\left(w^t r' + r^t w'\right)\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$

From there the bracket on the Lie algebra :

$[\upsilon(r,w), \upsilon(r',w')] = \upsilon(r,w) \cdot \upsilon(r',w') - \upsilon(r',w') \cdot \upsilon(r,w)$

$$[\upsilon(r,w), \upsilon(r',w')] = \upsilon\left(j(r)r' - j(w)w', j(w)r' + j(r)w'\right) \quad (A.1)$$

With signature (1,3) :

$\upsilon(r,w) \cdot \upsilon(r',w') = \frac{1}{4}\left(w^t w' - r^t r'\right) - \frac{1}{2}\upsilon\left(-j(r)r' + j(w)w', j(w)r' + j(r)w'\right) - \frac{1}{4}\left(w^t r' + r^t w'\right)\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$

From there the bracket on the Lie algebra :

$$[\upsilon(r,w), \upsilon(r',w')] = -\upsilon\left(j(r)r' - j(w)w', j(w)r' + j(r)w'\right) \quad (A.2)$$

In both signatures the basis of the Lie algebra is denoted :

$\overrightarrow{\kappa}_1 = \frac{1}{2}\varepsilon_3 \cdot \varepsilon_2,$
$\overrightarrow{\kappa}_2 = \frac{1}{2}\varepsilon_1 \cdot \varepsilon_3,$
$\overrightarrow{\kappa}_3 = \frac{1}{2}\varepsilon_2 \cdot \varepsilon_1,$
$\overrightarrow{\kappa}_4 = \frac{1}{2}\varepsilon_0 \cdot \varepsilon_1,$
$\overrightarrow{\kappa}_5 = \frac{1}{2}\varepsilon_0 \cdot \varepsilon_2,$
$\overrightarrow{\kappa}_6 = \frac{1}{2}\varepsilon_0 \cdot \varepsilon_3$
$a, b, c = 1, 2, 3$
$[\overrightarrow{\kappa}_a, \overrightarrow{\kappa}_b] = \epsilon(c, a, b)\, \overrightarrow{\kappa}_c$
$[\overrightarrow{\kappa}_a, \overrightarrow{\kappa}_{3+b}] = \epsilon(c, a, b)\, \overrightarrow{\kappa}_{3+c}$
$[\overrightarrow{\kappa}_{3+a}, \overrightarrow{\kappa}_{3+b}] = \epsilon(c, a, b)\, \overrightarrow{\kappa}_{3+c}$
More over : $\upsilon(x, y) \cdot \varepsilon_5 = \varepsilon_5 \cdot \upsilon(x, y) = \upsilon(-y, x)$
In $Cl(3, 1)$:
$\upsilon(x, y) = \frac{1}{2}(y_1\varepsilon_0 \cdot \varepsilon_1 + y_2\varepsilon_0 \cdot \varepsilon_2 + y_3\varepsilon_0 \cdot \varepsilon_3 + x_3\varepsilon_2 \cdot \varepsilon_1 + x_2\varepsilon_1 \cdot \varepsilon_3 + x_1\varepsilon_3 \cdot \varepsilon_2)$
$\frac{1}{2}(y_1\varepsilon_0 \cdot \varepsilon_1 + y_2\varepsilon_0 \cdot \varepsilon_2 + y_3\varepsilon_0 \cdot \varepsilon_3 + x_3\varepsilon_2 \cdot \varepsilon_1 + x_2\varepsilon_1 \cdot \varepsilon_3 + x_1\varepsilon_3 \cdot \varepsilon_2) \cdot \varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$
$= \frac{1}{2}(-y_1\varepsilon_3 \cdot \varepsilon_2 - y_2\varepsilon_1 \cdot \varepsilon_3 - y_3\varepsilon_2 \cdot \varepsilon_1 + x_3\varepsilon_0 \cdot \varepsilon_3 + x_2\varepsilon_0 \cdot \varepsilon_3 + x_1\varepsilon_0 \cdot \varepsilon_1)$
$= \upsilon(-y, x)$
$\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3 \cdot \frac{1}{2}(y_1\varepsilon_0 \cdot \varepsilon_1 + y_2\varepsilon_0 \cdot \varepsilon_2 + y_3\varepsilon_0 \cdot \varepsilon_3 + x_3\varepsilon_2 \cdot \varepsilon_1 + x_2\varepsilon_1 \cdot \varepsilon_3 + x_1\varepsilon_3 \cdot \varepsilon_2)$
$= \frac{1}{2}(-y_1\varepsilon_3 \cdot \varepsilon_2 - y_2\varepsilon_1 \cdot \varepsilon_3 - y_3\varepsilon_2 \cdot \varepsilon_1 + x_3\varepsilon_0 \cdot \varepsilon_3 + x_2\varepsilon_0 \cdot \varepsilon_3 + x_1\varepsilon_0 \cdot \varepsilon_1)$
$= \upsilon(-y, x)$
In $Cl(1, 3)$:
$\frac{1}{2}(y_1\varepsilon_0 \cdot \varepsilon_1 + y_2\varepsilon_0 \cdot \varepsilon_2 + y_3\varepsilon_0 \cdot \varepsilon_3 + x_3\varepsilon_2 \cdot \varepsilon_1 + x_2\varepsilon_1 \cdot \varepsilon_3 + x_1\varepsilon_3 \cdot \varepsilon_2) \cdot \varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$
$= \frac{1}{2}(-y_1\varepsilon_3 \cdot \varepsilon_2 - y_2\varepsilon_1 \cdot \varepsilon_3 - y_3\varepsilon_2 \cdot \varepsilon_1 + x_3\varepsilon_0 \cdot \varepsilon_3 + x_2\varepsilon_0 \cdot \varepsilon_3 + x_1\varepsilon_0 \cdot \varepsilon_1)$
$= \upsilon(-y, x)$
$\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3 \cdot \frac{1}{2}(y_1\varepsilon_0 \cdot \varepsilon_1 + y_2\varepsilon_0 \cdot \varepsilon_2 + y_3\varepsilon_0 \cdot \varepsilon_3 + x_3\varepsilon_2 \cdot \varepsilon_1 + x_2\varepsilon_1 \cdot \varepsilon_3 + x_1\varepsilon_3 \cdot \varepsilon_2)$
$= \frac{1}{2}(-y_1\varepsilon_3 \cdot \varepsilon_2 - y_2\varepsilon_1 \cdot \varepsilon_3 - y_3\varepsilon_2 \cdot \varepsilon_1 + x_3\varepsilon_0 \cdot \varepsilon_3 + x_2\varepsilon_0 \cdot \varepsilon_3 + x_1\varepsilon_0 \cdot \varepsilon_1)$
$= \upsilon(-y, x)$

Product on $Spin(3,1)$

Because they belong to $Cl_0(3,1)$ the elements of $Spin(3,1)$ can be written :
$s = a + \frac{1}{2}\left(w^1\varepsilon_0 \cdot \varepsilon_1 + w^2\varepsilon_0 \cdot \varepsilon_2 + w^3\varepsilon_0 \cdot \varepsilon_3 + r^3\varepsilon_2 \cdot \varepsilon_1 + r^2\varepsilon_1 \cdot \varepsilon_3 + r^1\varepsilon_3 \cdot \varepsilon_2\right) + b\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$
where $a, (w^j, r^j)_{j=1}^3, b$ are real scalar which are related. That we will write with

$$\varepsilon_5 = \varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3 \tag{A.3}$$

$$s = a + \upsilon(r, w) + b\varepsilon_5 \tag{A.4}$$

And similarly in $Cl(1,3)$
$s = a + \upsilon(r, w) + b\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$
The product of two elements of the spin group expressed as :
$s = a + \upsilon(r, w) + b\varepsilon_5$
$s' = a' + \upsilon(r', w') + b'\varepsilon_5$
can be computed with the previous formulas.
$(a + \upsilon(r, w) + b\varepsilon_5) \cdot (a' + \upsilon(r', w') + b'\varepsilon_5)$
$= aa' + a'\upsilon(r, w) + a'b\varepsilon_5 + a\upsilon(r', w') + \upsilon(r, w) \cdot \upsilon(r', w') + b\varepsilon_5 \cdot \upsilon(r', w') + ab'\varepsilon_5 + b'\upsilon(r, w) \cdot \varepsilon_5 + bb'\varepsilon_5 \cdot \varepsilon_5$
$= aa' - bb' + a'\upsilon(r, w) + a\upsilon(r', w') + \upsilon(r, w) \cdot \upsilon(r', w') + b\upsilon(-w', r') + b'\upsilon(-w, r) \cdot \varepsilon_5 + (a'b + ab')\varepsilon_5$

$= aa' - bb' + \upsilon(a'r + ar' - bw' - b'w, a'w + aw' + br' + b'r) + \upsilon(r,w) \cdot \upsilon(r',w') + (a'b + ab')\varepsilon_5$

i) With signature (3,1)

$\upsilon(r,w) \cdot \upsilon(r',w') = \frac{1}{4}(w^t w' - r^t r') + \frac{1}{2}\upsilon(j(r)r' - j(w)w', j(w)r' + j(r)w') - \frac{1}{4}(w^t r' + r^t w')\varepsilon_5$

$(a + \upsilon(r,w) + b\varepsilon_5) \cdot (a' + \upsilon(r',w') + b'\varepsilon_5) = a" + \upsilon(r", w") + b"\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$

$a" = aa' - bb' + \frac{1}{4}(w^t w' - r^t r')$

$b" = (a'b + ab') - \frac{1}{4}(w^t r' + r^t w')$

$r" = a'r + ar' - bw' - b'w + \frac{1}{2}(j(r)r' - j(w)w')$

$w" = a'w + aw' + br' + b'r + \frac{1}{2}(j(w)r' + j(r)w')$

ii) With signature (1,3)

$\upsilon(r,w) \cdot \upsilon(r',w') = \frac{1}{4}(w^t w' - r^t r') - \frac{1}{2}\upsilon(-j(r)r' + j(w)w', j(w)r' + j(r)w') - \frac{1}{4}(w^t r' + r^t w')\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$

$(a + \upsilon(r,w) + b\varepsilon_5) \cdot (a' + \upsilon(r',w') + b'\varepsilon_5) = a" + \upsilon(r", w") + b"\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$

$a" = aa' - bb' + \frac{1}{4}(w^t w' - r^t r')$

$b" = (a'b + ab') - \frac{1}{4}(w^t r' + r^t w')$

$r" = a'r + ar' - bw' - b'w + a'r + ar' - bw' - b'w + \frac{1}{2}(j(r)r' - j(w)w')$

$w" = a'w + aw' + br' + b'r - \frac{1}{2}(j(w)r' + j(r)w')$

A.1.2 Characterization of the elements of the Spin group

Inverse

The elements of $Spin(3,1)$ are the product of an even number of vectors of norm ± 1. Consequently we have :

$s \cdot s^t = (v_1 \cdot ... v_{2p}) \cdot (v_{2p} \cdot ... \cdot v_1) = 1$

The transposition is an involution on the Clifford algebra, thus :

$(a + \upsilon(r,w) + b\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3) \cdot \left(a + \upsilon(r,w)^t + b\varepsilon_3 \cdot \varepsilon_2 \cdot \varepsilon_1 \cdot \varepsilon_0\right) = 1$

$(a + \upsilon(r,w) + b\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3) \cdot (a - \upsilon(r,w) + b\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3) = 1$

$\Leftrightarrow (a + \upsilon(r,w) + b\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3)^{-1} = (a - \upsilon(r,w) + b\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3)$

and we have the same result in $Cl(1,3)$

$$(a + \upsilon(r,w) + b\varepsilon_5)^{-1} = a - \upsilon(r,w) + b\varepsilon_5 \tag{A.5}$$

Relation between a,b, r, w

By a straightforward computation this identity gives the following relation between a,b,r,w :

1. With signature (3,1)

$(a + \upsilon(r,w) + b\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3) \cdot (a - \upsilon(r,w) + b\varepsilon_3 \cdot \varepsilon_2 \cdot \varepsilon_1 \cdot \varepsilon_0) = 1$

$= a" + \upsilon(r", w") + b"\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$

with :

$a" = a^2 - b^2 + \frac{1}{4}(-w^t w + r^t r) = 1$

$b" = ab + ba - \frac{1}{4}(-w^t r - r^t w) = 0$

$r" = \frac{1}{2}(-j(r)r + j(w)w) + ar - ar - bw + bw = 0$

$w" = \frac{1}{2}(-j(w)r - j(r)w) + aw - aw + br - br = 0$

$a^2 - b^2 = 1 + \frac{1}{4}(w^t w - r^t r)$

So, for any element : $a + \upsilon(r,w) + b\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$

we have :

$$a^2 - b^2 = 1 + \frac{1}{4}\left(w^t w - r^t r\right) \tag{A.6}$$

$$ab = -\frac{1}{4} r^t w \tag{A.7}$$

and we keep only 6 free parameters. a, b are defined from $r, w,$ up to sign, with the conditions:
i) $r^t w \neq 0 : b = -\frac{1}{4a} r^t w$
$a^2 = \frac{1}{2}\left(\left(1 + \frac{1}{4}\left(w^t w - r^t r\right)\right) + \sqrt{\left(1 + \frac{1}{4}\left(w^t w - r^t r\right)\right)^2 + \frac{1}{4}\left(r^t w\right)^2}\right)$
ii) $r^t w = 0 :$
$\left(w^t w - r^t r\right) \geq -4 : a = \epsilon\sqrt{1 + \frac{1}{4}\left(w^t w - r^t r\right)}; b = 0$
$\left(w^t w - r^t r\right) \leq -4 : b = \epsilon\sqrt{-\left(1 + \frac{1}{4}\left(w^t w - r^t r\right)\right)}; a = 0$
So :
if $r = 0$ then : $s = \epsilon\sqrt{1 + \frac{1}{4} w^t w} + \upsilon(0, w)$
if $w = 0$ then
$r^t r \leq 4 : s = \epsilon\sqrt{1 - \frac{1}{4} r^t r} + \upsilon(r, 0)$
$r^t r \geq 4 : s = \upsilon(r, 0) + \epsilon\sqrt{\frac{1}{4} r^t r - 1}\varepsilon_5$

2. With signature (1,3)
$(a - \upsilon(r, w) + b\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3) \cdot (a + \upsilon(r, w) + b\varepsilon_3 \cdot \varepsilon_2 \cdot \varepsilon_1 \cdot \varepsilon_0) = 1$
$= a" + \upsilon(r", w") + b"\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$
with :
$r" = \frac{1}{2}(-j(r) r + j(w) w) + ar - ar + bw - bw = 0$
$w" = -\frac{1}{2}(-j(w) r - j(r) w) + aw - aw + br - br'$
$a" = a^2 - b^2 + \frac{1}{4}\left(-w^t w + r^t r\right)$
$b" = ab + ba - \frac{1}{4}\left(-w^t r - r^t w\right)$
we get the same relations.

A.1.3 Adjoint map

The adjoint map : $Ad : Spin(3,1) \times Cl(3,1) \rightarrow Cl(3,1) :: Ad_s X = s \cdot X \cdot s^{-1}$ has a specific action for vectors of F and for elements of the Lie algebra $T_1 Spin(3,1)$.

Action on vectors of F

$\forall X \in F, s \in Spin(3,1) : \mathbf{Ad}_s X = s \cdot X \cdot s^{-1}$
$X = X_0\varepsilon_0 + X_1\varepsilon_1 + X_2\varepsilon_2 + X_3\varepsilon_3$
$s = a + \upsilon(r, w) + b\varepsilon_5$
$\mathbf{Ad}_s X = (a + \upsilon(r, w) + b\varepsilon_5) \cdot X \cdot (a - \upsilon(r, w) + b\varepsilon_5)$
$= (a + \upsilon(r, w) + b\varepsilon_5) \cdot (aX - X \cdot \upsilon(r, w) + bX \cdot \varepsilon_5)$
$= a^2 X + ab(X \cdot \varepsilon_5 + \varepsilon_5 \cdot X) + b^2\varepsilon_5 \cdot X \cdot \varepsilon_5 + a(\upsilon(r, w) \cdot X - X \cdot \upsilon(r, w))$
$+ b(\upsilon(r, w) \cdot X \cdot \varepsilon_5 - \varepsilon_5 \cdot X \cdot \upsilon(r, w)) - \upsilon(r, w) \cdot X \cdot \upsilon(r, w)$

$X \cdot \varepsilon_5 = -X_0\varepsilon_1\varepsilon_2\varepsilon_3 - X_1\varepsilon_0\varepsilon_2\varepsilon_3 + X_2\varepsilon_0\varepsilon_1\varepsilon_3 - X_3\varepsilon_0\varepsilon_1\varepsilon_2$
$\varepsilon_5 \cdot X = X_0\varepsilon_1\varepsilon_2\varepsilon_3 + X_1\varepsilon_0\varepsilon_2\varepsilon_3 - X_2\varepsilon_0\varepsilon_1\varepsilon_3 + X_3\varepsilon_0\varepsilon_1\varepsilon_2$
$X \cdot \varepsilon_5 + \varepsilon_5 \cdot X = 0$
$\varepsilon_5 \cdot X \cdot \varepsilon_5 = -X\varepsilon_5\varepsilon_5 = X$

$\upsilon(r,w)\cdot X\cdot\varepsilon_5-\varepsilon_5\cdot X\cdot\upsilon(r,w)$
$=-\upsilon(r,w)\cdot\varepsilon_5\cdot X+X\cdot\varepsilon_5\cdot\upsilon(r,w)$
$=-\upsilon(-w,r)\cdot X+X\cdot\upsilon(-w,r)$

$\mathbf{Ad}_s X=\left(a^2+b^2\right)X+a\left(\upsilon(r,w)\cdot X-X\cdot\upsilon(r,w)\right)$
$-b\left(\upsilon(-w,r)\cdot X-X\cdot\upsilon(-w,r)\right)-\upsilon(r,w)\cdot X\cdot\upsilon(r,w)$

$2\upsilon(r,w)\cdot X$
$=X_0\left(y_1\varepsilon_1+y_2\varepsilon_2+y_3\varepsilon_3-x_3\varepsilon_0\varepsilon_1\varepsilon_2+x_2\varepsilon_0\varepsilon_1\varepsilon_3-x_1\varepsilon_0\varepsilon_2\varepsilon_3\right)$
$+X_1\left(y_1\varepsilon_0-y_2\varepsilon_0\varepsilon_1\varepsilon_2-y_3\varepsilon_0\varepsilon_1\varepsilon_3+x_3\varepsilon_2-x_2\varepsilon_3-x_1\varepsilon_1\varepsilon_2\varepsilon_3\right)$
$+X_2\left(y_1\varepsilon_0\varepsilon_1\varepsilon_2+y_2\varepsilon_0-y_3\varepsilon_0\varepsilon_2\varepsilon_3-x_3\varepsilon_1-x_2\varepsilon_1\varepsilon_2\varepsilon_3+x_1\varepsilon_3\right)$
$+X_3\left(y_1\varepsilon_0\varepsilon_1\varepsilon_3+y_2\varepsilon_0\varepsilon_2\varepsilon_3+y_3\varepsilon_0-x_3\varepsilon_1\varepsilon_2\varepsilon_3+x_2\varepsilon_1-x_1\varepsilon_2\right)$
$2X\cdot\upsilon(r,w)$
$=X_0\left(-y_1\varepsilon_1-y_2\varepsilon_2-y_3\varepsilon_3-x_3\varepsilon_0\varepsilon_1\varepsilon_2+x_2\varepsilon_0\varepsilon_1\varepsilon_3-x_1\varepsilon_0\varepsilon_2\varepsilon_3\right)$
$+X_1\left(-y_1\varepsilon_0-y_2\varepsilon_0\varepsilon_1\varepsilon_2-y_3\varepsilon_0\varepsilon_1\varepsilon_3-x_3\varepsilon_2+x_2\varepsilon_3-x_1\varepsilon_1\varepsilon_2\varepsilon_3\right)$
$+X_2\left(y_1\varepsilon_0\varepsilon_1\varepsilon_2-y_2\varepsilon_0-y_3\varepsilon_0\varepsilon_2\varepsilon_3+x_3\varepsilon_1-x_2\varepsilon_1\varepsilon_2\varepsilon_3-x_1\varepsilon_3\right)$
$+X_3\left(y_1\varepsilon_0\varepsilon_1\varepsilon_3+y_2\varepsilon_0\varepsilon_2\varepsilon_3-y_3\varepsilon_0-x_3\varepsilon_1\varepsilon_2\varepsilon_3-x_2\varepsilon_1+x_1\varepsilon_2\right)$
$\left(\upsilon(r,w)\cdot X-X\cdot\upsilon(r,w)\right)$
$=X_0\left(w_1\varepsilon_1+w_2\varepsilon_2+w_3\varepsilon_3\right)+X_1\left(w_1\varepsilon_0+r_3\varepsilon_2-r_2\varepsilon_3\right)+X_2\left(w_2\varepsilon_0-r_3\varepsilon_1+r_1\varepsilon_3\right)+X_3\left(w_3\varepsilon_0+r_2\varepsilon_1-r_1\varepsilon_2\right)$
$4\upsilon(r,w)\cdot X\cdot\upsilon(r,w)$
$=\varepsilon_0\left(-X_0\left(r^2+w^2\right)-2X_1\left(r_3w_2-r_2w_3\right)+2X_3\left(r_1w_2-r_2w_1\right)\right)$
$+\varepsilon_1\left(-2X_0\left(r_2w_3-r_3w_2\right)-X_1\left(r_1^2-r_2^2-r_3^2+w_1^2-w_2^2-w_3^2\right)-2X_2\left(r_1r_2+w_1w_2\right)-2X_3\left(r_1r_3+w_1w_3\right)\right)+$
$+\varepsilon_2\left(-2X_0\left(r_3w_1-r_1w_3\right)-2X_1\left(r_1r_2+w_1w_2\right)+X_2\left(r_1^2-r_2^2+r_3^2+w_1^2-w_2^2+X_2w_3^2\right)-2X_3\left(r_2r_3+w_2w_3\right)\right)$
$+\varepsilon_3\left(-2X_0\left(r_1w_2-r_2w_1\right)-2X_1\left(r_1r_3+w_1w_3\right)-2X_2\left(r_2r_3+w_2w_3\right)+X_3\left(r_1^2+r_2^2-r_3^2+w_1^2+w_2^2-w_3^2\right)\right)$

$\mathbf{Ad}_s X$
$=\varepsilon_0\{X_3\left(aw_3-br_3-\frac{1}{2}r_1w_2+\frac{1}{2}r_2w_1\right)+X_2\left(aw_2-br_2\right)+X_1\left(aw_1-br_1-\frac{1}{2}r_2w_3+\frac{1}{2}r_3w_2\right)+$
$\frac{1}{4}X_0\left(r^2+w^2\right)\}$
$+\varepsilon_1\{X_3\left(ar_2+bw_2+\frac{1}{2}r_1r_3+\frac{1}{2}w_1w_3\right)+X_2\left(\frac{1}{2}r_1r_2-bw_3-ar_3+\frac{1}{2}w_1w_2\right)$
$+X_1\left(\frac{1}{4}r_1^2-\frac{1}{4}r_2^2-\frac{1}{4}r_3^2+\frac{1}{4}w_1^2-\frac{1}{4}w_2^2-\frac{1}{4}w_3^2\right)+X_0\left(aw_1-br_1+\frac{1}{2}r_2w_3-\frac{1}{2}r_3w_2\right)\}$
$+\varepsilon_2\{X_3\left(\frac{1}{2}r_2r_3-bw_1-ar_1+\frac{1}{2}w_2w_3\right)+X_2\left(\frac{1}{4}r_2^2-\frac{1}{4}r_1^2-\frac{1}{4}r_3^2-\frac{1}{4}w_1^2+\frac{1}{4}w_2^2-\frac{1}{4}w_3^2\right)$
$+X_1\left(ar_3+bw_3+\frac{1}{2}r_1r_2+\frac{1}{2}w_1w_2\right)+X_0\left(aw_2-br_2-\frac{1}{2}r_1w_3+\frac{1}{2}r_3w_1\right)\}$
$+\varepsilon_3\{X_3\left(\frac{1}{4}r_3^2-\frac{1}{4}r_2^2-\frac{1}{4}r_1^2-\frac{1}{4}w_1^2-\frac{1}{4}w_2^2+\frac{1}{4}w_3^2\right)+X_2\left(ar_1+bw_1+\frac{1}{2}r_2r_3+\frac{1}{2}w_2w_3\right)$
$+X_1\left(\frac{1}{2}r_1r_3-bw_2-ar_2+\frac{1}{2}w_1w_3\right)+X_0\left(aw_3-br_3+\frac{1}{2}r_1w_2-\frac{1}{2}r_2w_1\right)\}$
In matrix form :
$[\mathbf{h}(s)]=$
$$\begin{bmatrix} a^2+b^2+\frac{1}{4}\left(r^tr+w^tw\right) & aw^t-br^t+\frac{1}{2}w^tj(r)\\ aw-br+\frac{1}{2}j(r)w & a^2+b^2+\frac{1}{4}\left(r^tr+w^tw\right)+aj(r)+bj(w)+\frac{1}{2}\left(j(r)j(r)+j(w)j(w)\right)\end{bmatrix}$$

Action on the Lie algebra

With
$g=a+\upsilon(r,w)+b\varepsilon_5$
$Z=\upsilon(x,y)$
$Ad_gX=\left(a+\upsilon(r,w)+b\varepsilon_5\right)\cdot\upsilon(x,y)\cdot\left(a-\upsilon(r,w)+b\varepsilon_5\right)$
A straightforward computation gives :
$\mathbf{Ad}_g\upsilon(x,y)=\left(a+\upsilon(r,w)+b\varepsilon_5\right)\cdot\upsilon(x,y)\cdot\left(a-\upsilon(r,w)+b\varepsilon_5\right)$
$=\left(a^2-b^2\right)\upsilon(x,y)+2ab\upsilon(-y,x)+a\left[\upsilon(r,w),\upsilon(x,y)\right]+b\left[\upsilon(r,w),\upsilon(-y,x)\right]-\upsilon(r,w)\upsilon(x,y)\upsilon(r,w)$
$=\upsilon\{\left[a^2-b^2+aj(r)-bj(w)\right]x-\left[2ab+aj(w)+bj(r)\right]y,$

$[2ab + aj(w) + bj(r)]\,x + \left[a^2 - b^2 + aj(r) - bj(w)\right] y - \upsilon(r,w)\,\upsilon(x,y)\,\upsilon(r,w)\}$
with
$\upsilon(x,y)\,\varepsilon_5 = \varepsilon_5\upsilon(x,y) = \upsilon(-y,x)$
$\varepsilon_5\upsilon(x,y)\,\varepsilon_5 = \varepsilon_5\upsilon(-y,x) = \upsilon(-x,-y) = -\upsilon(x,y)$
$\upsilon(r,w)\cdot\upsilon(x,y)\cdot\upsilon(r,w)$
$= \frac{1}{2}\upsilon\{\left(j(w)j(w) - j(r)j(r) + 2\left(a^2 - b^2 - 1\right)\right)x + \left(j(r)j(w) + j(w)j(r) - 4ab\right)y,$
$-\left(j(r)j(w) + j(w)j(r) - 4ab\right)x + \left(j(w)j(w) - j(r)j(r) + 2\left(a^2 - b^2 - 1\right)\right)y\}$

$\mathbf{Ad}_g =$
$$\begin{bmatrix} 1 + aj(r) - bj(w) + \frac{1}{2}\left(j(r)j(r) - j(w)j(w)\right) & -\left(aj(w) + bj(r) + \frac{1}{2}\left(j(r)j(w) + j(w)j(r)\right)\right) \\ aj(w) + bj(r) + \frac{1}{2}\left(j(r)j(w) + j(w)j(r)\right) & 1 + aj(r) - bj(w) + \frac{1}{2}\left(j(r)j(r) - j(w)j(w)\right) \end{bmatrix}$$
With $s_w = a_w + \upsilon(0,w)$
$$[\mathbf{Ad}_s] = \begin{bmatrix} \left[1 - \frac{1}{2}j(w)j(w)\right] & -\left[a_w j(w)\right] \\ \left[a_w j(w)\right] & \left[1 - \frac{1}{2}j(w)j(w)\right] \end{bmatrix}$$
With $s_r = a_r + \upsilon(r,0)$
$$[\mathbf{Ad}_s] = \begin{bmatrix} \left[1 + a_r j(r) + \frac{1}{2}j(r)j(r)\right] & 0 \\ 0 & \left[1 + a_r j(r) + \frac{1}{2}j(r)j(r)\right] \end{bmatrix}$$

A.1.4 Homogeneous Space

The Clifford algebras and Spin Group structures are built from the product of vectors. The Clifford Algebras as well as the corresponding Spin groups, for any vector space F of the same dimension and bilinear form of the same signature are algebraically isomorphic.

The structure $Cl(3)$ can be defined from a set of vectors only if their scalar product is always definite positive. So, in a given vector space $(F, \langle\rangle)$ with Clifford Algebra isomorphic to $Cl(3,1)$ the set isomorphic to $Cl(3)$ is not unique : there is one set for each choice of a vector $\varepsilon_0 \in F$ such that $\langle\varepsilon_0, \varepsilon_0\rangle = -1$. In each set isomorphic to $Cl(3)$ there is a unique group withe the algebraic structure $Spin(3)$ .The Clifford Algebra $Cl(3)$ is a subalgebra of $Cl(3,1)$ and $Spin(3)$ a subgroup of $Spin(3,1)$.

The sets isomorphic to $Spin(3)$

Let us choose a vector $\varepsilon_0 \in F : \langle\varepsilon_0, \varepsilon_0\rangle = -1$ (+1 for the signature (1 3)). In F let be $F^\perp$ the orthogonal complement to $\varepsilon_0 : F^\perp = \{u \in F : \langle\varepsilon_0, u\rangle = 0\}$. This is a 3 dimensional vector space. The scalar product induced on $F^\perp$ by $\langle\rangle$ is definite positive : in a basis of $F^\perp$ its matrix has 3 positive eigen values, otherwise with ε_0 we would have another signature. The Clifford Algebra $Cl\left(F^\perp, \langle\rangle_\perp\right)$ generated by $\left(F^\perp, \langle\rangle_\perp\right)$ is a subset of $Cl(F, \langle\rangle)$, Clifford isomorphic to $Cl(3)$. The Spin group of $Cl\left(F^\perp, \langle\rangle_\perp\right)$ is algebraically isomorphic to $Spin(3)$.

Theorem 111 *The Spin group $Spin(3)$ of $Cl\left(F^\perp, \langle\rangle_\perp\right)$ is the set of elements of the spin group $Spin(3,1)$ of $Cl(F, \langle\rangle)$ which leave ε_0 unchanged : $\mathbf{Ad}_{s_r}\varepsilon_0 = s_r \cdot \varepsilon_0 \cdot s_r^{-1} = \varepsilon_0$.They read : $s = \epsilon\sqrt{1 - \frac{1}{4}r^t r} + \upsilon(r,0)$*

Proof. i) In any orthonormal basis the elements of $Spin(3)$ are a subgroup of $Spin(3,1)$. They read :
$s_r = a + \upsilon(r,w) + b\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$
but $b = 0, w = 0$ because they are built without ε_0 and then
$a^2 = 1 - \frac{1}{4}r^t r$
$s_r \cdot \varepsilon_0 \cdot s_r^{-1} = \mathbf{Ad}_{s_r}\varepsilon_0$

$[\mathbf{Ad}_{s_r}] = \begin{bmatrix} 1 & 0 \\ 0 & 1 + aj(r) + \frac{1}{2}(j(r)j(r)) \end{bmatrix}$

$\mathbf{Ad}_{s_r}\varepsilon_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

ii) Conversely let us show that $E = \{s \in Spin(3,1) : s \cdot \varepsilon_0 = \varepsilon_0 \cdot s\} = Spin(3)$

if $s_r \cdot \varepsilon_0 = \varepsilon_0 \cdot s_r$

$s_r = a + \upsilon(r,w) + b\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$

$s_r \cdot \varepsilon_0 = \varepsilon_0 \cdot s_r$

In $Cl(3,1)$:

$s \cdot \varepsilon_0 = a\varepsilon_0 + \upsilon(r,w)\varepsilon_0 - b\varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3 = \varepsilon_0 \cdot s = a\varepsilon_0 + \varepsilon_0\upsilon(r,w) + b\varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$

$\upsilon(r,w)\varepsilon_0 =$

$= \frac{1}{2}\left(w^1\varepsilon_1 + w^2\varepsilon_2 + w^3\varepsilon_3 - r^3\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 + r^2\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_3 - r^1\varepsilon_0 \cdot \varepsilon_2 \cdot \varepsilon_3\right)$

$\varepsilon_0\upsilon(r,w)$

$= \frac{1}{2}\left(-w^1\varepsilon_1 - w^2\varepsilon_2 - w^3\varepsilon_3 - r^3\varepsilon_0\varepsilon_1 \cdot \varepsilon_2 + r^2\varepsilon_0\varepsilon_1 \cdot \varepsilon_3 - r^1\varepsilon_0\varepsilon_2 \cdot \varepsilon_3\right)$

$a\varepsilon_0 + \frac{1}{2}\left(w^1\varepsilon_1 + w^2\varepsilon_2 + w^3\varepsilon_3 - r^3\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 + r^2\varepsilon_0\varepsilon_1 \cdot \varepsilon_3 - r^1\varepsilon_0\varepsilon_2 \cdot \varepsilon_3\right) - b\varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$

$= a\varepsilon_0 + \frac{1}{2}\left(-w^1\varepsilon_1 - w^2\varepsilon_2 - w^3\varepsilon_3 - r^3\varepsilon_0\varepsilon_1 \cdot \varepsilon_2 + r^2\varepsilon_0\varepsilon_1 \cdot \varepsilon_3 - r^1\varepsilon_0\varepsilon_2 \cdot \varepsilon_3\right) + b\varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$

$\Rightarrow w = 0, b = 0$

In $Cl(1,3)$:

$s \cdot \varepsilon_0 = a\varepsilon_0 - \upsilon(g)\varepsilon_0 - b\varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3 = \varepsilon_0 \cdot s = a\varepsilon_0 - \varepsilon_0\upsilon(g) + b\varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3 \Rightarrow b = 0$

$\upsilon(g)\varepsilon_0 =$

$= \frac{1}{2}\left(-w^4\varepsilon_1 - w^2\varepsilon_2 - w^3\varepsilon_3 - r^3\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 + r^2\varepsilon_0\varepsilon_1 \cdot \varepsilon_3 - r^1\varepsilon_0\varepsilon_2 \cdot \varepsilon_3\right)$

$\varepsilon_0\upsilon(g)$

$= \frac{1}{2}\left(w^{41}\varepsilon_1 + w^2\varepsilon_2 + w^3\varepsilon_3 - r^3\varepsilon_0\varepsilon_1 \cdot \varepsilon_2 + r^2\varepsilon_0\varepsilon_1 \cdot \varepsilon_3 - r^1\varepsilon_0\varepsilon_2 \cdot \varepsilon_3\right)$

$\Rightarrow w = 0$

So the elements such that $s = \upsilon(r,0) + \epsilon\sqrt{\frac{1}{4}r^t r - 1}\varepsilon_5$ are excluded and we are left with

$E = \{s \in Spin(3,1) : s \cdot \varepsilon_0 = \varepsilon_0 \cdot s\} = \left\{\epsilon\sqrt{1 - \frac{1}{4}r^t r} + \upsilon(r,0)\right\}$

E has a group structure with $\cdot$ as it can be easily checked :

$\left(\epsilon\sqrt{1 - \frac{1}{4}r^t r} + \upsilon(r,0)\right) \cdot \left(\epsilon'\sqrt{1 - \frac{1}{4}r'^t r'} + \upsilon(r',0)\right)$

$= \epsilon\sqrt{1 - \frac{1}{4}r^t r}\epsilon'\sqrt{1 - \frac{1}{4}r'^t r'} - \frac{1}{4}r^t r' + \upsilon\left(\frac{1}{2}j(r)r' + r\epsilon'\sqrt{1 - \frac{1}{4}r'^t r'} + r'\epsilon\sqrt{1 - \frac{1}{4}r^t r}, 0\right)$

It is comprised of products of vectors of $(\varepsilon_i)_{i=1}^3$, so it belongs to $Cl\left(F^\perp, \langle\rangle_\perp\right)$, it is a Lie group of dimension 3 and so $E = Spin(3)$. ■

The scalars $\epsilon = \pm 1$ belong to the group. The group is not connected. The elements $s = \sqrt{1 - \frac{1}{4}r^t r} + \upsilon(r,0)$ constitute the component of the identity.

Homogeneous space

The quotient space $SW = Spin(3,1)/Spin(3)$ (called a homogeneous space) is not a group but a 3 dimensional manifold. It is characterized by the equivalence relation :

$s = a + \upsilon(r,w) + b\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3 \sim s' = a' + \upsilon(r',w') + b'\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$

$\Leftrightarrow \exists s_r \in Spin(3) : s' = s \cdot s_r$

As any quotient space its elements are *subsets* of $Spin(3,1)$.

Theorem 112 *In each class of the homogeneous space there are two elements, defined up to sign, which read :* $s_w = \pm(a_w + \upsilon(0,w))$

Proof. Each coset $[s] \in SW$ is in bijective correspondence with $Spin(3)$.

Any element of $Spin(3)$ reads $\epsilon\sqrt{1 - \frac{1}{4}\rho^t\rho} + \upsilon(\rho,0)$.

So $[s]=\left\{s'=s\cdot\left(\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho}+\upsilon(\rho,0)\right),\rho^t\rho\leq 4\right\}$

i) In $Spin(3,1)$:

$s=a+\upsilon(r,w)+b\varepsilon_5$

$s'=a'+\upsilon(r',w')+b'\varepsilon_5$

$a'=a\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho}-\frac{1}{4}r^t\rho$

$b'=b\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho}-\frac{1}{4}w^t\rho$

$r'=\frac{1}{2}j(r)\rho+r\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho}+a\rho$

$w'=\frac{1}{2}j(w)\rho+w\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho}+b\rho$

$a^2-b^2=1+\frac{1}{4}(w^tw-r^tr)$

$ab=-\frac{1}{4}r^tw$

ii) We can always choose in the class an element s' such that : $r'=0$. It requires : $\left(\frac{1}{2}j(r)+aI\right)\rho=-r\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho}$

$x=\left(\frac{1}{a}-\frac{b}{a^2+b^2r^tr}j(r)-\frac{b^2}{a(a^2+b^2r^tr)}j(r)j(r)\right)y$

This linear equation in ρ has always a unique solution :

$\rho=-\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho}\frac{1}{a}r$

$\rho^t\rho=\left(1-\frac{1}{4}\rho^t\rho\right)\frac{1}{a^2}(r^tr)\Rightarrow$

$\left(a^2+\frac{1}{4}(r^tr)\right)\rho^t\rho=(r^tr)$

$\rho^t\rho=\frac{4(r^tr)}{4a^2+(r^tr)}\leq 4$

$\sqrt{1-\frac{1}{4}\rho^t\rho}=\sqrt{\frac{4a^2}{4a^2+r^tr}}=\frac{2a}{\sqrt{4a^2+r^tr}}$

$\rho=-\epsilon\frac{2}{\sqrt{4a^2+r^tr}}r$

$\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho}+\upsilon(\rho,0)=\epsilon\frac{2a}{\sqrt{4a^2+r^tr}}-\upsilon\left(\epsilon\frac{2}{\sqrt{4a^2+r^tr}}r,0\right)$

$=\epsilon\left(\frac{2a}{\sqrt{4a^2+r^tr}}-\upsilon\left(\frac{2}{\sqrt{4a^2+r^tr}}r,0\right)\right)$

$a'=a\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho}-\frac{1}{4}r^t\rho=\frac{1}{2}\frac{\epsilon}{\sqrt{4a^2+r^tr}}\left(4a^2+r^tr\right)=\frac{1}{2}\epsilon\sqrt{4a^2+r^tr}$

$w'=\frac{1}{2}j(w)\rho+w\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho}+b\rho=\epsilon\frac{2}{\sqrt{4a^2+r^tr}}\left(\frac{1}{2}j(r)w+aw-br\right)$

$b'=b\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho}-\frac{1}{4}w^t\rho=\epsilon\frac{2}{\sqrt{4a^2+r^tr}}\left(ab+\frac{1}{4}w^tr\right)=0$

$s'=s_w=\frac{1}{2}\epsilon\sqrt{4a^2+r^tr}+\upsilon\left(0,\epsilon\frac{2}{\sqrt{4a^2+r^tr}}\left(\frac{1}{2}j(r)w+aw-br\right)\right)$

$=\epsilon\left(\frac{1}{2}\sqrt{4a^2+r^tr}+\upsilon\left(0,\frac{2}{\sqrt{4a^2+r^tr}}\left(\frac{1}{2}j(r)w+aw-br\right)\right)\right)$

$s'=s\cdot\left(\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho}+\upsilon(\rho,0)\right)$

$s=s'\cdot\left(\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho}+\upsilon(\rho,0)\right)^{-1}=s_w.\left(\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho}-\upsilon(\rho,0)\right)$

$=\epsilon\left(\frac{1}{2}\sqrt{4a^2+r^tr}+\upsilon\left(0,\frac{2}{\sqrt{4a^2+(r^tr)}}\left(\frac{1}{2}j(r)w+aw-br\right)\right)\right)\cdot\epsilon\left(\frac{2a}{\sqrt{4a^2+r^tr}}+\upsilon\left(\frac{2}{\sqrt{4a^2+(r^tr)}}r,0\right)\right)$

$s=a+\upsilon(r,w)+b\varepsilon_5=s_w\cdot s_r$

iii) In $Cl(1,3)$ we have the same decomposition with the same components.

$s=a+\upsilon(r,w)+b\varepsilon_5=s_w\cdot s_r$

$r"=\frac{1}{2}\epsilon\sqrt{4a^2+r^tr}\epsilon\frac{2}{\sqrt{4a^2+(r^tr)}}r=r$

$w"=\frac{1}{2}j\left(\left(\epsilon\sqrt{4a^2+r^tr}\right)\epsilon\frac{2}{4a^2+(r^tr)}\left(\frac{1}{2}j(r)w+aw-br\right)\right)\left(\epsilon\frac{2}{\sqrt{4a^2+(r^tr)}}\right)r$

$+\left(\epsilon\frac{2}{\sqrt{4a^2+(r^tr)}}\right)a\left(\epsilon\sqrt{4a^2+r^tr}\right)\epsilon\frac{2}{4a^2+(r^tr)}\left(\frac{1}{2}j\left(r\right)w+aw-br\right)$
$=2j\left(\epsilon\frac{1}{4a^2+(r^tr)}\left(\frac{1}{2}j\left(r\right)w+aw-br\right)\right)r+a\epsilon\frac{4}{4a^2+(r^tr)}\left(\frac{1}{2}j\left(r\right)w+aw-br\right)$
$=\left(\epsilon\frac{2}{4a^2+(r^tr)}\right)\left(j\left(\left(\frac{1}{2}j\left(r\right)w+aw-br\right)\right)r+a2\left(\frac{1}{2}j\left(r\right)w+aw-br\right)\right)$
$=\left(\epsilon\frac{2}{4a^2+(r^tr)}\right)\left(\frac{1}{2}j\left(j\left(r\right)w\right)r-aj\left(w\right)r+aj\left(r\right)w+2a^2w-2abr\right)$
$=\left(\epsilon\frac{2}{4a^2+(r^tr)}\right)\left(\frac{1}{2}\left(wr^t-rw^t\right)r+2a^2w+\frac{1}{2}\left(r^tw\right)r\right)$
$=\left(\epsilon\frac{2}{4a^2+(r^tr)}\right)\left(\frac{1}{2}w\left(r^tr\right)-\frac{1}{2}r\left(w^tr\right)+2a^2w+\frac{1}{2}\left(r^tw\right)r\right)$
$=\left(\epsilon\frac{1}{4a^2+(r^tr)}\right)\left(\left(4a^2+\left(r^tr\right)\right)w\right)=w$ ■

So any element of $Spin\left(3,1\right)$ can be written uniquely (up to sign) :
$s=a+\upsilon\left(r,w\right)+b\varepsilon_5=\epsilon s_w\cdot\epsilon s_r=\epsilon\left(a_w+\upsilon\left(0,w_w\right)\right)\cdot\epsilon\left(a_r+\upsilon\left(0,r_r\right)\right)$
$s_w=a_w+\upsilon\left(0,w_w\right)=\frac{1}{2}\sqrt{4a^2+r^tr}+\upsilon\left(0,\frac{2}{\sqrt{4a^2+(r^tr)}}\left(\frac{1}{2}j\left(r\right)w+aw-br\right)\right)$
$s_r=\left(a_r+\upsilon\left(0,r_r\right)\right)=\frac{2a}{\sqrt{4a^2+r^tr}}+\upsilon\left(\frac{2}{\sqrt{4a^2+(r^tr)}}r,0\right)$
$\epsilon a_ra_wa>0$
Remark : the elements $\pm s_w$ are equivalent :
$\left(a_w+\upsilon\left(0,w_w\right)\right)\sim-\left(a_w+\upsilon\left(0,w_w\right)\right)$
Take $s_r=-1\in Spin\left(3\right):-s_w=s_w\cdot s_r$
So $\pm s_w$ belong to the same class of equivalence. In the decomposition : $s=\epsilon s_w\cdot\epsilon s_r$, ϵs_w is a specific projection of s on the homogenous space.

A.1.5 Exponential on T_1Spin

The exponential on a Lie algebra is the flow of left invariant vector fields (Maths.22.2.6).
i) Left invariant vector fields on $Spin(3,1)$
As $Spin(3,1)\subset Cl\left(3,1\right)$ which is a vector space, a vector field $X\in\mathfrak{X}\left(TSpin\left(3,1\right)\right)$ reads $X\left(\sigma\right)\in Cl\left(3,1\right)$ with the relation :
$L_g'\sigma\left(X\left(\sigma\right)\right)=X\left(L_g\sigma\right)=g\cdot X\left(\sigma\right)=X\left(g\cdot\sigma\right)$
Thus the left invariant vector fields read :
$X\left(\sigma\right)=\sigma\cdot\upsilon\left(R,W\right)$ with $\upsilon\left(R,W\right)\in T_1Spin\left(3,1\right)$
ii) The flow of $X=\sigma\cdot\upsilon\left(R,W\right)\in\mathfrak{X}\left(TSpin\left(3,1\right)\right)$ reads:
$\Phi_X\left(t,1\right)=a\left(t\right)+\upsilon\left(r\left(t\right),w\left(t\right)\right)+b\left(t\right)\varepsilon_5\in Spin(3,1)$
$\Phi_X\left(t,1\right)=\exp tX=\exp\upsilon\left(tR,tW\right)$
$\exp\upsilon\left(tR,tW\right)=a\left(t\right)+\upsilon\left(r\left(t\right),w\left(t\right)\right)+b\left(t\right)\varepsilon_5$
$\frac{d}{dt}\exp\upsilon\left(tR,tW\right)|_{t=\theta}=\exp\upsilon\left(\theta R,\theta W\right)\cdot\upsilon\left(R,W\right)$
$\frac{d}{dt}\left(a\left(t\right)+\upsilon\left(r\left(t\right),w\left(t\right)\right)+b\left(t\right)\varepsilon_5\right)|_{t=\theta}=\left(a\left(\theta\right)+\upsilon\left(r\left(\theta\right),w\left(\theta\right)\right)+b\left(\theta\right)\varepsilon_5\right)\cdot\upsilon\left(R,W\right)$
with :
$a^2-b^2=1+\frac{1}{4}\left(w^tw-r^tr\right)$
$ab=-\frac{1}{4}r^tw$
$1=a\left(0\right)+\upsilon\left(r\left(0\right),w\left(0\right)\right)+b\left(0\right)\varepsilon_5$
The derivations give :
$\frac{\partial a}{\partial t}+\upsilon\left(\frac{\partial r}{\partial t},\frac{\partial w}{\partial t}\right)+\frac{\partial b}{\partial t}\varepsilon_5|_{t=\theta}=\left(a\left(\theta\right)+\upsilon\left(r\left(\theta\right),w\left(\theta\right)\right)+b\left(\theta\right)\varepsilon_5\right)\cdot\upsilon\left(R,W\right)$
$=a\left(\theta\right)\upsilon\left(R,W\right)+b\left(\theta\right)\upsilon\left(R,-W\right)+\upsilon\left(r\left(\theta\right),w\left(\theta\right)\right)\cdot\upsilon\left(R,W\right)$
$=\upsilon\left(\left(a\left(\theta\right)+b\left(\theta\right)\right)R,\left(a\left(\theta\right)-b\left(\theta\right)\right)W\right)$
$+\frac{1}{4}\left(W^tw-R^tr\right)+\frac{1}{2}\upsilon\left(-j\left(R\right)r+j\left(W\right)w,-j\left(W\right)r-j\left(R\right)w\right)-\frac{1}{4}\left(W^tr+R^tw\right)\varepsilon_5$

$\frac{\partial a}{\partial t}|_{t=\theta} = \frac{1}{4}\left(W^t w - R^t r\right)$
$\frac{\partial b}{\partial t}|_{t=\theta} = -\frac{1}{4}\left(W^t r + R^t w\right)$
$\frac{\partial r}{\partial t}|_{t=\theta} = \left(a\left(\theta\right) + b\left(\theta\right)\right) R + \frac{1}{2}\left(-j\left(R\right) r + j\left(W\right) w\right)$
$\frac{\partial w}{\partial t}|_{t=\theta} = \left(a\left(\theta\right) - b\left(\theta\right)\right) W + \frac{1}{2}\left(-j\left(W\right) r - j\left(R\right) w\right)$
$a\frac{\partial a}{\partial t}|_{t=\theta} - b\frac{\partial b}{\partial t}|_{t=\theta} = \frac{1}{4}\left(w^t \frac{\partial w}{\partial t} - r^t \frac{\partial r}{\partial t}\right)$
$\frac{\partial a}{\partial t}b + a\frac{\partial b}{\partial t} = -\frac{1}{4}r^t \frac{\partial w}{\partial t} - \frac{1}{4}w^t \frac{\partial r}{\partial t}$
The last two equations give :
$b\left(W + R\right)^t \left(w + r\right) = 0$
$b\left(W + R\right)^t \left(w - r\right) = 0$
iii) We have the morphism :
$\Pi : Spin(3,1) \to SO(3,1) :: \Pi\left(\pm\sigma\right) = \left[h\left(\sigma\right)\right]$ such that :
$\forall u \in \mathbb{R}^4 : \mathbf{Ad}_\sigma u = \sigma \cdot u \cdot \sigma^{-1} = \left[h\left(\sigma\right)\right] u = \Pi\left(\pm\sigma\right) u$
Take a vector field $X\left(\sigma\right) = \sigma \cdot \upsilon\left(R, W\right) \in \mathfrak{X}\left(TSpin\left(3,1\right)\right)$ then (Maths.1460) :
$\Pi \circ \Phi_X = \Phi_{\Pi_* X} \circ \Pi$
$\Pi_* X\left(\left[h\left(\sigma\right)\right]\right) = \Pi'\left(\sigma\right) X\left(\sigma\right)$
$\Pi'\left(1\right) X\left(1\right) = K\left(W\right) + J\left(R\right)$
$\Pi\left(\exp t\upsilon\left(R, W\right)\right) = \Phi_{\Pi'(1)\upsilon(R,W)}\left(t, \Pi\left(1\right)\right) = \exp t\left(K\left(W\right) + J\left(R\right)\right) = \exp tK\left(W\right) \exp tJ\left(R\right)$
$\Pi\left(\exp t\upsilon\left(R, 0\right)\right) = \exp tJ\left(R\right)$
$\Pi\left(\exp t\upsilon\left(0, W\right)\right) = \exp tK\left(W\right)$
$\Pi\left(\exp t\upsilon\left(R, W\right)\right) = \Pi\left(\exp t\upsilon\left(0, W\right)\right) \Pi\left(\exp t\upsilon\left(R, 0\right)\right)$
and because this is a morphism :
$\exp t\upsilon\left(R, W\right) = \exp t\upsilon\left(0, W\right) \cdot \exp t\upsilon\left(R, 0\right)$
iv) Coming back to the previous equations :
For $\exp \upsilon\left(0, tW\right)$:
$\frac{\partial a}{\partial t}|_{t=\theta} = \frac{1}{4}W^t w$
$\frac{\partial b}{\partial t}|_{t=\theta} = -\frac{1}{4}\left(W^t r\right)$
$\frac{\partial r}{\partial t}|_{t=\theta} = \frac{1}{2}j\left(W\right) w$
$\frac{\partial w}{\partial t}|_{t=\theta} = \left(a\left(\theta\right) - b\left(\theta\right)\right) W - \frac{1}{2}j\left(W\right) r$
$bW^t\left(w + r\right) = 0$
$bW^t\left(w - r\right) = 0$
if $b \neq 0$:
$W^t w = -W^t r = W^t r = 0$
$b\frac{\partial b}{\partial t}|_{t=\theta} = -\frac{1}{4}b\left(W^t r\right) = 0 \Rightarrow b^2 = Ct \Rightarrow b = Ct \Rightarrow W^t r = 0,$
$\Rightarrow W^t w = 0 \Rightarrow a = Ct$
$\Rightarrow r, w = Ct$
Thus $b = 0 \Rightarrow W^t r = 0$
$\frac{d^2 w}{dt^2} = \frac{1}{4}\left(W^t w\right) W - \frac{1}{4}j\left(W\right) j\left(W\right) w = \frac{1}{4}\left(W^t W\right) w$
$w\left(t\right) = w_1 \exp \frac{1}{2}t\sqrt{W^t W} + w_2 \exp\left(-\frac{1}{2}t\sqrt{W^t W}\right)$
$w\left(0\right) = 0 = w_1 + w_2$
$\frac{dw}{dt}\left(0\right) = W = w_1 - w_2$
$w\left(t\right) = \frac{1}{2}\left(\exp \frac{1}{2}t\sqrt{W^t W} - \exp\left(-\frac{1}{2}t\sqrt{W^t W}\right)\right) W = W \sinh \frac{1}{2}t\sqrt{W^t W}$
$\frac{\partial r}{\partial t}|_{t=\theta} = \frac{1}{2}j\left(W\right) W \sinh \frac{1}{2}t\sqrt{W^t W} = 0$
$r\left(0\right) = R = 0 \Rightarrow r\left(t\right) = 0$
$w^t w = W^t W \sinh^2 \frac{1}{2}t\sqrt{W^t W}$
$a^2 - b^2 = a^2 = 1 + \frac{1}{4}\left(W^t W \sinh^2 \frac{1}{2}t\sqrt{W^t W}\right)$
$\exp \upsilon\left(0, tW\right) = a_w\left(t\right) + \upsilon\left(0, w\left(t\right)\right)$

with
$a_w(t) = \sqrt{1 + \frac{1}{4}\left(W^t W \sinh^2 \frac{1}{2} t\sqrt{W^t W}\right)}$
$w(t) = \left(\sinh \frac{1}{2} t\sqrt{W^t W}\right)(W)$
$\frac{d}{dt}\left(a_w(t) + \upsilon(0, w(t))\right)|_{t=\theta} = \left(a_w(\theta) + \upsilon(0, w(\theta))\right) \cdot \upsilon(0, W)$
For $\exp \upsilon(tR, 0)$
$\frac{\partial a}{\partial t}|_{t=\theta} = -\frac{1}{4} R^t r$
$\frac{\partial b}{\partial t}|_{t=\theta} = -\frac{1}{4} R^t w$
$\frac{\partial r}{\partial t}|_{t=\theta} = (a(\theta) + b(\theta)) R - \frac{1}{2} j(R) r$
$\frac{\partial w}{\partial t}|_{t=\theta} = -\frac{1}{2} j(R) w$
$bR^t(w + r) = 0$
$bR^t(w - r) = 0$
$\Rightarrow b = 0, R^t w = 0$
$\frac{d^2 r}{dt^2} = -\frac{1}{4}(R^t r) R - \frac{1}{2} j(R)\left(aR - \frac{1}{2} j(R) r\right)$
$\frac{d^2 r}{dt^2} = -\frac{1}{4}(R^t r) R + \frac{1}{4}\left((R^t r) R - R^t R r\right)$
$\frac{d^2 r}{dt^2} = -\frac{1}{4}(R^t R) r$
$r(t) = r_1 \exp it\frac{1}{2}\sqrt{R^t R} + r_1 \exp\left(-it\frac{1}{2}\sqrt{R^t R}\right)$
$r(0) = 0 = r_1 + r_2$
$\frac{dr}{dt}(0) = R = r_1 - r_2$
$r(t) = R \sin t\frac{1}{2}\sqrt{R^t R}$
$r^t r = R^r R \sin^2 t\frac{1}{2}\sqrt{R^t R}$
$a^2 - b^2 = a^2 = 1 - \frac{1}{4} R^r R \sin^2 t\frac{1}{2}\sqrt{R^t R}$
$\exp t\upsilon(R, 0) = a_r(t) + \upsilon(r(t), 0)$
with :
$a_r(t) = \sqrt{1 - \frac{1}{4} R^r R \sin^2 t\frac{1}{2}\sqrt{R^t R}}$
$r(t) = \sin\left(t\frac{1}{2}\sqrt{R^t R}\right)(R)$

A.2 LIE DERIVATIVE

For the definition and properties of Lie derivatives see Maths.16.2.

One form

The Lie derivative of a 1 form on M $\lambda(m) = \sum_{a=1}^{m} \sum_{\alpha=0}^{3} \lambda_\alpha^a(m) d\xi^\alpha \otimes \overrightarrow{\theta}_a \in \Lambda_2(M; T_1 U)$ valued in a fixed vector space ($T_1 U$ can be replaced by any *fixed* vector space) with respect to the vector field $V = \sum_{\alpha=0}^{3} V^\alpha \partial\xi_\alpha \in \mathfrak{X}(TM)$ reads :

$\pounds_V \lambda(m) = \sum_{a=1}^{m} \sum_{\{\alpha\beta\}} \pounds_V\left(\lambda_\alpha^a(m) d\xi^\alpha\right) \otimes \overrightarrow{\theta}_a$
Using the properties of the Lie derivative :
$\pounds_V\left(\lambda_\alpha^a(m) d\xi^\alpha\right)$
$= \sum_{\alpha=0}^{3}\left(\pounds_V \lambda_\alpha^a(m)\right) d\xi^\alpha + \lambda_\alpha^a(m) \pounds_V(d\xi^\alpha)$
$\pounds_V \lambda_\alpha^a(m) = \sum_{\gamma=0}^{3}\left(V^\gamma \partial_\gamma \lambda_\alpha^a\right)$
$\pounds_V(d\xi^\alpha) = i_V d(d\xi^\alpha) + d(i_V d\xi^\alpha) = dV^\alpha = \sum_\gamma \partial_\gamma V^\alpha d\xi^\gamma$
$\pounds_V\left(\lambda_\alpha^a(m) d\xi^\alpha\right) = \sum_{\alpha=0}^{3}\left(\sum_{\gamma=0}^{3}\left(V^\gamma \partial_\gamma \lambda_\alpha^a\right)\right) d\xi^\alpha + \lambda_\alpha^a \sum_{\gamma=0}^{3} \partial_\gamma V^\alpha d\xi^\gamma$
$= \sum_{\alpha=0}^{3}\left(\sum_{\gamma=0}^{3}\left(V^\gamma \partial_\gamma \lambda_\alpha^a\right)\right) d\xi^\alpha + \sum_{\alpha=0}^{3} \lambda_\gamma^a \sum_{\gamma=0}^{3} \partial_\alpha V^\gamma d\xi^\alpha$

$= \sum_{\alpha,\gamma=0}^{3} \left(V^\gamma \partial_\gamma \lambda_\alpha^a + \lambda_\gamma^a \partial_\alpha V^\gamma\right) d\xi^\alpha$

$\pounds_V \lambda(m) = \sum_{a=1}^{m} \sum_{\alpha,\gamma=0}^{3} \left(V^\gamma \partial_\gamma \lambda_\alpha^a + \lambda_\gamma^a \partial_\alpha V^\gamma\right) d\xi^\alpha \otimes \overrightarrow{\theta}_a$

2 form

The Lie derivative of a 2 form on M $\mathcal{F}(m) = \sum_{a=1}^{m} \sum_{\{\alpha\beta\}} \mathcal{F}_{\alpha\beta}^a(m) \, d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\theta}_a \in \Lambda_2(M; T_1U)$ valued in a fixed vector space (T_1U can be replaced by any *fixed* vector space) with respect to the vector field $V = \sum_{\alpha=0}^{3} V^\alpha \partial\xi_\alpha \in \mathfrak{X}(TM)$ reads :

$\pounds_V \mathcal{F}(m) = \sum_{a=1}^{m} \sum_{\{\alpha\beta\}} \pounds_V \left(\mathcal{F}_{\alpha\beta}^a(m) \, d\xi^\alpha \wedge d\xi^\beta\right) \otimes \overrightarrow{\theta}_a$

Using the properties of the Lie derivative :

$\pounds_V \left(\mathcal{F}_{\alpha\beta}^a(m) \, d\xi^\alpha \wedge d\xi^\beta\right)$

$= \sum_{\{\alpha\beta\}} \left(\pounds_V \mathcal{F}_{\alpha\beta}^a(m)\right) d\xi^\alpha \wedge d\xi^\beta + \mathcal{F}_{\alpha\beta}^a(m) \, \pounds_V \left(d\xi^\alpha \wedge d\xi^\beta\right)$

$= \sum_{\{\alpha\beta\}} \left(\pounds_V \mathcal{F}_{\alpha\beta}^a(m)\right) d\xi^\alpha \wedge d\xi^\beta + \mathcal{F}_{\alpha\beta}^a(m) \left((\pounds_V d\xi^\alpha) \wedge d\xi^\beta + d\xi^\alpha \wedge \pounds_V d\xi^\beta\right)$

$= \sum_{\{\alpha\beta\}} \left(\sum_{\gamma=0}^{3} V^\gamma \partial_\gamma \mathcal{F}_{\alpha\beta}^a d\xi^\alpha \wedge d\xi^\beta\right) + \mathcal{F}_{\alpha\beta}^a \left((\pounds_V d\xi^\alpha) \wedge d\xi^\beta + d\xi^\alpha \wedge \pounds_V d\xi^\beta\right)$

$\pounds_V (d\xi^\alpha) = i_V d(d\xi^\alpha) + d(i_V d\xi^\alpha) = dV^\alpha = \sum_\gamma \partial_\gamma V^\alpha d\xi^\gamma$

$\pounds_V \left(d\xi^\alpha \wedge d\xi^\beta\right) = \left(\sum_\gamma \partial_\gamma V^\alpha d\xi^\gamma\right) \wedge d\xi^\beta + d\xi^\alpha \wedge \sum_\gamma \partial_\gamma V^\beta d\xi^\gamma$

We get the general formula :

$$\pounds_V \mathcal{F}(m) = \sum_{a=1}^{m} \left(\sum_{\{\alpha\beta\}} \sum_{\gamma=0}^{3} V^\gamma \partial_\gamma \mathcal{F}_{\alpha\beta}^a d\xi^\alpha \wedge d\xi^\beta + \mathcal{F}_{\alpha\beta}^a \left(\partial_\gamma V^\alpha d\xi^\gamma \wedge d\xi^\beta + \partial_\gamma V^\beta d\xi^\alpha \wedge d\xi^\gamma\right)\right) \otimes \overrightarrow{\theta}_a \tag{A.8}$$

For $a = 1...m, \{\alpha, \beta\} = 0...3$:

$\mathcal{F}(m) =$

$\sum_{a=1}^{m} \{\mathcal{F}_{32} d\xi^3 \wedge d\xi^2 + \mathcal{F}_{13} d\xi^1 \wedge d\xi^3 + \mathcal{F}_{21} d\xi^2 \wedge d\xi^1 + \mathcal{F}_{01} d\xi^0 \wedge d\xi^1 + \mathcal{F}_{02} d\xi^0 \wedge d\xi^2 + \mathcal{F}_{03} d\xi^0 \wedge d\xi^3\} \otimes \overrightarrow{\theta}_a$

$\sum_{\{\alpha\beta\}} \sum_\gamma V^\gamma \partial_\gamma \mathcal{F}_{\alpha\beta}^a d\xi^\alpha \wedge d\xi^\beta$

$= \sum_\gamma V^\gamma \partial_\gamma \mathcal{F}_{32} d\xi^3 \wedge d\xi^2 + V^\gamma \partial_\gamma \mathcal{F}_{13} d\xi^1 \wedge d\xi^3 + V^\gamma \partial_\gamma \mathcal{F}_{21} d\xi^2 \wedge d\xi^1$

$+ V^\gamma \partial_\gamma \mathcal{F}_{01} d\xi^0 \wedge d\xi^1 + V^\gamma \partial_\gamma \mathcal{F}_{02} d\xi^0 \wedge d\xi^2 + V^\gamma \partial_\gamma \mathcal{F}_{03} d\xi^0 \wedge d\xi^3$

A straightforward computation gives :

$[\pounds_V \mathcal{F}]_{01}^a = \partial_0 V^0 [\mathcal{F}^{aw}]_1 + \partial_1 V^1 [\mathcal{F}^{aw}]_1 + \partial_1 V^2 [\mathcal{F}^{aw}]_2 + \partial_1 V^3 [\mathcal{F}^{aw}]_3 - \partial_0 V^3 [\mathcal{F}^{ar}]_2 + \partial_0 V^2 [\mathcal{F}^{ar}]_3 + \sum_{\gamma=0}^{3} V^\gamma [\partial_\gamma \mathcal{F}^w]_1$

$[\pounds_V \mathcal{F}]_{02}^a = \partial_2 V^1 [\mathcal{F}^{aw}]_1 + \partial_0 V^0 [\mathcal{F}^{aw}]_2 + \partial_2 V^2 [\mathcal{F}^{aw}]_2 + \partial_2 V^3 [\mathcal{F}^{aw}]_3 + \partial_0 V^3 [\mathcal{F}^{ar}]_1 - \partial_0 V^1 [\mathcal{F}^{ar}]_3 + \sum_{\gamma=0}^{3} V^\gamma [\partial_\gamma \mathcal{F}^w]_2$

$[\pounds_V \mathcal{F}]_{03}^a = \partial_3 V^1 [\mathcal{F}^{aw}]_1 + \partial_3 V^2 [\mathcal{F}^{aw}]_2 + \partial_0 V^0 [\mathcal{F}^{aw}]_3 + \partial_3 V^3 [\mathcal{F}^{aw}]_3 - \partial_0 V^2 [\mathcal{F}^{ar}]_1 + \partial_0 V^1 [\mathcal{F}^{ar}]_2 + \sum_{\gamma=0}^{3} V^\gamma [\partial_\gamma \mathcal{F}^w]_3$

$[\pounds_V \mathcal{F}]_{32}^a = \partial_3 V^0 [\mathcal{F}^{aw}]_2 - \partial_2 V^0 [\mathcal{F}^{aw}]_3 + \partial_3 V^3 [\mathcal{F}^{ar}]_1 + \partial_2 V^2 [\mathcal{F}^{ar}]_1 - \partial_2 V^1 [\mathcal{F}^{ar}]_2 - \partial_3 V^1 [\mathcal{F}^{ar}]_3 + \sum_{\gamma=0}^{3} V^\gamma [\partial_\gamma \mathcal{F}^r]_1$

$[\pounds_V \mathcal{F}]_{13}^a = -\partial_3 V^0 [\mathcal{F}^{aw}]_1 + \partial_1 V^0 [\mathcal{F}^{aw}]_3 - \partial_1 V^2 [\mathcal{F}^{ar}]_1 + \partial_1 V^1 [\mathcal{F}^{ar}]_2 + \partial_3 V^3 [\mathcal{F}^{ar}]_2 - \partial_3 V^2 [\mathcal{F}^{ar}]_3 + \sum_{\gamma=0}^{3} V^\gamma [\partial_\gamma \mathcal{F}^r]_2$

$[\pounds_V \mathcal{F}]_{21}^a = \partial_2 V^0 [\mathcal{F}^{aw}]_1 - \partial_1 V^0 [\mathcal{F}^{aw}]_2 - \partial_1 V^3 [\mathcal{F}^{ar}]_1 - \partial_2 V^3 [\mathcal{F}^{ar}]_2 + \partial_2 V^2 [\mathcal{F}^{ar}]_3 + \partial_1 V^1 [\mathcal{F}^{ar}]_3 + \sum_{\gamma=0}^{3} V^\gamma [\partial_\gamma \mathcal{F}^r]_3$

$[(\pounds_V \mathcal{F}(m))^r] = \sum_{\gamma=0}^3 V^\gamma [\partial_\gamma \mathcal{F}^r] +$
$[\mathcal{F}^{aw}] \begin{bmatrix} 0 & -\partial_3 V^0 & \partial_2 V^0 \\ \partial_3 V^0 & 0 & -\partial_1 V^0 \\ -\partial_2 V^0 & \partial_1 V^0 & 0 \end{bmatrix} - [\mathcal{F}^{ar}] \begin{bmatrix} [\partial_1 v]^1 & [\partial_1 v]^2 & [\partial_1 v]^3 \\ [\partial_2 v]^1 & +[\partial_2 v]^2 & [\partial_2 v]^3 \\ [\partial_3 v]^1 & [\partial_3 v]^2 & [\partial_3 v]^3 \end{bmatrix}$
$+ [\mathcal{F}^{ar}] \left([\partial_3 v]^3 + [\partial_2 v]^2 + [\partial_1 v]^1 \right) I_3$

$$[(\pounds_V \mathcal{F})^r] = \sum_{\gamma=0}^3 V^\gamma [\partial_\gamma \mathcal{F}^r] + [\mathcal{F}^{aw}] j \left(\partial V^0\right) + [\mathcal{F}^{ar}] \left(-[\partial v]^t + (div(v)) I_3 \right) \tag{A.9}$$

$[(\pounds_V \mathcal{F}(m))^w] = \sum_{\gamma=0}^3 V^\gamma [\partial_\gamma \mathcal{F}^w] + \partial_0 V^0 [\mathcal{F}^{aw}] + [\mathcal{F}^{aw}] [\partial v]$
$+ [\mathcal{F}^{ar}] \begin{bmatrix} 0 & [\partial_0 V]^3 & -[\partial_0 V]^2 \\ -[\partial_0 V]^3 & 0 & [\partial_0 V]^1 \\ [\partial_0 V]^2 & -[\partial_0 V]^1 & 0 \end{bmatrix}$

$$[(\pounds_V \mathcal{F})^w] = \sum_{\gamma=0}^3 V^\gamma [\partial_\gamma \mathcal{F}^w] + [\mathcal{F}^{aw}] \left(\partial_0 V^0 + [\partial v]\right) - [\mathcal{F}^{ar}] j (\partial_0 V) \tag{A.10}$$

with :
$[\partial v] = \begin{bmatrix} \partial_1 V^1 & \partial_2 V^1 & \partial_3 V^1 \\ \partial_1 V^2 & \partial_2 V^2 & \partial_3 V^2 \\ \partial_1 V^3 & \partial_2 V^3 & \partial_3 V^3 \end{bmatrix}$
$[\partial_0 V] = \begin{bmatrix} \partial_0 V^1 & \partial_0 V^2 & \partial_0 V^3 \end{bmatrix}$
$\left[\partial V^0\right] = \begin{bmatrix} \partial_1 V^0 \\ \partial_2 V^0 \\ \partial_3 V^0 \end{bmatrix}$
$[(\pounds_V \mathcal{F})^r] = \sum_{\gamma=0}^3 V^\gamma [\partial_\gamma \mathcal{F}^r] + [\mathcal{F}^{aw}] j \left(\partial V^0\right) + [\mathcal{F}^{ar}] \left(-[\partial v]^t + (div(v)) I_3 \right)$
$[(\pounds_V \mathcal{F})^w] = \sum_{\gamma=0}^3 V^\gamma [\partial_\gamma \mathcal{F}^w] + [\mathcal{F}^{aw}] \left(\partial_0 V^0 + [\partial v]\right) - [\mathcal{F}^{ar}] j (\partial_0 V)$
$\begin{bmatrix} [(\pounds_V \mathcal{F})^r] & [(\pounds_V \mathcal{F})^w] \end{bmatrix} =$
$\sum_{\gamma=0}^3 V^\gamma \begin{bmatrix} [\partial_\gamma \mathcal{F}^r] & [\partial_\gamma \mathcal{F}^w] \end{bmatrix} + \begin{bmatrix} [\mathcal{F}^r] & [\mathcal{F}^w] \end{bmatrix} \begin{bmatrix} \left(-[\partial v]^t + (div(v)) I_3 \right) & -j(\partial_0 V) \\ j\left(\partial V^0\right) & \left(\partial_0 V^0 + [\partial v]\right) \end{bmatrix}$

A.3 INVERSE

$M = \begin{bmatrix} A & -B \\ B & A \end{bmatrix}$

let be $M' = \begin{bmatrix} A^{-1} & B^{-1} \\ -B^{-1} & A^{-1} \end{bmatrix}$

$N = M'M = \begin{bmatrix} A^{-1} & B^{-1} \\ -B^{-1} & A^{-1} \end{bmatrix} \begin{bmatrix} A & -B \\ B & A \end{bmatrix} = \begin{bmatrix} 2I & -\left(A^{-1}B - B^{-1}A\right) \\ \left(A^{-1}B - B^{-1}A\right) & 2I \end{bmatrix}$

$N' = \begin{bmatrix} 2I & \left(A^{-1}B - B^{-1}A\right) \\ -\left(A^{-1}B - B^{-1}A\right) & 2I \end{bmatrix}$

$N'N = \begin{bmatrix} 2I & \left(A^{-1}B - B^{-1}A\right) \\ -\left(A^{-1}B - B^{-1}A\right) & 2I \end{bmatrix} \begin{bmatrix} 2I & -\left(A^{-1}B - B^{-1}A\right) \\ \left(A^{-1}B - B^{-1}A\right) & 2I \end{bmatrix}$

$= \begin{bmatrix} 4I + \left(A^{-1}B - B^{-1}A\right)^2 & 0 \\ 0 & \left(A^{-1}B - B^{-1}A\right)^2 + 4I \end{bmatrix} = \begin{bmatrix} R & 0 \\ 0 & R \end{bmatrix}$

$$(N'N)^{-1} = \begin{bmatrix} R^{-1} & 0 \\ 0 & R^{-1} \end{bmatrix}$$
$$(N'N)^{-1}(N'N) = I = (N'N)^{-1} N'M'M$$
$$\Rightarrow M^{-1} = (N'N)^{-1} N'M'$$
$$\begin{bmatrix} A & -B \\ B & A \end{bmatrix}^{-1} = \begin{bmatrix} R^{-1}\left(A^{-1} + B^{-1}AB^{-1}\right) & R^{-1}\left(B^{-1} + A^{-1}BA^{-1}\right) \\ -R^{-1}\left(B^{-1} + A^{-1}BA^{-1}\right) & R^{-1}\left(A^{-1} + B^{-1}AB^{-1}\right) \end{bmatrix}$$

A.4 FORMULAS

A.4.1 ALGEBRA

Operator j

Let $r \in \mathbb{C}^3, w \in \mathbb{C}^3$:

$$[j(r)]w = \begin{bmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} r_2 w_3 - r_3 w_2 \\ -r_1 w_3 + r_3 w_1 \\ r_1 w_2 - r_2 w_1 \end{bmatrix}$$

$$w[j(r)] = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{bmatrix} = \begin{bmatrix} -r_2 w_3 + r_3 w_2 & r_1 w_3 - r_3 w_1 & -r_1 w_2 + r_2 w_1 \end{bmatrix}$$

$$[j(r)]^{\alpha}_{\beta} = -\epsilon(\alpha, \beta, \gamma) r_{\gamma}$$

$$[j(r)w]^a = \sum_{b,c=1}^3 \epsilon(a,b,c) r_b w_c$$

$$[wj(r)]_a = -\sum_{b,c=1}^3 \epsilon(a,b,c) r_b w_c$$

$$j(r)^t = -j(r) = j(-r)$$

$$j(x)y = -j(y)x$$

$$y^t j(x) = -x^t j(y)$$

$$j(x)y = 0 \Leftrightarrow \exists k \in \mathbb{R} : y = kx$$

$$Tr(j(x)j(y)) = -2x^t y$$

$$([j(r)]w)^t ([j(r)]w) = (r^t r)(w^t w) - (r^t w)^2$$

eigenvectors of j(r)

$r = \sqrt{r^t r}$

$$0 : \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}$$

$$ir : \begin{bmatrix} -(-r_1 r_2 + i r_3 r) \\ -\left(r_1^2 + r_3^2\right) \\ r_2 r_3 + i r_1 r \end{bmatrix}$$

$$-ir : \begin{bmatrix} -(r_1 r_2 + i r_3 r) \\ \left(r_1^2 + r_3^2\right) \\ -r_2 r_3 + r_1 i r \end{bmatrix}$$

Identities

$$j(x)j(y) = yx^t - (y^t x) I$$

$$j(x)j(x)j(x) = j(x)(xx^t - (x^t x) I) = -(x^t x) j(x)$$

$$j(j(x)y) = yx^t - xy^t = j(x)j(y) - j(y)j(x)$$

$$j(j(x)j(x)y) = (y^t x) j(x) - (x^t x) j(y)$$

$$j(x)j(y)j(x) = -(y^t x) j(x)$$

$$j(x)j(x)j(y) = j(x)yx^t - (y^t x) j(x) = -j(y)xx^t - (y^t x) j(x) = -j(y)(j(x)j(x) + x^t x) - (y^t x) j(x)$$

$$x^t j(r) j(s) y = x^t (sr^t - r^t s I) y = (x^t s)(r^t y) - (x^t y)(r^t s)$$

$[M], [X] \in L(3)$:

$$([M]_1)^t j([M]_2) [M]_3 = \det M$$

$$M^t j(Mx) M = (\det M) j(x)$$

$$j(Mx) = [M^t]^{-1} j(x) [M]^{-1} \det M$$

$$\left[\begin{array}{ccc} j\left([M]_2\right)[M]_3 & j\left([M]_3\right)[M]_1 & j\left([M]_1\right)[M]_2 \end{array}\right] = (\det M)\left[M^{-1}\right]^t$$
$$\left([X]_1\right)^t [M]^t j([M][X]_2)[M][X]_3 = (\det [M])(\det [X])$$
$$M \in O(3) : j(Mx)My = Mj(x)y \Leftrightarrow Mx \times My = M(x \times y)$$
$$k > 0 : j(r)^{2k} = \left(-r^t r\right)^{k-1}\left(rr^t - \left(r^t r\right) I\right) = \left(-r^t r\right)^{k-1} j(r)j(r)$$
$$k \geq 0 : J(r)^{2k+1} = \left(-r^t r\right)^k j(r)$$
$$\left[\begin{array}{ccc} 0 & j(z) & -j(y) \\ -j(z) & 0 & j(x) \\ j(y) & -j(x) & 0 \end{array}\right]^{-1} = \frac{1}{2x^t j(y) z}\left[\begin{array}{ccc} xx^t & 2yx^t - xy^t & 2zx^t - xz^t \\ 2xy^t - yx^t & yy^t & 2zy^t - yz^t \\ 2xz^t - zx^t & 2yz^t - zy^t & zz^t \end{array}\right]$$
$$= \frac{1}{2x^t j(y) z}\left(2\left[\begin{array}{ccc} xx^t & yx^t & zx^t \\ xy^t & yy^t & zy^t \\ xz^t & yz^t & zz^t \end{array}\right] - \left[\begin{array}{c} x \\ y \\ z \end{array}\right]_{9\times 1}\left[\begin{array}{ccc} x^t & y^t & z^t \end{array}\right]_{1\times 9}\right)$$
$$x^t j(y) z = -\det\left[\begin{array}{ccc} x_1 & y_2 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{array}\right]$$

Polynomials

The set of polynomials of matrices $P(z) = aI + bj(z) + cj(z)j(z)$ where $z \in \mathbb{C}^3$ is fixed, $a, b, c \in \mathbb{C}$ is a commutative ring.

$$P(z) = \left[\begin{array}{ccc} a - c\left(\left(z^t z\right) - z_1^2\right) & -br_3 + cz_1 z_2 & bz_2 + cz_1 z_3 \\ bz_3 + cz_1 z_2 & a - c\left(\left(z^t z\right) - z_2^2\right) & -bz_1 + cz_2 z_3 \\ -bz_2 + cz_1 z_3 & bz_1 + cz_2 z_3 & a - c\left(\left(z^t z\right) - z_3^2\right) \end{array}\right]$$
$$\left(a + bj(z) + cj(z)j(z)\right)\left(a' + b'j(z) + c'j(z)j(z)\right)$$
$$= aa' + \left(ab' + a'b - \left(z^t z\right)\left(c'b + b'c\right)\right) j(z) + \left(ac' + a'c + b'b - \left(z^t z\right) c'c\right) j(z)j(z)$$
$$\det\left(a + bj(z) + cj(z)j(z)\right) = a\left(a^2 + \left(b^2 + c^2\left(z^t z\right) - 2ac\right)\left(z^t z\right)\right)$$
$$\left[a + bj(z) + cj(z)j(z)\right]^{-1} = \left[\frac{1}{a}I - \frac{ab}{\det P} j(r) - \frac{\left(ac - b^2 - c^2\left(z^t z\right)\right)}{\det P} j(z)j(z)\right]$$
$$\det\left(a + bj(z)\right) = a\left(a^2 + b^2\left(z^t z\right)\right)$$
$$\left[a + bj(z)\right]^{-1} = \left[\frac{1}{a}I - \frac{ab}{\det P} j(r) + \frac{b^2}{a\left(a^2 + b^2\left(z^t z\right)\right)} j(z)j(z)\right]$$
$$\left[a + cj(z)j(z)\right]^{-1} = \left[\frac{1}{a}I - \frac{\left(ac - c^2\left(z^t z\right)\right)}{a\left(a^2 + \left(c^2\left(z^t z\right) - 2ac\right)\left(z^t z\right)\right)} j(z)j(z)\right]$$

eigenvectors of $P(r)$: the only real eigen value is a with eigen vector r

Matrices of SO(3,1)

signature $(3,1) : \langle u, v\rangle = u^1 v^1 + u^2 v^2 + u^3 v^3 - u^0 v^0$

signature $(1,3) : \langle u, v\rangle = -u^1 v^1 - u^2 v^2 - u^3 v^3 + u^0 v^0$

$$[\kappa]^t [\eta][\kappa] = [\eta] \Leftrightarrow [\kappa] \in SO(3,1) \equiv SO(1,3)$$
$$[\chi] = \exp[K(w)] \exp[J(r)]$$
$$\exp[K(w)] = I_4 + \frac{\sinh\sqrt{w^t w}}{\sqrt{w^t w}} K(w) + \frac{\cosh\sqrt{w^t w} - 1}{w^t w} K(w)K(w)$$
$$[\eta] = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Basis of $so(3,1) \equiv so(1,3)$

$$[\kappa_1] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}; [\kappa_2] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}; [\kappa_3] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[\kappa_4] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; [\kappa_5] = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; [\kappa_6] = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$[\kappa] = [J(r)] + [K(w)] \in so(3,1)$ with

$$[J(r)] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -r_3 & r_2 \\ 0 & r_3 & 0 & -r_1 \\ 0 & -r_2 & r_1 & 0 \end{bmatrix}; [K(w)] = \begin{bmatrix} 0 & w_1 & w_2 & w_3 \\ w_1 & 0 & 0 & 0 \\ w_2 & 0 & 0 & 0 \\ w_3 & 0 & 0 & 0 \end{bmatrix}$$

Dirac's matrices

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}; \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}; \sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$\sigma^* = \sigma$

$\sigma_i \sigma_j + \sigma_j \sigma_i = 2\delta_{ij} \sigma_0$

$\sigma_i^2 = \sigma_0$

$j \neq k, l = 1,2,3 : \sigma_j \sigma_k = \epsilon(j,k,l) i\sigma_l$

$\sigma_1 \sigma_2 \sigma_3 = i\sigma_0$

$\left(u^b\right)_{b=1}^3 \in \mathbb{C}^3 : \sigma(u) = \sum_{b=1}^3 u^b \sigma_b,$

$$\sigma(u) = \begin{bmatrix} u_3 & u_1 - iu_2 \\ u_1 + iu_2 & -u_3 \end{bmatrix}$$

then :

$\sigma(u) \times \sigma(v) = i\sigma(j(u)v) + (u^t v)\sigma_0$

$\sigma(u)^{-1} = \frac{1}{u^t u} \sigma(u)$

$(\sigma(z))^* = \sigma(\overline{z})$

$\sigma(z)\sigma(z') = \sigma(j(z)z') + z^t z' \sigma_0$

γ matrices

$$\gamma_0 = \begin{bmatrix} 0 & -i\sigma_0 \\ i\sigma_0 & 0 \end{bmatrix}; \gamma_1 = \begin{bmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{bmatrix}; \gamma_2 = \begin{bmatrix} 0 & \sigma_2 \\ \sigma_2 & 0 \end{bmatrix}; \gamma_3 = \begin{bmatrix} 0 & \sigma_3 \\ \sigma_3 & 0 \end{bmatrix};$$

$\gamma_i \gamma_j + \gamma_j \gamma_i = 2\delta_{ij} I_4$

$\gamma_i = \gamma_i^* = \gamma_i^{-1}$

$$j = 1,2,3 : \gamma_j \gamma_0 = -\gamma_0 \gamma_j = i \begin{bmatrix} \sigma_j & 0 \\ 0 & -\sigma_j \end{bmatrix} = i\gamma_5 \widetilde{\gamma}_j$$

$$\gamma_1 \gamma_2 = -\gamma_2 \gamma_1 = i \begin{bmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{bmatrix}; \gamma_2 \gamma_3 = -\gamma_3 \gamma_2 = i \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_1 \end{bmatrix};$$

$$\gamma_3 \gamma_1 = -\gamma_1 \gamma_3 = i \begin{bmatrix} \sigma_2 & 0 \\ 0 & \sigma_2 \end{bmatrix}$$

$$\gamma_1 \gamma_2 \gamma_3 = i \begin{bmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{bmatrix}$$

$$\gamma_5 = \gamma_0 \gamma_1 \gamma_2 \gamma_3 = \begin{bmatrix} \sigma_0 & 0 \\ 0 & -\sigma_0 \end{bmatrix}$$

$\gamma_5 \gamma_a = -\gamma_a \gamma_5$

$j=1,2,3:\widetilde{\gamma}_j=\begin{bmatrix}\sigma_j & 0\\ 0 & \sigma_j\end{bmatrix}$
$j\neq k,l=1,2,3:\gamma_j\gamma_k=-\gamma_k\gamma_j=i\epsilon\left(j,k,l\right)\widetilde{\gamma}_l$
$\widetilde{\gamma}_a\widetilde{\gamma}_a=I_4$
$\widetilde{\gamma}_a\widetilde{\gamma}_b=i\epsilon\left(a,b,c\right)\widetilde{\gamma}_c$
$\widetilde{\gamma}_a\widetilde{\gamma}_b+\widetilde{\gamma}_b\widetilde{\gamma}_a=2\delta_{ab}I_4$
$\gamma_a\widetilde{\gamma}_b=i\epsilon\left(a,b,c\right)\gamma_c$
$\gamma_a\widetilde{\gamma}_a=\begin{bmatrix}0 & \sigma_0\\ \sigma_0 & 0\end{bmatrix}$
$\gamma_0\widetilde{\gamma}_a=\begin{bmatrix}0 & -i\sigma_a\\ i\sigma_a & 0\end{bmatrix}=\left(\gamma_0\widetilde{\gamma}_a\right)^*=\begin{bmatrix}0 & -i\sigma_a^*\\ i\sigma_a^* & 0\end{bmatrix}$
$Cl\left(3,1\right):\gamma C\left(\varepsilon_j\right)=\gamma_j,j=1,2,3;\gamma C\left(\varepsilon_0\right)=i\gamma_0;\gamma C\left(\varepsilon_5\right)=i\gamma_5$
$a=1,2,3:\gamma C\left(\overrightarrow{\kappa}_a\right)=-\frac{1}{2}i\widetilde{\gamma}_a$
$a=4,5,6:\gamma C\left(\overrightarrow{\kappa}_a\right)=\frac{1}{2}\begin{bmatrix}\sigma_a & 0\\ 0 & -\sigma_a\end{bmatrix}=-\frac{1}{2i}\gamma_0\gamma_j$
$\gamma C\left(\upsilon\left(r,w\right)\right)=-i\frac{1}{2}\sum_{a=1}^3\left(w_a\gamma_a\gamma_0+r_a\widetilde{\gamma}_a\right)$
$=\frac{1}{2}\begin{bmatrix}\sigma\left(w-ir\right) & 0\\ 0 & -\sigma\left(w+ir\right)\end{bmatrix}$
$\gamma C\left(a+\upsilon\left(r,w\right)+b\varepsilon_5\right)=aI+ib\gamma_5-i\frac{1}{2}\sum_{a=1}^3\left(w_a\gamma_a\gamma_0+r_a\widetilde{\gamma}_a\right)$
$Cl\left(1,3\right):\gamma C'\left(\varepsilon_j\right)=i\gamma_j,j=1,2,3;\gamma C'\left(\varepsilon_j\right)=\gamma_0;\gamma C'\left(\varepsilon_5\right)=\gamma_5$
In $Cl(1,3)$:
$\gamma C'\left(\upsilon\left(r,w\right)\right)=-i\frac{1}{2}\sum_{a=1}^3\left(w_a\gamma_a\gamma_0-r_a\widetilde{\gamma}_a\right)$

A.4.2 CLIFFORD ALGEBRA

$\varepsilon_i\cdot\varepsilon_j+\varepsilon_j\cdot\varepsilon_i=2\eta_{ij}$
$\varepsilon_5\cdot\varepsilon_5=-1$
$X\cdot\varepsilon_5+\varepsilon_5\cdot X=0$
$\forall s\in Spin\left(3,1\right):\mathbf{Ad}_s\varepsilon_5=\varepsilon_5$

Adjoint map

$\forall X\in Cl\left(3,1\right),s\in Spin\left(3,1\right):\mathbf{Ad}_sX=s\cdot X\cdot s^{-1}$
$\left\langle\mathbf{Ad}_sX,\mathbf{Ad}_sY\right\rangle=\left\langle X,Y\right\rangle$
$\mathbf{Ad}_s\circ\mathbf{Ad}_{s'}=\mathbf{Ad}_{s\cdot s'}$

Action of the Adjoint map on vectors :

$v=\sum_{i=0}^3v^i\varepsilon_i\to\widetilde{v}=\mathbf{Ad}_sv=\sum_{i=0}^3v^is\cdot\varepsilon_i\cdot s^{-1}=\sum_{i=0}^3\widetilde{v}^i\varepsilon_i$
$\widetilde{v}^i=\sum_{j=0}^3\left[h\left(s\right)\right]_j^iv^j$
$\left[h\left(s\right)\right]$
$=\begin{bmatrix}a^2+b^2+\frac{1}{4}\left(r^tr+w^tw\right) & aw^t-br^t+\frac{1}{2}w^tj\left(r\right)\\ aw-br+\frac{1}{2}j\left(r\right)w & a^2+b^2+\frac{1}{4}\left(r^tr+w^tw\right)+aj\left(r\right)+bj\left(w\right)+\frac{1}{2}\left(j\left(r\right)j\left(r\right)+j\left(w\right)j\left(w\right)\right)\end{bmatrix}$
$\left[h(s)\right]^t\left[\eta\right]\left[h\left(s\right)\right]=\left[\eta\right]$
If $s=a_w+\upsilon\left(0,w\right)$
$\left[h(s)\right]=\begin{bmatrix}2a_w^2-1 & a_ww^t\\ a_ww & 2a_w^2-1+\frac{1}{2}j\left(w\right)j\left(w\right)\end{bmatrix}$
If $s=a_r+\upsilon\left(r,0\right)$
$\left[h(s)\right]=\begin{bmatrix}1 & 0\\ 0 & 1+a_rj\left(r\right)+\frac{1}{2}j\left(r\right)j\left(r\right)\end{bmatrix}$

$[C(r)] = 1 + a_r j(r) + \frac{1}{2} j(r) j(r) \in SO(3)$

Action of the adjoint map on the Lie algebra:

$Z = \sum_{a=1}^{6} Z_a \overrightarrow{\kappa}_a \rightarrow \widetilde{Z} = \sum_{a=1}^{6} Z_a \mathbf{Ad}_s(\overrightarrow{\kappa}_a) = \sum_{a=1}^{6} Z_a s \cdot (\overrightarrow{\kappa}_a) \cdot s^{-1} = \sum_{a=1}^{6} Z_a \widetilde{\overrightarrow{\kappa}_a} = \sum_{a=1}^{6} \widetilde{Z}_a \overrightarrow{\kappa}_a$

With :

$Z = \upsilon(X,Y) \rightarrow \widetilde{Z} = \upsilon\left(\widetilde{X}, \widetilde{Y}\right)$

$\begin{bmatrix} \widetilde{X} \\ \widetilde{Y} \end{bmatrix} = [\mathbf{Ad}_s] \begin{bmatrix} X \\ Y \end{bmatrix}$

$[\mathbf{Ad}_s]$

$= \begin{bmatrix} 1 + aj(r) - bj(w) + \frac{1}{2}(j(r)j(r) - j(w)j(w)) & -\left(aj(w) + bj(r) + \frac{1}{2}(j(r)j(w) + j(w)j(r))\right) \\ aj(w) + bj(r) + \frac{1}{2}(j(r)j(w) + j(w)j(r)) & 1 + aj(r) - bj(w) + \frac{1}{2}(j(r)j(r) - j(w)j(w)) \end{bmatrix}$

With $s_w = a_w + \upsilon(0,w)$

$[\mathbf{Ad}_s] = \begin{bmatrix} \left[1 - \frac{1}{2} j(w) j(w)\right] & -[a_w j(w)] \\ [a_w j(w)] & \left[1 - \frac{1}{2} j(w) j(w)\right] \end{bmatrix} = \begin{bmatrix} A & -B \\ B & A \end{bmatrix}$

and the identities :

$A = A^t, B^t = -B, AB = BA$

$A^2 + B^2 = I$

$[\mathbf{Ad}_{s_w}]^{-1} = \left[\mathbf{Ad}_{s_w^{-1}}\right] = \begin{bmatrix} A & B \\ -B & A \end{bmatrix}$

With $s_r = a_r + \upsilon(r,0)$

$[\mathbf{Ad}_s] = \begin{bmatrix} \left[1 + a_r j(r) + \frac{1}{2} j(r) j(r)\right] & 0 \\ 0 & \left[1 + a_r j(r) + \frac{1}{2} j(r) j(r)\right] \end{bmatrix} = \begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix}$

and the identities :

$CC^t = C^t C = I_3$

$[\mathbf{Ad}_{s_r}]^{-1} = \left[\mathbf{Ad}_{s_r^{-1}}\right] = \begin{bmatrix} C^t & 0 \\ 0 & C^t \end{bmatrix}$

Lie Algebras

$\upsilon(r,w) = \frac{1}{2}\left(w^1 \varepsilon_0 \cdot \varepsilon_1 + w^2 \varepsilon_0 \cdot \varepsilon_2 + w^3 \varepsilon_0 \cdot \varepsilon_3 + r^3 \varepsilon_2 \cdot \varepsilon_1 + r^2 \varepsilon_1 \cdot \varepsilon_3 + r^1 \varepsilon_3 \cdot \varepsilon_2\right)$

$\overrightarrow{\kappa}_1 = \upsilon((1,0,0),(0,0,0)) = \frac{1}{2}\varepsilon_3 \cdot \varepsilon_2,$

$\overrightarrow{\kappa}_2 = \upsilon((0,1,0),(0,0,0)) = \frac{1}{2}\varepsilon_1 \cdot \varepsilon_3,$

$\overrightarrow{\kappa}_3 = \upsilon((0,0,1),(0,0,0)) = \frac{1}{2}\varepsilon_2 \cdot \varepsilon_1,$

$\overrightarrow{\kappa}_4 = \upsilon((0,0,0),(1,0,0)) = \frac{1}{2}\varepsilon_0 \cdot \varepsilon_1,$

$\overrightarrow{\kappa}_5 = \upsilon((0,0,0),(0,1,0)) = \frac{1}{2}\varepsilon_0 \cdot \varepsilon_2,$

$\overrightarrow{\kappa}_6 = \upsilon((0,0,0),(0,0,1)) = \frac{1}{2}\varepsilon_0 \cdot \varepsilon_3$

$\upsilon(r,w) \cdot \varepsilon_0 = \frac{1}{2}\left(w^1 \varepsilon_1 + w^2 \varepsilon_2 + w^3 \varepsilon_3 + r^3 \varepsilon_2 \cdot \varepsilon_1 \cdot \varepsilon_0 + r^2 \varepsilon_1 \cdot \varepsilon_3 \cdot \varepsilon_0 + r^1 \varepsilon_3 \cdot \varepsilon_2 \cdot \varepsilon_0\right)$

$\varepsilon_0 \cdot \upsilon(r,w) = \frac{1}{2}\left(-w^1 \varepsilon_1 - w^2 \varepsilon_2 + w^3 \varepsilon_3 + r^3 \varepsilon_2 \cdot \varepsilon_1 \cdot \varepsilon_0 + r^2 \varepsilon_1 \cdot \varepsilon_3 \cdot \varepsilon_0 + r^1 \varepsilon_3 \cdot \varepsilon_2 \cdot \varepsilon_0\right)$

$\upsilon(r,w) \cdot \varepsilon_0 - \varepsilon_0 \cdot \upsilon(r,w) = w$

$\upsilon(r,w) \cdot \varepsilon_0 + \varepsilon_0 \cdot \upsilon(r,w) = 2\upsilon(r,0) \cdot \varepsilon_0$

$\upsilon(r,w) \cdot \varepsilon_5 = \varepsilon_5 \cdot \upsilon(r,w) = \upsilon(-w,r)$

In $Cl(3,1)$:

$\upsilon(r',w') \cdot \upsilon(r,w)$

$= \frac{1}{4}(w^t w' - r^t r') + \frac{1}{2}\upsilon(-j(r)r' + j(w)w', -j(w)r' - j(r)w') - \frac{1}{4}(w^t r' + r^t w')\varepsilon_5$

$[\upsilon(r,w), \upsilon(r',w')] = \upsilon(j(r)r' - j(w)w', j(w)r' + j(r)w')$

$\upsilon(r',w') \cdot \upsilon(r,w) = -\frac{1}{2}[\upsilon(r,w), \upsilon(r',w')] + \frac{1}{4}(w^t w' - r^t r') - \frac{1}{4}(w^t r' + r^t w')\varepsilon_5$

In $Cl(1,3)$:

$\upsilon(r,w)\cdot\upsilon(r',w')$
$=\frac{1}{4}(w^tw'-r^tr')-\frac{1}{2}\upsilon(-j(r)r'+j(w)w',j(w)r'+j(r)w')-\frac{1}{4}(w^tr'+r^tw')\varepsilon_5$
$[\upsilon(r,w),\upsilon(r',w')]=-\upsilon(j(r)r'-j(w)w',j(w)r'+j(r)w')$
Scalar product :
$\langle\upsilon(r,w),\upsilon(r',w')\rangle_{Cl}=\frac{1}{4}(r^tr'-w^tw')$

Spin groups

$s=a+\upsilon(r,w)+b\varepsilon_5$
$a^2-b^2=1+\frac{1}{4}(w^tw-r^tr)$
$ab=-\frac{1}{4}r^tw$
if $r=0$ then $a=\epsilon\sqrt{1+\frac{1}{4}w^tw};b=0$
if $w=0$ then
$r^tr\leq 4:a=\epsilon\sqrt{1-\frac{1}{4}r^tr};b=0$
$r^tr\geq 4:b=\epsilon\sqrt{-1+\frac{1}{4}r^tr};a=0$

Product :
$(a+\upsilon(r,w)+b\varepsilon_5)^{-1}=a-\upsilon(r,w)+b\varepsilon_5$
$s\cdot s'=a"+\upsilon(r",w")+b"\varepsilon_0\cdot\varepsilon_1\cdot\varepsilon_2\cdot\varepsilon_3$
with :
$a"=aa'-b'b+\frac{1}{4}(w^tw'-r^tr')$
$b"=ab'+ba'-\frac{1}{4}(w^tr'+r^tw')$

i) In $Cl(3,1)$:
$r"=\frac{1}{2}(j(r)r'-j(w)w')+a'r+ar'-b'w-bw'$
$w"=\frac{1}{2}(j(w)r'+j(r)w')+a'w+aw'+b'r+br'$
$(a+\upsilon(0,w))\cdot(a'+\upsilon(0,w'))=aa'+\frac{1}{4}w^tw'+\upsilon\left(-\frac{1}{2}(j(w)w',a'w+aw')\right)$
$(a+\upsilon(r,0))\cdot(a'+\upsilon(r',0))=aa'-\frac{1}{4}r^tr'+\upsilon\left(\frac{1}{2}j(r)r'+(a'r+ar'),0\right)$
$(a_w+\upsilon(0,w))\cdot(a_r+\upsilon(r,0))=a_wa_r+\upsilon\left(a_wr,\frac{1}{2}j(w)r+a_rw\right)-\frac{1}{4}(w^tr)\varepsilon_5$

ii) In $Cl(1,3)$:
$r"=\frac{1}{2}(j(r)r'-j(w)w')+a'r+ar'+b'w+bw'$
$w"=-\frac{1}{2}(j(w)r'+j(r)w')+a'w+aw'+b'r+br'$

A.4.3 DIFFERENTIAL GEOMETRY

Pull back, push forward

M,N manifolds, $f\in C_1(M;N)$
$Tf:TM\to TN::Tf(m,u_m)=(f(m),f'(m)u_m)$
push forward of a vector field :
$f_*:V\in TM\to f_*V\in TN::f_*V(f(m))=f'(m)V(m)\Leftrightarrow f_*V=Tf(V)$
pull back of a 1 form :
$f^*::\lambda\in TN^*\to f^*\lambda\in TM^*::f^*\lambda(m)(u_m)=\lambda(f(m))f'(m)u_m\Leftrightarrow f^*\lambda=\lambda(Tf)$
If f is a diffeomorphism :
pull back of a vector field :
$f^*::V\in TN\to f^*V\in TM::f^*V(m)=(f')^{-1}(n)V(n)\Leftrightarrow f^*V=(Tf)^{-1}(V)$
push forward of a 1 form :

$f_* : \lambda \in TM^* \to f_*\lambda \in TN^* :: f_*\lambda\left(f\left(m\right)\right)\left(u_{f(m)}\right) = \lambda\left(f^{-1}\left(n\right)\right)\left(f'\left(n\right)\right)^{-1}\left(u_n\right) \Leftrightarrow f_*\lambda = \lambda\left(\left(Tf\right)^{-1}\right)$

$f^* = \left(f_*\right)^{-1}$

r Forms

2 forms

$\mathcal{F} = \mathcal{F}^r + \mathcal{F}^w$

$\mathcal{F}^r = \mathcal{F}_{32}d\xi^3 \wedge d\xi^2 + \mathcal{F}_{13}d\xi^1 \wedge d\xi^3 + \mathcal{F}_{21}d\xi^2 \wedge d\xi^1$

$\mathcal{F}^w = \mathcal{F}_{01}d\xi^0 \wedge d\xi^1 + \mathcal{F}_{02}d\xi^0 \wedge d\xi^2 + \mathcal{F}_{03}d\xi^0 \wedge d\xi^3$

$$[\mathcal{F}]^{\alpha=0..3}_{\beta=0..3} = \begin{bmatrix} 0 & \mathcal{F}_{01} & \mathcal{F}_{02} & \mathcal{F}_{03} \\ -\mathcal{F}_{01} & 0 & -\mathcal{F}_{21} & \mathcal{F}_{13} \\ -\mathcal{F}_{02} & \mathcal{F}_{21} & 0 & -\mathcal{F}_{32} \\ -\mathcal{F}_{03} & -\mathcal{F}_{13} & \mathcal{F}_{32} & 0 \end{bmatrix} = \begin{bmatrix} 0 & [\mathcal{F}^w] \\ -[\mathcal{F}^w]^t & [j([\mathcal{F}^r])] \end{bmatrix}$$

$\mathcal{F} \wedge K = -\left([\mathcal{F}^r][K^w]^t + [\mathcal{F}^w][K^r]^t\right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$\mathcal{F}^w \wedge K^r = -\left([\mathcal{F}^r][K^w]^t\right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$i_V\mathcal{F} = [\mathcal{F}^w][v]\,d\xi^0 + [j(\mathcal{F}^r)v]^1 d\xi^1 + [j(\mathcal{F}^r)v]^2 d\xi^2 + [j(\mathcal{F}^r)v]^3 d\xi^3 - V^0\left(\mathcal{F}_{01}d\xi^1 + \mathcal{F}_{02}d\xi^2 + \mathcal{F}_{03}d\xi^3\right)$

$\mathcal{F}(V,W) = [\mathcal{F}^w]\left(W^0[v] - V^0[w]\right) + [w]^t j(\mathcal{F}^r)v$

$d(i_V\mathcal{F}) =$

$+\partial_0\left(\mathcal{F}_{13}V^3 - \mathcal{F}_{21}V^2 - \mathcal{F}_{01}V^0\right) - \partial_1\left(\mathcal{F}_{01}V^1 + \mathcal{F}_{02}V^2 + \mathcal{F}_{03}V^3\right) d\xi^0 \wedge d\xi^1$

$+\partial_0\left(\mathcal{F}_{21}V^1 - \mathcal{F}_{32}V^3 - \mathcal{F}_{02}V^0\right) - \partial_2\left(\mathcal{F}_{01}V^1 + \mathcal{F}_{02}V^2 + \mathcal{F}_{03}V^3\right) d\xi^0 \wedge d\xi^2$

$+\partial_0\left(\mathcal{F}_{32}V^2 - \mathcal{F}_{13}V^1 - \mathcal{F}_{03}V^0\right) - \partial_3\left(\mathcal{F}_{01}V^1 + \mathcal{F}_{02}V^2 + \mathcal{F}_{03}V^3\right) d\xi^0 \wedge d\xi^3$

$+\partial_3\left(\mathcal{F}_{21}V^1 - \mathcal{F}_{32}V^3 - \mathcal{F}_{02}V^0\right) - \partial_2\left(\mathcal{F}_{32}V^2 - \mathcal{F}_{13}V^1 - \mathcal{F}_{03}V^0\right) d\xi^3 \wedge d\xi^2$

$+\partial_1\left(\mathcal{F}_{32}V^2 - \mathcal{F}_{13}V^1 - \mathcal{F}_{03}V^0\right) - \partial_3\left(\mathcal{F}_{13}V^3 - \mathcal{F}_{21}V^2 - \mathcal{F}_{01}V^0\right) d\xi^1 \wedge d\xi^3$

$+\partial_2\left(\mathcal{F}_{13}V^3 - \mathcal{F}_{21}V^2 - \mathcal{F}_{01}V^0\right) - \partial_1\left(\mathcal{F}_{21}V^1 - \mathcal{F}_{32}V^3 - \mathcal{F}_{02}V^0\right) d\xi^2 \wedge d\xi^1$

Change of chart :

$\left[\widetilde{\mathcal{F}}^r\right] = [\mathcal{F}^r](\det k)\left[k^{-1}\right]^t + [\mathcal{F}^w][k]\,j([K_0])$

$\left[\widetilde{\mathcal{F}}^w\right] = [\mathcal{F}^w]\left(\left([K]^0_0\right)[k] - [K_0]\left[K^0\right]\right) - [\mathcal{F}^r]\,j([K_0])[k]$

Exterior differential of a 2 form

$d\{\mathcal{F}_{32}d\xi^3 \wedge d\xi^2 + \mathcal{F}_{13}d\xi^1 \wedge d\xi^3 + \mathcal{F}_{21}d\xi^2 \wedge d\xi^1 + \mathcal{F}_{01}d\xi^0 \wedge d\xi^1 + \mathcal{F}_{02}d\xi^0 \wedge d\xi^2 + \mathcal{F}_{03}d\xi^0 \wedge d\xi^3\}$

$= (-\partial_0\mathcal{F}_{21} + \partial_2\mathcal{F}_{01} - \partial_1\mathcal{F}_{02})\,d\xi^0 \wedge d\xi^1 \wedge d\xi^2 + (\partial_0\mathcal{F}_{13} + \partial_3\mathcal{F}_{01} - \partial_1\mathcal{F}_{03})\,d\xi^0 \wedge d\xi^1 \wedge d\xi^3$

$+ (-\partial_0\mathcal{F}_{32} + \partial_3\mathcal{F}_{02} - \partial_2\mathcal{F}_{03})\,d\xi^0 \wedge d\xi^2 \wedge d\xi^3 - (\partial_1\mathcal{F}_{32} + \partial_2\mathcal{F}_{13} + \partial_3\mathcal{F}_{21})\,d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

Scalar product of forms

$\forall\mu \in \Lambda_r(M) : *\mu \wedge \lambda = G_r(\mu,\lambda)\,\varpi_4 = *\lambda \wedge \mu$

Scalar product of 1-forms :

$G_1(\lambda,\mu) = \sum_{\alpha\beta=0}^3 g^{\alpha\beta}\lambda_\alpha\mu_\beta$

Scalar product of 2-forms :

$G_2(\mathcal{F},K) = -\frac{1}{\det P'}\left([*\mathcal{F}^w][K^r]^t + [*\mathcal{F}^r][K^w]^t\right) = \sum_{\{\alpha\beta\}}\mathcal{F}^{\alpha\beta}K_{\alpha\beta}$

$= (\det Q)^2[\mathcal{F}^r][g_3][K^r]^t - [\mathcal{F}^w]\left[g_3^{-1}\right][K^w]^t$

Hodge dual

$*\mathcal{F}^r = -\left(\mathcal{F}^{01}d\xi^3 \wedge d\xi^2 + \mathcal{F}^{02}d\xi^1 \wedge d\xi^3 + \mathcal{F}^{03}d\xi^2 \wedge d\xi^1\right)\det P'$

$*\mathcal{F}^w = -\left(\mathcal{F}^{32}d\xi^0 \wedge d\xi^1 + \mathcal{F}^{13}d\xi^0 \wedge d\xi^2 + \mathcal{F}^{21}d\xi^0 \wedge d\xi^3\right)\det P'$

$\mathcal{F}^{\alpha\beta} = \sum_{\lambda\mu=0}^3 g^{\alpha\lambda}g^{\beta\mu}\mathcal{F}_{\lambda\mu}$

$[*\mathcal{F}^r] = \left([\mathcal{F}^w]\left(-g^{00}\left[g_3^{-1}\right] + \left[g^{-1}\right]_0\left[g^{-1}\right]^0\right) + [\mathcal{F}^r]\, j\left(\left[g^{-1}\right]_0\right)\left[g_3^{-1}\right]\right)\det P'$

$[*\mathcal{F}^w] = -\left([\mathcal{F}^w]\left[g_3^{-1}\right] j\left(\left[g^{-1}\right]_0\right) + (\det Q)^2[\mathcal{F}^r][g_3]\right)\det P'$

In a standard basis :

$[*\mathcal{F}^r] = [\mathcal{F}^w]\left[g_3^{-1}\right]\det Q'$

$[*\mathcal{F}^w] = -[\mathcal{F}^r][g_3]\det Q$

$[*\mathcal{F}_r^r] = [\mathcal{F}_r^w]\left[g_3^{-1}\right]\det Q'$

$[*\mathcal{F}_w^r] = [\mathcal{F}_w^w]\left[g_3^{-1}\right]\det Q'$

$[*\mathcal{F}_r^w] = -[\mathcal{F}_r^r][g_3]\det Q$

$[*\mathcal{F}_w^w] = -[\mathcal{F}_w^r][g_3]\det Q$

$[*\mathcal{F}_A^r] = [\mathcal{F}_A^w]\left[g_3^{-1}\right]\det Q'$

$[*\mathcal{F}_A^w] = -[\mathcal{F}^r][g_3]\det Q$

$**\lambda_r = -(-1)^{r(n-r)}\lambda \Rightarrow **\lambda_2 = -\lambda_2$

Chern identity : $\langle\mathcal{F}_{01},\mathcal{F}_{32}\rangle + \langle\mathcal{F}_{02},\mathcal{F}_{13}\rangle + \langle\mathcal{F}_{03},\mathcal{F}_{21}\rangle = 0$

Lie derivative of 2 form

$£_V\mathcal{F} = \sum_{a=1}^m\left(\sum_{\{\alpha\beta\}}\sum_{\gamma=0}^3 V^\gamma\partial_\gamma\mathcal{F}_{\alpha\beta}^a d\xi^\alpha \wedge d\xi^\beta + \mathcal{F}_{\alpha\beta}^a\left(\partial_\gamma V^\alpha d\xi^\gamma \wedge d\xi^\beta + \partial_\gamma V^\beta d\xi^\alpha \wedge d\xi^\gamma\right)\right)\otimes \overrightarrow{\theta}_a$

$[(£_V\mathcal{F})^r] = \sum_{\gamma=0}^3 V^\gamma[\partial_\gamma\mathcal{F}^r] + [\mathcal{F}^{aw}]\, j\left(\partial V^0\right) + [\mathcal{F}^{ar}]\left(-[\partial v]^t + (div(v))I_3\right)$

$[(£_V\mathcal{F})^w] = \sum_{\gamma=0}^3 V^\gamma[\partial_\gamma\mathcal{F}^w] + [\mathcal{F}^{aw}]\left(\partial_0 V^0 + [\partial v]\right) - [\mathcal{F}^{ar}]\, j(\partial_0 V)$

with

$$[\partial v] = \begin{bmatrix} \partial_1 V^1 & \partial_2 V^1 & \partial_3 V^1 \\ \partial_1 V^2 & \partial_2 V^2 & \partial_3 V^2 \\ \partial_1 V^3 & \partial_2 V^3 & \partial_3 V^3 \end{bmatrix};$$

$div(v) = \partial_1 V^1 + \partial_2 V^2 + \partial_3 V^3$

$$\partial V^0 = \begin{bmatrix} \partial_1 V^0 \\ \partial_2 V^0 \\ \partial_3 V^0 \end{bmatrix}$$

$$j(\partial_0 V) = \begin{bmatrix} 0 & -\partial_0 V^3 & \partial_0 V^2 \\ \partial_0 V^3 & 0 & -\partial_0 V^1 \\ -\partial_0 V^2 & \partial_0 V^1 & 0 \end{bmatrix}$$

A.4.4 RELATIVIST GEOMETRY

Divergence of a vector field V :

$divV = \frac{1}{\det P'}\sum_{\alpha=0}^3\frac{d}{d\xi^\alpha}\left(V^\alpha\det P'\right) = \sum_{\alpha=0}^3\frac{dV^\alpha}{d\xi^\alpha} + \frac{1}{\det P'}\sum_{\alpha=0}^3 V^\alpha\frac{d\det P'}{d\xi^\alpha}$

$\frac{d\det P'}{d\xi^\alpha} = (\det P')\, Tr\left([\partial_\alpha P'][P]\right)$

Integral curve

$\frac{\partial}{\partial s}\Phi_V(s,x)|_{s=s_0} = V(\Phi_V(s_0,x))$

$\Phi_V(0,x) = x$

For a motion on an integral curve of the vector field V :
$\frac{1}{\det P'}\sum_{\alpha=0}^{3}V^{\alpha}\frac{d\det P'}{d\xi^{\beta}}=\frac{1}{\det P'}\frac{d\det P'}{d\tau}$

Trajectory of a particle

Velocity u of the particle, measured in its proper time :
$u=\frac{dp}{d\tau}$
$\langle u,u\rangle=-c^{2}$
Speed V of a particle as measured by an observer :
$p(t)=\Phi_{O}(ct,x(t))=\varphi_{o}(t,x(t))$
$V(t)=\frac{dp}{dt}=\sum_{\alpha=0}^{3}c\frac{\partial p^{\alpha}}{\partial t}\partial\xi_{\alpha}=c\partial\xi_{0}+\overrightarrow{v}=c\varepsilon_{0}(q(t))+\overrightarrow{v}$
$\overrightarrow{v}=\frac{\partial}{\partial x}\Phi_{O}(ct,x(t))\frac{\partial x}{\partial t}=\sum_{\alpha=1}^{3}\frac{d\xi^{\alpha}}{dt}\partial\xi_{\alpha}$
Between the proper time τ of a particle and the time t of an observer :
$\frac{d\tau}{dt}=\sqrt{1-\frac{\|\overrightarrow{v}\|^{2}}{c^{2}}}=\frac{1}{c}\sqrt{-\langle V,V\rangle}=\frac{c}{u^{0}}=\frac{1}{2a_{w}^{2}-1}$
Between the velocity u of a particle and the speed V as measured by an observer :
$u=\frac{dp}{d\tau}=V\frac{c}{\sqrt{-\langle V,V\rangle}}=\frac{1}{\sqrt{1-\frac{\|\overrightarrow{v}\|^{2}}{c^{2}}}}\left(\overrightarrow{v}+c\varepsilon_{0}(p(\tau))\right)$

Tetrad

$\varepsilon_{i}(m)=\sum_{\alpha=0}^{3}P_{i}^{\alpha}(m)\partial\xi_{\alpha}\Leftrightarrow\partial\xi_{\alpha}=\sum_{i=0}^{3}P_{\alpha}^{\prime i}(m)\varepsilon_{i}(m)$
$\varepsilon^{i}(m)=\sum_{i=0}^{3}P_{\alpha}^{\prime i}(m)d\xi^{\alpha}\Leftrightarrow d\xi^{\alpha}=\sum_{i=0}^{3}P_{i}^{\alpha}(m)\varepsilon^{i}(m)$
Standard chart :
$\xi^{0}=ct$
$\mathbf{O}(m)=\partial\xi_{0}$
$\varphi_{o}\left(\xi^{0},\xi^{1},\xi^{2},\xi^{3}\right)=\Phi_{O}(ct,x)$
$[P']=\begin{bmatrix}P'_{00}&P'_{10}&P'_{20}&P'_{30}\\P'_{10}&P'_{11}&P'_{12}&P'_{13}\\P'_{20}&P'_{21}&P'_{22}&P'_{23}\\P'_{30}&P'_{31}&P'_{32}&P'_{33}\end{bmatrix};[P]=\begin{bmatrix}P_{00}&P_{01}&P_{02}&P_{03}\\P_{10}&P_{11}&P_{12}&P_{13}\\P_{20}&P_{21}&P_{22}&P_{23}\\P_{30}&P_{31}&P_{32}&P_{33}\end{bmatrix}$
$[Q]=\begin{bmatrix}P_{11}&P_{12}&P_{13}\\P_{21}&P_{22}&P_{23}\\P_{31}&P_{32}&P_{33}\end{bmatrix}$
$[Q(\varphi_{o}(t,x))]=[Q(\varphi_{o}(0,x))][C_{q}(\varphi_{o}(t,x))]$
with $[C_{q}(\varphi_{o}(t,x))]\in GL(\mathbb{R},3),[C_{q}(\varphi_{o}(0,x))]=I_{3}$.
$[Q'((\varphi_{o}(t,x)))]=[C_{q}(\varphi_{o}(t,x))]^{-1}[Q'((\varphi_{o}(0,x)))]$

Metric

$[g]^{-1}=\underline{[P][\eta]}[P]^{t}\Leftrightarrow[g]=[P']^{t}[\eta][P']$
$\sqrt{-\det[g]}=\det P'$
$\partial_{\alpha}\det P'=(\det P')Tr([\partial_{\alpha}P'][P])$
$[g]=[P']^{t}[\eta][P']=\begin{bmatrix}-1&0\\0&[g]_{3}\end{bmatrix}$
$[g_{3}]=[Q']^{t}[Q']$
$[g_{3}]^{-1}=[Q][Q]^{t}$
$\det[g_{3}]=-\det g=(\det Q')^{2}$

$\det[g_3]^{-1} = (\det Q)^2$
$\varpi_4 = \det[P']\, d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$
$\varpi_3 = \det[P']\, d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

A.4.5 FIBER BUNDLES

i) $P_G(M, Spin_0(3,1), \pi_G)$:
$\mathbf{p}(m) = \varphi_G(m,1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m)\cdot\chi(m)^{-1}$:
$\sigma(m) = \varphi_G(m,\sigma(m)) = \widetilde{\varphi}_G(m,\chi(m)\cdot\sigma(m)) = \widetilde{\varphi}_G(m,\widetilde{\sigma}(m))$
$\sigma(m) \to \widetilde{\sigma}(m) = \chi(m)^{-1}\cdot\sigma(m)$
ii) $P_G\left[\mathbb{R}^4, \mathbf{Ad}\right]$:
$\varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i) \to \widetilde{\varepsilon}_i(m) = \mathbf{Ad}_{\chi(m)^{-1}}\varepsilon_i(m)$
$\partial\xi_\alpha = \sum_{i=0}^3 P'^i_\alpha \varepsilon_i(m) = \sum_{i=0}^3 \widetilde{P}'^i_\alpha \widetilde{\varepsilon}_i(m)$
$\Rightarrow \widetilde{P}'^i_\alpha = \sum_{j=0}^3 [h(\chi(m))]^i_j P'^j_\alpha$
$\left[\widetilde{P'(m)}\right] = [h(\chi(m))][P'(m)]$
iii) $P_G\left[\left(\mathbb{R}^4\right)^*, \mathbf{Ad}^{-1}\right]$:
$\varepsilon^i(m) = \left(\mathbf{p}(m), \varepsilon^i\right) \to \widetilde{\varepsilon}^i(m) = \mathbf{Ad}_{\chi(m)}\varepsilon^i(m)$
$d\xi^\alpha = \sum_{i=0}^3 P^\alpha_i \varepsilon^i(m) = \sum_{i=0}^3 \widetilde{P}^\alpha_i \widetilde{\varepsilon}^i(m)$
$\Rightarrow \widetilde{P}^\alpha_i = \sum_{j=0}^3 \left[h\left(\chi(m)^{-1}\right)\right]^j_i P^\alpha_j$
$\left[\widetilde{P(m)}\right] = [P(m)]\left[h\left(\chi(m)^{-1}\right)\right] = [P(m)][h(\chi(m))]^t$
iv) $P_G[E, \gamma C]$
$\mathbf{S}(m) = (\mathbf{p}(m), S_m) = \left(\mathbf{p}(m)\cdot\chi(m)^{-1}, \gamma C(\chi(m)) S_m\right)$
v) $P_G[T_1 Spin(3,1), \mathbf{Ad}]$:
$\kappa_a(m) = (\mathbf{p}(m), \kappa_a) \to \widetilde{\kappa}_a(m) = \mathbf{Ad}_{\chi(m)^{-1}}\kappa_a(m)$
$\upsilon(r(m), w(m)) = \widetilde{\upsilon}(\widetilde{r}(m), \widetilde{w}(m))$
$$\begin{bmatrix} \widetilde{r}(m) \\ \widetilde{w}(m) \end{bmatrix} = \left[\mathbf{Ad}_{\chi(m)}\right]\begin{bmatrix} r(m) \\ w(m) \end{bmatrix} = \begin{bmatrix} A & -B \\ B & A \end{bmatrix}\begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix}\begin{bmatrix} r(m) \\ w(m) \end{bmatrix}$$
vi) $P_G\left[T_1 Spin(3,1)^*, \mathbf{Ad}^{-1}\right]$:
$\kappa^a(m) = (\mathbf{p}(m), \kappa^a) \to \widetilde{\kappa}^a(m) = \mathbf{Ad}_{\chi(m)}\kappa^a(m)$
$\upsilon^*(X(m), Y(m)) = \widetilde{\upsilon}^*\left(\widetilde{X}(m), \widetilde{Y}(m)\right)$
$$\begin{bmatrix} \widetilde{X}(m) \\ \widetilde{Y}(m) \end{bmatrix} = \begin{bmatrix} X(m) \\ Y(m) \end{bmatrix}\left[\mathbf{Ad}_{\chi(m)^{-1}}\right] = \begin{bmatrix} X(m) \\ Y(m) \end{bmatrix}\begin{bmatrix} C^t & 0 \\ 0 & C^t \end{bmatrix}\begin{bmatrix} A & B \\ -B & A \end{bmatrix}$$
vii) P_U
$\mathbf{p}_U(m) = \varphi_U(m,1) \to \widetilde{\mathbf{p}}_U(m) = \widetilde{\varphi}_U(m,1) = \mathbf{p}_U(m)\cdot\chi(m)^{-1}$
$\varkappa(m) = \varphi_U(m, \varkappa(m)) = \widetilde{\varphi}_U(m, \chi(m)\cdot\varkappa(m))$
viii) $P_U[F, \varrho]$
$\mathbf{f}_j(m) = (\mathbf{p}(m), f_j) \to \widetilde{\mathbf{f}}_j(m) = \varrho\left(\chi(m)^{-1}\right)(\mathbf{f}_j(m))$
$\phi(m) \to \widetilde{\phi}(m) = \varrho(\chi(m))\phi(m)$

A.4.6 SPINORS and PARTICLES

Spinors

$S = \sum_{i=1}^4 S^i e_i \in E$

$S=\gamma C\left(\sigma_w\cdot\sigma_r\right)S_0$
$S_0=\begin{bmatrix}S_R\\S_L\end{bmatrix}=\begin{bmatrix}v\\\epsilon iv\end{bmatrix},v\in\mathbb{C}^2$
$\left\langle S_0,S_0\right\rangle=\epsilon M_p^2c^2$
$S_R=\frac{M_pc}{\sqrt{2}}\begin{bmatrix}e^{i\alpha_1}\cos\alpha_0\\e^{i\alpha_2}\sin\alpha_0\end{bmatrix}$
Relation between the speed V of a particle and an element $\sigma_w\in PW$:
$u\left(m\right)=\mathbf{Ad}_{\sigma_w(m)}\varepsilon_0\left(m\right)=\left(\mathbf{p}\left(m\right),\mathbf{Ad}_{\sigma_w}\varepsilon_0\right)=\left(\mathbf{p}\left(m\right),u\right)$
$u=\frac{dp}{d\tau}=c\left(\left(2a_w^2-1\right)\varepsilon_0+\epsilon a_w\sum_{i=1}^3w_i\varepsilon_i\right)$
$V=\frac{dp}{dt}=\overrightarrow{v}+c\varepsilon_0\left(m\right)=c\left(\varepsilon_0+\epsilon\frac{a_w}{2a_w^2-1}\sum_{i=1}^3w_i\varepsilon_i\right)$

Momentum

$\mathcal{M}=\sum_{\alpha=0}^3\gamma C\left(\sigma^{-1}\cdot\partial_\alpha\sigma\right)S_0\otimes d\xi^\alpha\in\Lambda_1\left(M;M_0\right)$
$\partial_\alpha S=\gamma C\left(\sigma\right)\gamma C\left(v\left(\left(\left[C\left(r\right)\right]^t\left(\frac{1}{2}j\left(w\right)\partial_\alpha w\right)+\left[D\left(r\right)\right]^t\partial_\alpha r\right),\left[C\left(r\right)\right]^t\left(\frac{1}{4a_w}\left[4-j\left(w\right)j\left(w\right)\right]\partial_\alpha w\right)\right)\right)S_0$
$\left[C\left(r\right)\right]=\left[1+a_rj\left(r\right)+\frac{1}{2}j\left(r\right)j\left(r\right)\right]$
$\left[D\left(r\right)\right]=\left[\frac{1}{a_r}+\frac{1}{2}j\left(r\right)+\frac{1}{4a_r}j\left(r\right)j\left(r\right)\right]$ and $\left[D\left(r\right)\right]=\left[C\left(r\right)\right]\left[D\left(r\right)\right]^t$

Inertial vector

$\frac{1}{i}\left\langle S,\frac{dS}{d\tau}\right\rangle=k^tX$
$X=\left[C\left(r\right)\right]^t\left(\frac{1}{2}j\left(w\right)\frac{dw}{dt}\right)+\left[D\left(r\right)\right]^t\frac{dr}{dt}$
$k^a=-\epsilon\left(S_R^*\sigma_aS_R\right)$
$k=-\epsilon\frac{1}{2}M_p^2c^2k_0$
$k_0=\begin{bmatrix}\left(\sin2\alpha_0\right)\cos\left(\alpha_2-\alpha_1\right)\\\left(\sin2\alpha_0\right)\sin\left(\alpha_2-\alpha_1\right)\\\cos2\alpha_0\end{bmatrix};k_0^tk_0=1$

State of a particle

$\psi=\sum_{i=1}^4\sum_{j=1}^n\psi^{ij}e_i\otimes f_j\in E\otimes F$
$\left[\psi\right]_{j=1...m}^{i=1,2,3,4}$
$\psi=\vartheta\left(\sigma,\varkappa\right)\left[\psi_0\right]=\left[\gamma C\left(\sigma\right)\right]\left[\psi_0\right]\left[\varrho\left(\varkappa\right)\right]\in Q\left[E\otimes F,\vartheta\right]$
$\left[\psi_0\right]=\begin{bmatrix}\psi_R\\\epsilon i\psi_R\end{bmatrix}$
$\left\langle\psi,\psi'\right\rangle=Tr\left(\left[\psi\right]^*\left[\gamma_0\right]\left[\psi'\right]\right)$
$\left\langle\vartheta\left(\sigma,\varkappa\right)\psi,\vartheta\left(\sigma,\varkappa\right)\psi'\right\rangle=\left\langle\psi,\psi'\right\rangle$
$\left(f_i\right)_{i=1}^n$ is an orthonormal basis of F
$\left(\overrightarrow{\theta}_a\right)_{a=1}^m$ is a basis of the *Lie algebra* T_1U
$\left[\theta_a\right]$ is the matrix of $\varrho'\left(1\right)\left(\overrightarrow{\theta}_a\right)$ expressed in the basis $\left(f_i\right)_{i=1}^n:\left[\theta_a\right]=-\left[\theta_a\right]^*$
Mass at rest : $M_p=\frac{1}{c}\sqrt{\epsilon\left\langle\psi_0,\psi_0\right\rangle}=\frac{1}{c}\sqrt{3\epsilon Tr\left(\psi_R^*\psi_R\right)}$
Momentum :
$\mathcal{M}=\sum_{\alpha=0}^3\left(\left[\gamma C\left(\sigma^{-1}\partial_\alpha\sigma\right)\right]\left[\psi_0\right]+\left[\psi_0\right]\left[\varrho'\left(1\right)\left(g^{-1}\partial_\alpha g\right)\right]\right)\otimes d\xi^\alpha\in\Lambda_1\left(M;E\otimes F\right)$
Continuity equation : $\mu divV+\frac{d\mu}{dt}=0$

A.4.7 FORCE FIELDS

Connections :

On P_U :

$\grave{A} \in \Lambda_1(M; T_1U) : TM \to T_1U :: \grave{A}(m) = \sum_{\alpha=0}^{3} \sum_{a=1}^{m} \grave{A}_\alpha^a(m) \overrightarrow{\theta}_a \otimes d\xi^\alpha$

$\left[\grave{A}_\alpha\right] = \sum_{a=1}^{m} \grave{A}_\alpha^a [\theta_a]$

$\grave{A}(m) \to \widetilde{\grave{A}}(m) = Ad_\chi \left(\grave{A}(m) - L'_{\chi^{-1}}(\chi)\chi'(m)\right)$

$\nabla^U \mathbf{p}_g = \left(L'_{g^{-1}} g\right)(g'(m)) + \sum_{\alpha=0}^{3} Ad_{g^{-1}} \grave{A}_\alpha(m) d\xi^\alpha \in \Lambda_1(M, T_1U)$

$\nabla^F \phi = \sum_{\alpha=0}^{3} \left(\partial_\alpha \phi^i + \sum_{i=1}^{n} \left[\grave{A}_\alpha\right]_j^i \phi^j\right) \mathbf{f}_i(m) \otimes d\xi^\alpha \in \Lambda_1(M, P_U[F, \varrho])$

On P_G :

$G(m) = \sum_{\alpha=0}^{3} \upsilon(G_{r\alpha}(m), G_{w\alpha}(m)) d\xi^\alpha G(m) = \sum_{\alpha=0}^{3} \upsilon(G_{r\alpha}(m), G_{w\alpha}(m)) d\xi^\alpha \in \Lambda_1(M; T_1 Spin(3,1))$

$G(m) \to \widetilde{G}(m) = \mathbf{Ad}_\chi \left(G(m) - L'_{\chi^{-1}}(\chi)\chi'(m)\right)$

$\nabla^G \sigma = \sigma^{-1} \cdot \sigma' + \mathbf{Ad}_{\sigma^{-1}} G = \upsilon(X_\alpha, Y_\alpha) d\xi^\alpha$

with

$X_\alpha = [C(r)]^t \left(([D(r)]\partial_\alpha r + \frac{1}{2} j(w)\partial_\alpha w + [A(w)] G_{r\alpha} + [B(w)] G_{w\alpha}\right)$

$Y_\alpha = [C(r)]^t \left(\frac{1}{4a_w}[4 - j(w)j(w)]\partial_\alpha w - [B(w)] G_{r\alpha} + [A(w)] G_{w\alpha}\right)$

$[C(r)] = \left[1 + a_r j(r) + \frac{1}{2} j(r) j(r)\right]$

$[D(r)] = \left[\frac{1}{a_r} + \frac{1}{2} j(r) + \frac{1}{4a_r} j(r) j(r)\right]$ and $[D(r)] = [C(r)][D(r)]^t$

$[A(w)] = \left[1 - \frac{1}{2} j(w) j(w)\right]$

$[B(w)] = a_w [j(w)]$

$\nabla^M V = \sum_{\alpha i=0}^{3} \left(\partial_\alpha V^i + \sum_{j=0}^{3} [\Gamma_{M\alpha}(m)]_j^i V^j\right) \varepsilon_i(m) \otimes d\xi^\alpha$

$$[\Gamma_{M\alpha}] = \sum_{a=1}^{6} G_\alpha^a [\kappa_a] = \begin{bmatrix} 0 & G_{w\alpha}^1 & G_{w\alpha}^2 & G_{w\alpha}^3 \\ G_{w\alpha}^1 & 0 & -G_{r\alpha}^3 & G_{r\alpha}^2 \\ G_{w\alpha}^2 & G_{r\alpha}^3 & 0 & -G_\alpha^1 \\ G_{w\alpha}^3 & -G_{r\alpha}^2 & G_{r\alpha}^1 & 0 \end{bmatrix}$$

$\nabla^S S = \sum_{\alpha=0}^{3} (\partial_\alpha S + \gamma C(G_\alpha) S) d\xi^\alpha = \sum_{\alpha=0}^{3} (\partial_\alpha S + \gamma C(\upsilon(G_{r\alpha}, G_{w\alpha})) S) d\xi^\alpha$

$\gamma C(\upsilon(G_{r\alpha}, G_{w\alpha})) = -i\frac{1}{2} \sum_{a=1}^{3} (G_{w\alpha}\gamma_a\gamma_0 + G_{r\alpha}\widetilde{\gamma}_a)$

Total connection

$\nabla_\alpha \psi \in \Lambda_1(M, Q[E \otimes F, \vartheta])$:

$[\nabla_\alpha \psi] = \sum_{\alpha=0}^{3} [\partial_\alpha \psi] + [\gamma C(G_\alpha)][\psi] + [\psi]\left[\grave{A}_a\right]$

$\nabla_\alpha \psi = \vartheta(\sigma, \varkappa)\left(\left[\gamma C\left(\sigma^{-1}\partial_\alpha \sigma + \mathbf{Ad}_{\sigma^{-1}} G_\alpha\right)\right][\psi_0] + [\psi_0]\left[Ad_\varkappa\left(\grave{A}_\alpha\right)\right]\right)$

Energy of a particle

$\operatorname{Im}\langle\psi, \nabla_\alpha \psi\rangle = \frac{1}{i}\left(\langle\psi, \partial_\alpha \psi\rangle + \left\langle\psi, [\psi]\left[\grave{A}_\alpha\right]\right\rangle + \langle\psi, [\gamma C(G_\alpha)]\psi\rangle\right)$

kinetic energy :

$\frac{1}{i}\left\langle\psi, \frac{d\psi}{dt}\right\rangle = \left\langle\psi_0, \gamma C\left(\sigma^{-1}\frac{d\sigma}{dt}\right)\psi_0\right\rangle = k^t\left([C(r)]^t\left([D(r)]\frac{dr}{dt} + \frac{1}{2} j(w)\frac{dw}{dt}\right)\right)$

action of the fields :

$\frac{1}{i}\left\langle\psi, [\psi]\left[\widehat{A}\right]\right\rangle + \left\langle\psi, \gamma C\left(\widehat{G}\right)\psi\right\rangle = k^t [C(r)]^t \left([A(w)]\widehat{G}_r + [B(w)]\widehat{G}_w\right) + \frac{1}{i}\left\langle\psi_0, [\psi_0]\left[Ad_\varkappa \widehat{A}\right]\right\rangle$

$[A(w)] = \left[1 - \frac{1}{2} j(w) j(w)\right]$
$[B(w)] = a_w [j(w)]$
$[C(r)] = \left[1 + a_r j(r) + \frac{1}{2} j(r) j(r)\right]$
$[D(r)] = \left[\frac{1}{a_r} + \frac{1}{2} j(r) + \frac{1}{4a_r} j(r) j(r)\right]$
$[C(r)]^t [D(r)] = [D(r)]^t$

Propagation of fields

Gravitational Field

$\mathcal{F}_G = \sum_{a=1}^{6} \left(dG^a + \sum_{\alpha\beta=0}^{3} [G_\alpha, G_\beta]^a d\xi^\alpha \wedge d\xi^\beta \right) \otimes \overrightarrow{\kappa}_a$
$= \sum_{a=1}^{6} \sum_{\{\alpha,\beta\}=0}^{3} \left(\partial_\alpha G_\beta^a - \partial_\beta G_\alpha^a + 2 [G_\alpha, G_\beta]^a \right) d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\kappa}_a$
$= \sum_{a=1}^{6} \sum_{\alpha,\beta=0}^{3} \mathcal{F}_{G\alpha\beta}^a d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\kappa}_a$
$= \sum_{\{\alpha,\beta\}=0}^{3} \upsilon(\mathcal{F}_{r\alpha\beta}, \mathcal{F}_{w\alpha\beta}) d\xi^\alpha \wedge d\xi^\beta$
$[\mathcal{F}_{G\alpha\beta}] = \sum_{a=1}^{6} \mathcal{F}_{G\alpha\beta}^a [\kappa_a] = [K(\mathcal{F}_{w\alpha\beta})] + [J(\mathcal{F}_{r\alpha\beta})]$
$[\mathcal{F}_{\alpha\beta}] = [\partial_\alpha \Gamma_{M\beta}] - [\partial_\beta \Gamma_{M\alpha}] + [\Gamma_{M\alpha}][\Gamma_{M\beta}] - [\Gamma_{M\beta}][\Gamma_{M\alpha}]$
with the signature (3,1) :
$\mathcal{F}_{r\alpha\beta} = \upsilon(\partial_\alpha G_{r\beta} - \partial_\beta G_{r\alpha} + 2(j(G_{r\alpha}) G_{r\beta} - j(G_{w\alpha}) G_{w\beta}), 0)$
$\mathcal{F}_{w\alpha\beta} = \upsilon(0, \partial_\alpha G_{w\beta} - \partial_\beta G_{w\alpha} + 2(j(G_{w\alpha}) G_{r\beta} + j(G_{r\alpha}) G_{w\beta}))$
Change of gauge : $\mathbf{p}_G(m) = \varphi_G(m,1) \to \widetilde{\mathbf{p}}_G(m) = \mathbf{p}_G(m) \cdot s(m)^{-1}$:
$\mathcal{F}_{G\alpha\beta} \to \widetilde{\mathcal{F}}_{G\alpha\beta}(m) = \mathbf{Ad}_{s(m)} \mathcal{F}_{G\alpha\beta}$
Matrix representation :

$$[\mathcal{F}]_{6\times6} = \begin{bmatrix} \mathcal{F}_r^r & \mathcal{F}_r^w \\ \mathcal{F}_w^r & \mathcal{F}_w^w \end{bmatrix} = \left[\mathcal{F}_{G\alpha\beta}^a\right]$$

$$[\mathcal{F}_r^r]_{3\times3} = \begin{bmatrix} \mathcal{F}_{G32}^1 & \mathcal{F}_{G13}^1 & \mathcal{F}_{G21}^1 \\ \mathcal{F}_{G32}^2 & \mathcal{F}_{G13}^2 & \mathcal{F}_{G21}^2 \\ \mathcal{F}_{G32}^3 & \mathcal{F}_{G13}^3 & \mathcal{F}_{G21}^3 \end{bmatrix}$$

$$[\mathcal{F}_r^w]_{3\times3} = \begin{bmatrix} \mathcal{F}_{G01}^1 & \mathcal{F}_{G02}^1 & \mathcal{F}_{G03}^1 \\ \mathcal{F}_{G01}^2 & \mathcal{F}_{G02}^2 & \mathcal{F}_{G02}^2 \\ \mathcal{F}_{G01}^3 & \mathcal{F}_{G02}^3 & \mathcal{F}_{G03}^3 \end{bmatrix}$$

$$[\mathcal{F}_w^r]_{3\times3} = \begin{bmatrix} \mathcal{F}_{G32}^4 & \mathcal{F}_{G13}^4 & \mathcal{F}_{G21}^4 \\ \mathcal{F}_{G32}^5 & \mathcal{F}_{G13}^5 & \mathcal{F}_{G21}^5 \\ \mathcal{F}_{G32}^6 & \mathcal{F}_{G13}^6 & \mathcal{F}_{G21}^6 \end{bmatrix}$$

$$[\mathcal{F}_w^w]_{3\times3} = \begin{bmatrix} \mathcal{F}_{G01}^4 & \mathcal{F}_{G02}^4 & \mathcal{F}_{G03}^4 \\ \mathcal{F}_{G01}^5 & \mathcal{F}_{G02}^5 & \mathcal{F}_{G02}^5 \\ \mathcal{F}_{G01}^6 & \mathcal{F}_{G02}^6 & \mathcal{F}_{G03}^6 \end{bmatrix}$$

Hodge dual :

$[*\mathcal{F}_r^r] = \left([\mathcal{F}_r^w] \left(-g^{00} [g_3^{-1}] + [g^{-1}]_0 [g^{-1}]^0 \right) + [\mathcal{F}_r^r] j\left([g^{-1}]_0\right) [g_3^{-1}] \right) \det P'$
$[*\mathcal{F}_w^r] = \left([\mathcal{F}_w^w] \left(-g^{00} [g_3^{-1}] + [g^{-1}]_0 [g^{-1}]^0 \right) + [\mathcal{F}_w^r] j\left([g^{-1}]_0\right) [g_3^{-1}] \right) \det P'$
$[*\mathcal{F}_r^w] = -\left([\mathcal{F}_r^w] [g_3^{-1}] j\left([g^{-1}]_0\right) + (\det Q)^2 [\mathcal{F}_r^r] [g_3] \right) \det P'$
$[*\mathcal{F}_w^w] = -\left([\mathcal{F}_w^w] [g_3^{-1}] j\left([g^{-1}]_0\right) + (\det Q)^2 [\mathcal{F}_w^r] [g_3] \right) \det P'$

Scalar product :

$\langle \mathcal{F}, K \rangle_G$

$= \sum_{a=1}^{3} \sum_{\{\alpha\beta\}} \mathcal{F}_r^{a\alpha\beta} K_{r\alpha\beta}^a - \mathcal{F}_w^{a\alpha\beta} K_{w\alpha\beta}^a = \frac{1}{2} \sum_{a=1}^{3} \sum_{\alpha\beta} \mathcal{F}_r^{a\alpha\beta} K_{r\alpha\beta}^a - \mathcal{F}_w^{a\alpha\beta} K_{w\alpha\beta}^a$

$= -\frac{1}{\det P'} Tr \left(\left([*\mathcal{F}_r^w] [K_r^r]^t + [*\mathcal{F}_r^r] [K_r^w]^t \right) - \left([*\mathcal{F}_w^w] [K_w^r]^t + [*\mathcal{F}_w^r] [K_w^w]^t \right) \right)$

$\langle \mathcal{F}, \mathcal{F} \rangle_G = \frac{1}{\det P'} Tr \left(\left([*\mathcal{F}_r^w] [\mathcal{F}_r^r]^t + [*\mathcal{F}_r^r] [\mathcal{F}_r^w]^t \right) - \left([*\mathcal{F}_w^w] [\mathcal{F}_w^r]^t + [*\mathcal{F}_w^r] [\mathcal{F}_w^w]^t \right) \right)$

$= \sum_{a=1}^{3} \sum_{\{\alpha\beta\}} \mathcal{F}_r^{a\alpha\beta} \mathcal{F}_{r\alpha\beta}^a - \mathcal{F}_w^{a\alpha\beta} \mathcal{F}_{w\alpha\beta}^a$

$\langle \mathcal{F}, \mathcal{F} \rangle_G \varpi_4 = 2 \left(\sum_{a=1}^{3} *\mathcal{F}_r^{ar} \wedge \mathcal{F}_r^{aw} - *\mathcal{F}_w^{ar} \wedge \mathcal{F}_w^{aw} \right)$

Riemann tensor

$R = \sum_{\{\alpha\beta\} ij} [R_{\alpha\beta}]_j^i \, d\xi^\alpha \wedge d\xi^\beta \otimes \varepsilon_i(m) \otimes \varepsilon^j(m)$

$[R_{\alpha\beta}] = \sum_{a=1}^{6} \mathcal{F}_{G\alpha\beta}^a [P] [\kappa_a] [P']$

$R_{\alpha\beta\gamma\eta} = -R_{\alpha\beta\eta\gamma}$ with $R_{\alpha\beta\gamma\eta} = \sum_\lambda R_{\alpha\beta\gamma}^\lambda g_{\lambda\eta}$

Ricci tensor

$Ric = \sum_{\alpha\beta\gamma} ([P] [\mathcal{F}_{G\alpha\gamma}] [P'])_\beta^\gamma \, d\xi^\alpha \otimes d\xi^\beta$

Scalar curvature

$\mathbf{R} = \sum_{\alpha\beta} \sum_{a=1}^{6} \mathcal{F}_{G\alpha\gamma}^a \left([P] [\kappa_a] [\eta] [P]^t \right)_\alpha^\beta$

$\mathbf{R} = 2Tr \left(-(\det Q) [\mathcal{F}_r^r] \left[Q^{-1} \right]^t + [\mathcal{F}_w^r] \, j([P_0]) [Q] \right)$

Other Fields

Strength of the field

$\mathcal{F}_A = \sum_{a=1}^{m} \left(d \left(\sum_{\alpha=0}^{3} \grave{A}_\alpha^a d\xi^\alpha \right) + \sum_{\alpha\beta} \left[\grave{A}_\alpha, \grave{A}_\beta \right] d\xi^\alpha \wedge d\xi^\beta \right) \otimes \overrightarrow{\theta}_a$

$\mathcal{F}_A = \sum_{a=1}^{m} \sum_{\{\alpha,\beta\}} \left(\mathcal{F}_{A\alpha\beta}^a d\xi^\alpha \wedge d\xi^\beta \right) \otimes \overrightarrow{\theta}_a \in \Lambda_2(M; T_1U)$

$\mathcal{F}_{A\alpha\beta}^a = \partial_\alpha \grave{A}_\beta^a - \partial_\beta \grave{A}_\alpha^a + 2 \left[\grave{A}_\alpha, \grave{A}_\beta \right]^a$

Change of gauge : $\mathbf{p}_U(m) = \varphi_{P_U}(m, 1) \to \widetilde{\mathbf{p}}_U(m) = \mathbf{p}_U(m) \cdot \varkappa(m)^{-1}$:

$\mathcal{F}_{A\alpha\beta} \to \widetilde{\mathcal{F}}_{A\alpha\beta}(m) = Ad_{\varkappa(m)} \mathcal{F}_{A\alpha\beta}$

$[\mathcal{F}]_{m\times 6} = \begin{bmatrix} \mathcal{F}_A^r & \mathcal{F}_A^w \end{bmatrix} = \left[\mathcal{F}_{A\alpha\beta}^a \right]$

$[\mathcal{F}_A^r]_{m\times 3} = \begin{bmatrix} \mathcal{F}_{A32}^1 & \mathcal{F}_{A13}^1 & \mathcal{F}_{A21}^1 \\ ... & ... & ... \\ \mathcal{F}_{A32}^m & \mathcal{F}_{G13}^m & \mathcal{F}_{G21}^m \end{bmatrix}$

$[\mathcal{F}_A^w]_{m\times 3} = \begin{bmatrix} \mathcal{F}_{A01}^1 & \mathcal{F}_{A02}^1 & \mathcal{F}_{A03}^1 \\ ... & ... & ... \\ \mathcal{F}_{A01}^m & \mathcal{F}_{A02}^m & \mathcal{F}_{A03}^m \end{bmatrix}$

Hodge dual

$[*\mathcal{F}_A^r] = \left([\mathcal{F}_A^w] \left(-g^{00} \left[g_3^{-1} \right] + \left[g^{-1} \right]_0 \left[g^{-1} \right]^0 \right) + [\mathcal{F}_a^r] \, j\left(\left[g^{-1} \right]_0 \right) \left[g_3^{-1} \right] \right) \det P'$

$[*\mathcal{F}_A^w] = -\left([\mathcal{F}^w] \left[g_3^{-1} \right] j\left(\left[g^{-1} \right]_0 \right) + (\det Q)^2 [\mathcal{F}^r] [g_3] \right) \det P'$

Scalar product :

$$\langle \mathcal{F}, K \rangle_A = \sum_{a=1}^m \sum_{\{\alpha\beta\}} \mathcal{F}^{a\alpha\beta} K^a_{\alpha\beta} = \frac{1}{2} \sum_{a=1}^m \sum_{\alpha\beta=0}^3 \mathcal{F}^{a\alpha\beta} K^a_{\alpha\beta} = -\frac{1}{\det P'} Tr \left([*\mathcal{F}^w] [K^r]^t + [*\mathcal{F}^r] [K^w]^t \right)$$

$$\langle \mathcal{F}, \mathcal{F} \rangle_A = \sum_{a=1}^m \sum_{\{\alpha\beta\}} \mathcal{F}^{a\alpha\beta} \mathcal{F}^a_{\alpha\beta} = \frac{1}{\det P'} Tr \left([*\mathcal{F}^w] [\mathcal{F}^r]^t + [*\mathcal{F}^r] [\mathcal{F}^w]^t \right)$$

$$\langle \mathcal{F}, \mathcal{F} \rangle_A \varpi_4 = 2 \left(\sum_{a=1}^3 *\mathcal{F}^{ar}_A \wedge \mathcal{F}^{aw}_A \right)$$

Propagation

$$V(\Phi_V(s,n)) = \frac{dt}{ds} (c\varepsilon_0 + \overrightarrow{v})$$

$$divV = 0$$

$$F(m) = \mathcal{E}_0 - \langle \mathcal{F}, \mathcal{F} \rangle$$

$$\sum_{\lambda\mu} g^{\lambda\mu} (\partial_\lambda F)(\partial_\mu F) = 0$$

$\mathcal{F} = \widehat{\mathcal{F}} \otimes X$ with $\widehat{\mathcal{F}} \in \Lambda_2(M;\mathbb{R})$, and $X \in C_1(M;T_1U)$

$$\pounds_V \widehat{\mathcal{F}} = 0$$

Chern identity

$$Tr\left([\mathcal{F}(\varphi_o(t,x))^r_r]^t [\mathcal{F}(\varphi_o(t,x))^w_r] - [\mathcal{F}(\varphi_o(t,x))^r_w]^t [\mathcal{F}(\varphi_o(t,x))^w_w] \right) = 0$$

A.4.8 CONTINUOUS MODELS

Currents

$$\phi_G = \sum_{a=1}^6 \sum_{\beta=0}^3 \left[\mathcal{F}^{\alpha\beta}_r, G_{r\beta} \right]^a \overrightarrow{\kappa}_a \otimes \partial\xi_\alpha \in T_1 Spin(3,1) \otimes TM$$

$$\phi_A = \sum_{a=1}^m \sum_{\beta=0}^3 \left[\mathcal{F}^{\alpha\beta}_A, \grave{A}_\beta \right] \overrightarrow{\theta}_a \otimes \partial\xi_\alpha \in T_1U \otimes TM$$

$$J_G = \frac{C_I}{8C_G} \mu \frac{1}{i} \upsilon \left(\langle \psi, [\gamma C(\overrightarrow{\kappa}_a)][\psi] \rangle, -\langle \psi, [\gamma C(\overrightarrow{\kappa}_{a+3})][\psi] \rangle \right) \otimes V \in T_1 Spin(3,1) \otimes TM$$

$$J_G = \frac{C_I}{8C_G} \mu \upsilon \left([A(w(t))][C(r(t))]k, -[B(w(t))][C(r(t))]k \right) \otimes V$$

$$J_A = \frac{C_I}{8C_A} \mu \sum_{a=1}^m \left\langle \psi, [\psi] \frac{1}{i} [\theta_a] \right\rangle \overrightarrow{\theta}_a \otimes V \in T_1U \otimes TM$$

$$C_I \mu \operatorname{Im} \langle \psi, \nabla_V \psi \rangle = C_I \mu k^t \left([C(r)]^t \left([D(r)] \frac{dr}{dt} + \frac{1}{2} j(w) \frac{dw}{dt} \right) \right)$$

$$+8 \sum_{\beta=0}^3 \left(4C_G \left\langle J^\beta_G, G_\beta \right\rangle_{Cl} + C_A \left\langle J^\beta_A, \grave{A}_\beta \right\rangle_{T_1U} + C_{EM} J^\beta_{EM} \grave{A}_\beta \right)$$

Codifferential equation

$$\phi_{EM} = 0; J^*_{EM} = -\frac{1}{2} \delta \mathcal{F}_{EM}$$

$$J_A = \phi_A; J_G = \phi_G;$$

$$d(*\mathcal{F}_A) = 0; d(*\mathcal{F}_G) = 0;$$

$$\Delta\mathcal{F}_A = -\delta d\mathcal{F}_A; \Delta\mathcal{F}_G = -\delta d\mathcal{F}_G$$

Energy momentum tensor

$$T^\alpha_\beta = C_I \mu \frac{1}{i} V^\alpha \langle \psi, \partial_\beta \psi \rangle + 4 \sum_{\gamma=0}^3 4C_G \langle \mathcal{F}^{\alpha\gamma}_G, \partial_\beta G_\gamma \rangle_{Cl} + C_A \left\langle \mathcal{F}^{\alpha\gamma}_A, \partial_\beta \grave{A}_\gamma \right\rangle_{T_1U}$$

A.4.9 BOSONS

$$\Delta\mathcal{F}_A(V) = -\mathbf{p}^*(m) \Delta\grave{\mathbf{A}}(m)$$

$$B_0(s) = \sum_{\alpha=0}^3 V^\alpha \Delta\mathcal{F}_A(V)(\Phi_V(s,n)) \in T_1U$$

$$B = B_0 \otimes V \in T_1U \otimes TM$$

$\mathcal{E}_B = \langle \mathcal{F}_A(V), \Delta\mathcal{F}_A(V) \rangle$