THEORETICAL PHYSICS

Jean Claude Dutailly

Hal-id: 01 38 96 71 v1 – submitted 2016-10-28

ISBN-13: 978-1539889434

ISBN-10: 1539889432

Cover page picture: Nasa / JPL

Updated : July 19, 2017

This book proposes a new formulation 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 objects to deal with them, chosen among well-grounded mathematical theories. Proceeding this way, the reader will have a comprehensive, consistent and rigorous understanding of the main topics of the Physics of the XXI° century, together with many tools to do practical computations.

After a short introduction about the meaning of Theories 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, from 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 motion leads to associate a frame to all material bodies, whatever their scale, and to the representation of motion in Clifford Algebras. Momenta, translational and rotational, are then represented by spinors, which provide a clear explanation for the spin and the existence of anti-particles.

The force fields are introduced through connections, in the 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 solutions for the equations representing a system in the most general case.

The topic of the last chapter is discontinuous processes. The phenomenon of collision is studied, and we show that bosons can be understood as discontinuities in the fields.

In the Version Updated 7/19/2017 : the presentation of some important topics has been improved, and new results added.

Geometry : the definition of matter fields has been improved. Symmetries have been added.

Kinematics : the model of atoms has been added.

Fields : an introduction to the Einstein's Theory of gravitation has been added. The section on the phenomenon of propagation has been rewritten. It is shown that fields propagate on Killing curves. As a consequence a general specification of the metric can be given.

Contents

INTRODUCTION

With each new discovery Physics has expanded into new theories Mechanics, Thermodynamics, Electromagnetism, Fluid Mechanics,... Beyond their diversity, they share a common core of key concepts and First Principles. Relativity, then Quantum Mechanics have broken this unity. A century after their introduction these powerful theories have not brought what could have been expected, that is a unified framework, consistent, intellectually satisfying, and efficient. For some Physicists we have to accept the idea of two physics, based on unrecognizable visions of the world, for others we have to give up altogether the idea of a real world and contend ourselves with more or less consistent formal systems justified only by their immediate efficiency. For most this is not a great concern, as far as the computation works. But not everybody is satisfied by this sorry state of Physics. 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,...

The purpose of this book is not 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 in Physics. And indeed we do not answer to all questions in this book, some issues are still open, but I hope that their meaning will be clearer, leading the way to a better and stronger understanding of the real world. Its purpose is to propose a unified Theory of Physics. Not all the domains are covered, but it addresses the key topics : the Geometry of the Universe, the Kinematics of material body and Mechanics, the Theory of Fields, their propagation and their interaction with material bodies, and the bases of the classification of Elementary Particles. I propose a Theory which can be understood, starting from the concepts such as space, time, mass, momentum,... and the usual First Principles which are known by every Physicist and have been used for centuries. But one needs a new, candid, look at these concepts and the phenomena they describe, 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. The meaning of the concepts has often be lost, replaced by some mathematical expression. In most books any consideration of a location is quickly followed by "let x, y, z be the coordinates of the point", without much of a thought for the fact that nobody use practically orthonormal frames to locate a point. These formal substitutes acquire a life of their own, and can become a real burden when they impede the full understanding of new theories. In Special Relativity "inertial frames" are part of the mandatory equipment, and the Quantum Theory of Fields is deemed incompatible with General

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

Relativity because giving up orthonormal frames seems too big an effort. For a century studies in General Relativity (GR) have been based on the metric and the Levy-Civita connection, without much of a physical justification, because they seem more convenient, and have become a standard in the field. Theoretical Physics has its “generally accepted practices”, such as the “substitution rule” in quantization, or the renormalisation to get rid of divergent integrals, which would surprise any non professional. In looking beyond the usual formalism, in regaining the true meaning of the physical properties and laws, it is possible to get a more sensible and unified picture. But to develop its full potential, we need the most adequate mathematical tools. Each new step in the progress of Physics has been made with a simultaneous advance in the formalism. The new tools exist, they need some effort to master them, but it is worth of it.

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 building, step by step, 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 another book ("Mathematics for Theoretical Physics"). A great effort has been done to develop practical tools, which make the computations easier. For instance a dozen lines suffice to express the Einstein law of GR in 3 linear, computable, equations. And we give explicit specifications for the metric, and formulas for its computation, in the most general case. The objective is not only to give a beautiful picture, but to provide a manageable Theory.

The first chapter is devoted to a bit of philosophy. From many discussions with scientists I felt that it is appropriate. Because the purpose of this book is to provide a new theoretical framework for Physics, it is necessary to have a good understanding of what is meant by physical laws, theories, validation by experiments, models, representations,... Epistemology, a branch of Philosophy, 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. In this Chapter are reminded the First Principles of Physics, fundamental laws which are universally accepted, and their general meaning.

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, it does not state 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.

There are three main objects in Physics : the Universe, material bodies and fields. They are the topic of the 3 following chapters.

By Universe 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. By Universe we mean the container, seen at our scale, in which every other object live, and how we represent it. This is the topic of Geometry, and any Theory in Physics must have a physical Geometry, telling how one locates a point, measure vectors or tensors, and of course, how one deals with time. The usual Geometries are the Galilean Geometry, Special Relativity (SR) and General Relativity (GR). In this book we adopt the latter. This choice will be justified by building the Geometry from 5 basic, natural assumptions, from where all the usual theorems can be deduced. For instance we show that the existence of a Lorentz metric is the logical consequence of the Principle of Causality. This metric is actually the main physical property of the Universe. We will introduce mathematical tools, such as the flow of a vector field and fiber bundles, which help to understand the geometry and to make practical computations. We will see the necessity to introduce the observer, the physicist who proceeds to the measures, as an integral part of the system.

The second object of Physics is material bodies, usually seen as a collection of material points which behave in some coherent way. Material points have a location and a translational motion, but material bodies have also other geometric properties, related to the concept of rotation. And it seems that these properties exist for any material body, whatever the scale, from elementary particles to galaxies. After a review of the classical representations, we show that these physical properties can be represented through a Clifford Algebra, we introduce tools to deal practically with any motion in the GR framework. They enable to extend easily the concept of deformable solid, from atoms to stars systems. Moreover they give the basis for the concept of matter field, a collection of particles which have a similar behavior.

The fourth chapter addresses Kinematics, which, by the concept of momentum, is the bridge between forces and geometry. The revision of the concept of motion of a material body requires the introduction of a new representation of the momentum, based on Spinors. Spinors are not new in Physics, but we will see why they are necessary and what they mean. This leads naturally to the introduction of the spin, which has a clear and simple interpretation, of antiparticles and to the representation of particles by fields of spinors, which are one the faces of the duality wave / particles. With these tools it is then easy to build a model of atoms and their electrons shells, using the theorems of quantization proven in the 2nd Chapter.

Particles interact with force fields, according to additional properties such as their electric charge. After a short reminder of the Standard Model we see in the 4th Chapter how to extend Spinors to represent the state of elementary particles, with all their properties, including their interactions with force fields.

The third object of Physics is Force Fields. They have been introduced in the late XIX° to replace the idea of action at a distance between material bodies. There are represented, in any physical theory, either as continuous objects, or as resulting from the action of special particles, the bosons, which carry the field. Modern Physics have introduced, with "gauge fields", a powerful tool to represent force fields in their continuous manifestation. In the 4th Chapter we show how it can be used to represent any field, including gravitation. 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. We give a short presentation of the Einstein's theory of gravitation, usually seen as being part of General Relativity, but is actually an original theory which deserves a special look to be genuinely understood. One important property of force fields is that they propagate in the vacuum. It is usually seen through the equations at equilibrium, of which the Maxwell's equations re the paradigm, but the phenomenon of propagation, to which too little attention has been given by theoricians, requires a comprehensive study. We show that all force fields propagate along specific curves, and from this result we deduce a general specification for the metric.

The Principle of Least Action is the main tool to model force fields and their interactions with particles. In the 6th we review the problems, physical and mathematical, of its implementation, 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. We show how the variational calculus can be implemented, in particular with a rigorous introduction of functional derivatives. It gives a solid framework for the introduction and justification of the Energy-Momentum tensor and conservation laws.

The 7th 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 by developing two models, for a field of particles and for individual particles. In this chapter we introduce the concept of currents and prove some important theorems. We give comprehensive and detailed guidelines to solve the equations in the most general context. In particular we prove that non relativist particles follow geodesics, we give a practical method to compute the fields, and an explicit solution for the metric.

The eighth chapter is dedicated to discontinuous processes. They are common in the real world but their study is difficult. We show how one can solve the problem of collision of particles in the general framework of GR and rotating bodies. 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, and how the results can be used in electrodynamics and the interactions of elementary particles.

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.

Using all the basic work done by philosophers, I will try to draw a schematic view of epistemology, with words which are 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 $^{\circ}C$J,“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 enable 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 unification 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.

There is some esthetic in Science. It is common to say about a theory that it is beautiful. And simplicity usually brings more beauty. Quoting Jauch “in all properly formulated physical ideas there is an economy of thought which is beautiful to contemplate. I have always been convinced that this esthetic aspect of a well-expressed physical theory is just as indispensable as its agreement with experience." (in Foundations of Quantum Mechanics).

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 spaces 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 Theory. 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 a determinist model of irreversible elementary process.

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 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

[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 has prevented them to happen.

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 FIVE 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 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,

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

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 Measure

Measure is certainly one of the characteristics of Science. The capacity to measure is indeed the condition for the development of a formalism and models. For instance Economics has achieved its full status with accounting. From there Measure has acquired a kind of sacred status. After all, the verdict of scientificity through falsifiability is based on measure. However the process of measure is more complex than reading a figure on an instrument. Any measure is actually a comparison between systems which are assumed to behave similarly. The most basic measure of lengths, by surveying, assume that the standard keeps its length. The new definition of the meter is based on the assumption that light has always the same speed in the vacuum. The data collected from experiments show the relations between systems that we know, and systems that we probe. This is the most obvious in Particles Physics : the color, representing the charge in strong interactions, is just a classification to identify particles which have the same behavior.

As a consequence concepts such as mass, charge, lengths,.. [4]and the units in which they are measured, lose their intrinsic character. The concepts stay, but one cannot say that this object "has" this length, we can just say that by comparison with other objects it has a property which shows constant correlations.

In some interpretations of QM the properties of the objects of Physics are nothing more than the relations between phenomena, statistically checked by repeated experiments. Eventually one cannot say anything about a property before it has been submitted to the process of measure. Its existence becomes a metaphysical assertion, without physical justification. Mathematical objects are attached to these relations (the variables of a model), physical laws are just mathematical computations, and the formalism is at least as legitimate than the physical properties it represents. This interpretation, common in QM, has important consequences.

It does not see anything strange in probabilist laws, since their validation is a statistical process, this is just an extension of their expression. But, to reject the determinism is not without risk : if ultimately all physical phenomena occur randomly the criterion of falsifiability would loose most of

[4]The coordinates of space and time used to locate a point in the Universe are similarly conventional as it is seen in the Chapter 3.

its merit. The second consequence is that it gives an incommensurate importance to the way the experiments are done. Notably to the possibility to measure or not simultaneously two quantities (the role given to the commutation rules in the formalism). But, as Relativity shows, simultaneity is a subtle concept, and obviously, measures, based on the comparison between similar phenomena, are never simultaneous. The third consequence is that the link with the evidence is lost : the objects of the formalism have not necessarily a real content (the wave function), mysterious objects appear on a regular basis, and virtuality reigns. So, while pretending to stay close to the empirical facts, actually this interpretation gives preeminence to the formalism over the reality. Moreover this interpretation misses an important point : the mathematical objects used in a model have also properties of their own, as we will see in the next chapter.

1.4.5 Dogmatism and Hubris in Science

As the criteria for the validation of Scientific knowledge began to emerge, the implementation of the same criteria led 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 appropriateness 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.

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 who 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".

For a material body the momentum is related to the motion. Motion is a purely geometric concept, corresponding to the change in the location and disposition of a material body with time. If the translational motion can be easily understood and modelled, the rotational motion is simple only for solids. But it seems clear that the motion of objects at the atomic scale should incorporate in some way these two components. Moreover the usual representations based on orthonormal frames in a 3 dimensional space must be adjusted to the relativist context.

The link between motion and momentum is done through kinematic characteristics of material bodies, such as mass and inertial tensor. Their representation must be done in accordance with the representation of motion, and then Relativity requires a profound adjustment, which has been done only partially. In particular rotation has a clear meaning only for rigid solid, whose concept cannot be transposed as such in Relativity.

Actually, if the momentum can be computed, only the change of momentum has a physical meaning, it is related to the forces and torques exerted to the body to change its motion. In a continuous motion the link is clear but not so in discontinuous processes, such as those occurring at the atomic level. Momenta are represented by vectorial quantities, in accordance to the usual representation of forces. However the representation of torques is essentially conventional in Newtonian Mechanics.

With the advent of Electromagnetism it has been clear that we should reject the idea of action at a distance, and this led to the introduction of a new object in physics : the force fields. They have special properties : they exist everywhere, they propagate at a finite speed, they interact with particles and this interaction depends on specific properties of particles, their charge. Actually the only field which is well known is the Electromagnetic field (EM). The concept of field is consistent only in the Relativist framework, however its propagation raises several issues, such as its measure by different observers. Their interaction with particles introduce, at least formally, a discontinuity, as interactions occur at a point, and the field propagates everywhere.

The concept of Energy comes from the work done by a force. This mechanical energy has a translational and a rotational component. Moments are vectorial and localized quantities, energy is a expressed by a scalar and has a more versatile definition. It has been enlarged with Thermodynamics, but it is essentially rooted in Mechanics. Thermodynamics considers internal energy as a state variable, which has an absolute value[5], an interpretation which as taken traction with Relativity, however only the flow of energy during processes, or between states, can be measured.

Because force fields and particles interact, energy must be exchanged during this process, and so we have to define and measure the flow of energy of a field.

The concepts of motion, momentum, and energy require a clear definition of the time, which can depend on the observer. It is assumed that energy and momentum are conserved at each time for the observer. Which leads to the concept of "potential energy" : when we lift an object from the surface we spend some energy, which is "stored" and can be recovered when the object falls. Formally the balance is kept even at each time, by assuming that the energy is exchanged in the interaction between the object and the gravitational field.

One feature to notice about the Principle of Conservation of Energy and Momentum 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.

[5] Actually the most important property of internal energy U in Thermodynamics is that its changes can be expressed by a total differential dU.

1.5.3 Principle of Least Action

The scope of the Principle of Least Action is more limited : it concerns continuous processes, which are considered over a definite period of time (or area of Space-Time) and describes the conditions for the equilibrium of a system.

It assumes that, in any physical process, a 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. Equilibrium does not imply that the state of the system is frozen, it can change 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 however, by construct, the Principle of Least Action describes the characteristics of an equilibrium in a continuous process. But is does not assume anything about the mechanisms by which this equilibrium is reached.

From Mechanics, this principle is usually represented by the assumption 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, as described above.

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 :

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

This formulation is extensively used, and many 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 and are smooth. 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 (they are then known) in L. So the concept of equilibrium is that of a global balance between the physical objects considered.

An equilibrium can be static or dynamic. A **static equilibrium** is reached when the state of the system does not change : the variables defining the system keep the same values all over the time. So it can be seen as a special solution in the implementation of the Principle of Least Action, where all the derivatives with respect to the time are null. But, because the concept of equilibrium encompasses the case of systems whose state evolves with time, one can consider **dynamic equilibrium**. And indeed, because the key variables are the maps (and not their value at a given point), the implementation of the Principle of Least Action provides solutions in which, at any given time, the system is at the equilibrium, even if the state changes. An object in a ballistic trajectory is always in equilibrium. However one can consider special solutions, in which the same states appear regularly : they are **periodic states**, such as the motion of planets in a star system, tides, thermodynamic cycles,... The concept of periodic equilibrium is linked to the idea of **stable state** : a system which always "looks the same", even if it changes. Quite often one knows, or can assume, that the stability of the system, attested by the measure of some of its properties, covers actually some internal processes which are not static. For instance the stable state of an atom, the system comprised of a nucleus and its electrons, does not imply that the state of the individual particles is fixed, but that they change according to a fixed pattern. The picture of a stable state depends on the scale : scale of which the phenomenon is observed, but also the frequency at which the state is measured. When a physicist states that a system is stable, he assumes that the changes which may occur do not matter for his purpose. There are two ways to deal with the characterization of a stable state : either one assumes that it is the result of random behaviors, whose sum (over time and space) is on average null, or that it is the result of a periodic behavior, whose manifestation cannot be measured, or is considered not significant. The choice of the method depends on the problem, but it is clear that the Principle of Least Action is not efficient to deal with random processes, meanwhile it is well suited to study periodic states : it is the simplest generalization of static states. Indeed in a periodic equilibrium it is assumed that all the variables take the same value at some periodic moment, so they can be considered as depending on the time only : if the location, represented by some coordinates x, is an argument of some variable $X(x)$, then, because the location will be always the same over a period, one can consider $X(x(t))$ with only the argument t.

The Principle of Least Action gives the conditions for an equilibrium, it is different from the Principle of Conservation of Energy. If the lagrangian represents the balance of energy between all the physical elements interacting in the system in an action such as :

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

the conservation of energy tells that, for a given observer, the integral should be constant for any variation at a given time. It brings a condition, in addition to the stationarity of the integral.

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 that they 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.

A physical process is not necessarily continuous. The distinction between continuous / discontinuous processes is made clear when one considers the mathematical formalism : it is related to the properties assigned to the variables. But 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. And from a physical point of view one can say, in a similar way, that a discontinuous process is characterized by the existence of a transition between two equilibriums. 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 by the interaction of microsystems. And in the study of a discontinuous process what matters most for the physicist is the transition, when and in which conditions it occurs.

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. When dynamic equilibriums are considered, several interpretations are possible. If it is clear in the Principle of Least Action, other definitions exist in Physics. In Thermodynamics equilibrium is identified with reversible processes, seen as slow processes : at any moment the system is close to equilibrium. In Theoretical Physics a common definition is 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 consisting 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 $X = (X_a)_{a=1}^n$ 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 requires 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.

This idea 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.

The concept of Entropy, whatever its form, has a clear meaning in the study of systems consisting of microsystems interacting. Then it shows :

- that not all states of equilibrium of the whole system are equivalent, and there is a driving force towards one of them;

- there are quantities, which have a meaning and can be measured for the whole system (such as temperature) but not at the level of microsystems.

As such it has a great importance in Physics, however it does not address the laws which rule the behavior of the microsystems, only the possible outcomes of their interactions.

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). And 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 is more important.

1.5.5 Principle of Locality

It can be stated as : "the outcome of any physical process occurring at a location depends only on 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 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, meanwhile 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 rational 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 coordinate is usually sufficient to account for the potential causality. In Quantum Mechanics this is more sensitive : almost all reasonings are 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.

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

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 deduce scientific laws, as theorems. And this is the experimental verification of these laws which provides 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 physics of elementary particles. It uses additional concepts (such as gauge fields), 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 sets 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, by defining a formal system. 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 gaps. 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. They 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 × 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 variables 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 × 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. 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 : $X'_k = F_k(X_k)$

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 $\alpha X + \beta X'$ have 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 usually not 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 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) Because the variables are maps, belonging to infinite dimensional vector spaces, a single occurrence of the model, that is a single experiment, provides possibly infinitely many measures. If $X_k : M \to \mathbb{R}$ is a map defined on a set M, then the measures could be $X_k(m_1), ...X_k(m_n)$ at points $m_1...m_n$, for any $n \in \mathbb{N}$. *The variables are the maps* $X_k : M \to \mathbb{R}$ *and not their values* $X_k(m)$ *at a given point* $m \in M$. The usual case is when they represent the evolution of the system with the time t : then X_k is the function itself : $X_k : \mathbb{R} \to O_k :: X_k(t)$

iv) In the conditions 1 we stay at the very first step of modelling : the description of all the mathematical objects which will be considered. The purpose of a model is to represent, in a common formalism, a broad range of similar physical experiments, with varying parameters, which are themselves part of the variables X_k. The variables can be restricted to take only some range (for instance it must be positive) : this is the meaning of the condition $X_k \in O_k$.

v) A Fréchet space is a Hausdorff, complete, topological space endowed with a countable family of semi-norms (Maths.12.2.6). 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), the spaces $L^p(M, \mu, E)$ of integrable sections of a vector bundle (valued in a Banach E);
- the spaces of continuously differentiable sections on a vector bundle, the spaces of differentiable functions on a manifold.

A topological vector space is separable if it has a dense countable subset (Maths.10.1.3) which, for a Fréchet space, is equivalent to be second countable. 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. The spaces $L^p(\mathbb{R}^n, dx, \mathbb{C})$ of integrable functions for $1 \le p < \infty$, 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.

vi) 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 states 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 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 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 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.15.1.3) 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 :

$\pi_V : V_0 \to \mathbb{R}_0^I :: \pi_V \left(\sum_{i\in I} x_i e_i\right) = x = (x_i)_{i\in I}$

where the set $\mathbb{R}_0^I \subset \mathbb{R}^I$ 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.12.5.7).

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 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 :

$\pi_H : H_0 \to \mathbb{R}_0^I :: \pi_H \left(\sum_{i\in I} y_i \varepsilon_i\right) = y = (y_i)_{i\in I}$

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.

$\mathbb{R}_0^I \subset \ell^2(I)$

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 (Maths.12.5.2).

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, 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 :

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

$\langle \widetilde{e}_i, \widetilde{e}_j \rangle_V = \langle \Xi(\widetilde{e}_i), \Xi(\widetilde{e}_j) \rangle_H = \langle \widetilde{\varepsilon}_i, \widetilde{\varepsilon}_j \rangle_H = \delta_{ij}$
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 $\forall j \in I : \left| \widetilde{\psi}_j \right| < r$
$\psi_i = \sum_{j \in J} [\mathcal{L}_H]_i^j \widetilde{\psi}_j \Rightarrow |\psi_i| < \varepsilon \sum_{j \in I} \max \left| [\mathcal{L}_H]_i^j \right|$ and $\left(\left| [\mathcal{L}_H]_i^j \right| \right)_{j \in I}$ 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.952). 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]

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 :

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

[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 manifolds. This is a point to be checked by mathematicians.

$\forall \psi \in H : i\psi = J(\psi)$
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 :
$\psi = \sum_\alpha \psi^{2\alpha}\varepsilon_{2\alpha} + \psi^{2\alpha+1}\varepsilon_{2\alpha+1} = \sum_\alpha \left(\psi^{2\alpha} - i\psi^{2\alpha+1}\right)\varepsilon_{2\alpha}$
$= \sum_\alpha \left(-i\psi^{2\alpha} + \psi^{2\alpha+1}\right)\varepsilon_{2\alpha+1}$
$\|\psi\|^2 = \sum_\alpha \left|\psi^{2\alpha} - i\psi^{2\alpha+1}\right|^2$
$= \sum_\alpha \left|\psi^{2\alpha}\right|^2 + \left|\psi^{2\alpha+1}\right|^2 + i\left(-\overline{\psi}^{2\alpha}\psi^{2\alpha+1} + \psi^{2\alpha}\overline{\psi}^{2\alpha+1}\right)$
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 consists 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 $F : V \to [0,1] :: F(X) = f \circ \Upsilon(X)$ 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.628). 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.708). Ω_1, Ω_2 are disjoint and closed in H. Then, by the Urysohn's Theorem (Maths.600) 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 function $F : V \to [0,1] :: F(X) = f \circ \Upsilon(X)$ is continuous and $F(X) = 1$ in O_1 and $F(X) = 0$ in O_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 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 : $\widehat{\Phi} \in \mathcal{L}(H; H_\Phi)$

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

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 $(e_i)_{i\in I} = \prod_{k=1}^{N} \left(e_i^k\right)_{i\in I_k}$ is a basis of V.*

So the procedure can involve simultaneously several variables. It requires only *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*

$$Y_J(X) = \sum_{i\in I} \left\langle \phi_i, \widehat{Y}_J(\Upsilon(X)) \right\rangle_H e_i \qquad (2.2)$$

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.1147) : 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.1151) 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 (on these topics Maths.III.3.). 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, 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.12.5.6).

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 $\overline{A}$" ■

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 (Maths.13.2).

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 :

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

ii) $P(E) = Id$

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

One can show that there is a bijective correspondence between the spectral measures on H and the maps : $\chi : \sigma_E \to H$ 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 $\widehat{\chi}_\psi : \sigma_E \to \mathbb{R} :: \widehat{\chi}_\psi(\varpi) = \frac{\langle \widehat{\pi}_{\chi(\varpi)}\psi, \psi\rangle}{\|\psi\|^2} = \frac{\|\widehat{\pi}_{\chi(\varpi)}\psi\|^2}{\|\psi\|^2}$ is a probability law on (E, σ_E).

2. An application of standard theorems on spectral measures tells that, for any bounded measurable function $f : E \to \mathbb{R}$, 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 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 $\chi : \sigma_{\mathbb{R}} \to H$ 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 $\widehat{\mu}_\psi : \sigma_{\mathbb{R}} \to \mathbb{R} :: \widehat{\mu}_\psi(\varpi) = \frac{\langle \widehat{\pi}_{\chi(\varpi)}\psi, \psi \rangle}{\|\psi\|^2} = \frac{\left\| \widehat{\pi}_{\chi(\varpi)}\psi \right\|^2}{\|\psi\|^2}$ 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 $\left(Y_{J_p}\right)_{p=1}^N$ 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.

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.1146) 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. 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 :

$$\Phi(X) = \sum_{i \in I} \left\langle \phi_i, \widehat{\Phi}(\Upsilon(X)) \right\rangle_H e_i \tag{2.3}$$

$\widehat{\Phi}$ 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** 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 :
$\forall Z \in B : \exists \alpha \in [0.1], Y, Y' \in A : Z = \alpha Y + (1-\alpha) Y'$
$\Rightarrow \forall m \in \Omega : Z(m) = x_p$
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 : $\sum_{m\in\Omega} \left\|\widehat{X}(m) - x_p\right\|_E$ (there can be additional conditions).

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 is 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 $\widehat{B}$ 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 $2^{N-1} - 1$ 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 comprise uniquely 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, 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 :

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

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 :

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

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

$\alpha = \int_w \prod_{p=1}^N \left((1-\lambda)\, \rho (x_p) + \lambda \rho_2 (x_p) \right) (dx)^N$

$\beta = \Pr(x \in w | H_1) = \int_w \left(\prod_{p=1}^N \rho (x_p) \right) (dx)^N$

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.

2.2.7 Observables defined by distributions

A measure is, in one way or another, the result of an experiment in which the unknown variable X acts on some known other variables φ to produce a finite number of data. For instance a field is measured by testing the behavior of known particles. We can model this measure as follows :

The variable $X \in V$ belongs to a vector space of maps, the "test functions" $\varphi \in W$ belongs to a vector space of maps W, with all the nice properties that we wish (they are smooth, compactly supported, and defined by a small number of parameters), the result of the experiment is expressed as a linear map :

$T : V \times W \to F :: T(X)(\varphi) = u$

where F is a finite dimensional vector space. T is linear and continuous in both variables.

The observable is then $T(X)$. Because φ is known and simple, and F finite dimensional it is expected that we have a good knowledge of X by doing enough experiments.

When X, φ are scalar complex functions defined on some manifold M, then $T(X) \in W'$: it belongs to the topological dual of W, and it can be shown that in the most general cases T can be expressed by an integral :

$T(X)(\varphi) = \int_M X(m)\varphi(m)\mu(m)$

with some fixed measure μ on M. Then $T(X)$ is called a "distribution" (or "generalized function"). And this can be extended to the case of maps $X \in C(M; \mathcal{L}(E; F)), \varphi \in C_{\infty,c}(M; E)$: at each point $m \in M$ the quantity $X(m)(\varphi(m))$ is a vector $u(m)$ of F and the integral is a vector of F. The key point in the definition is that $T(X)$ becomes a linear map, valued in F, and acting on the vector space $C_{\infty,c}(M; E)$ of smooth and compactly supported maps on M valued in E. In Mathematics the interest of distributions is that they enable to extend some operations (such as derivation) to any function (see Maths.7.2.2). But they have a physical meaning.

In our models the variables are maps, defined at any point of M, we can expect to measure their values $x_1, ...x_N$ at some points $m_1, ...m_N$ and, using a specification and a statistical method, estimate X. However a measure, as any experiment, is not done at a point (that is at a definite place and time), but in the neighborhood of a point and a short period of time. The test function φ is assumed to be compactly supported, which means that it is null outside of a compact area ω around a point. And to assign a value $T(X)$ to the observable states that, whatever the test function φ, the result will be given by $\int_M X(m)(\varphi(m))\mu(m)$. It is then customary to write : $T(X) = X$ that is to identify the variable X with the linear operator $T(X)$, and the identification is understood "in the meaning of distributions".

The interest of this construct is that one meets equations of the kind $X = Y$ where X, Y are defined on the same set M, and valued in the same vector space, but with different support. If Y is null out of a line N, then X should be null on the line, which is usually impossible if it is continuous, or worse smooth. By understanding the equation in the meaning of distributions we state that :

$\forall \varphi : \int_M X(m)(\varphi(m))\mu(m) = \int_N Y(m)(\varphi(m))\mu(m)$

Whatever the physical measure of the variables X, Y we get the same result.

The set $C_{\infty,c}(M; E)$ is chosen by the Physicist, it is as nice as he wishes and we can assume that this is a finite dimensional Hilbert space, as well as F Then, if X meets the conditions 1 there is an associated vector $\Upsilon(X) \in H$ and $\widehat{T}(\Upsilon(X))$ is a linear map : $\widehat{T} : H \times C_{\infty,c}(M; E) \to F ::$ $\widehat{T}(\Upsilon(X))(\varphi) = T(X)(\varphi)$

$\widehat{T} \circ \Upsilon = T \Leftrightarrow \widehat{T} = T \circ \Upsilon^{-1}$

$\widehat{T}$ is a linear map between Hilbert spaces, and its restriction to any finite dimensional vector space H_J is defined by a finite number of parameters. This is equivalent to take $T(Y_J)$ as observable, with a primary observable Y_J of X.

2.2.8 Structure defined by an observable

Let a model with variables $X \in O \subset V$ meeting the conditions 1. Then, for any occurrence of X a primary observable defines a map :

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

with $p = card(J)$ real parameters $x = (x_j)_{j \in J} : x_j = \left\langle \phi_j, \widehat{Y}_J(\Upsilon(X)) \right\rangle_H$. Or equivalently a map $f_J : O \to \mathbb{R}^p :: f_J(X) = x$. Because $\widehat{Y}_J$ is compact the image $\widehat{O}_J = f_J(O)$ is an open subset of $\mathbb{R}^p$.

Conversely, for any set x of parameters such that $x \in \widehat{O}_J$ the map $\widehat{X_j}(x) = \sum_{j \in J} x_j e_j$ provides an estimate of X, a map defined by p scalars which can be used in other problems.

Primary observables, as well as the maps $f_J, \widehat{X_j}$ are uniquely defined by the choice of p linearly independent vectors $(e_j)_{j \in J} \in V^J$. If we denote F_p the set of all maps f_J for all possible primary observables and C_p the set of all associated maps $\widehat{X_J} : \mathbb{R}^p \to V$ such that $\widehat{X_J}(x) \in O$, then the triple (O, C_p, F_p) define a p-structure on the model [2], which is linear : the model is characterized by p scalars.

$\forall X \in O : \widehat{X_J} \circ f_J(X) = Y_J(X)$

$\forall x \in \widehat{O} : f_J \circ \widehat{X_J}(x) = x$

However this structure has 2 limitations.

i) For a given set $(e_j)_{j \in J}$, the map $Y_J(X)$ is an optimal estimator of X but there is a discrepancy between X and $Y_J(X)$. It would be useful to know more about the quality of the estimation, and how it depends on the choice of $(e_j)_{j \in J}$. This point is addressed in the next section.

ii) A state of the system is represented by a single vector $\psi = \Upsilon(X) \in H$ but can provide infinitely many observations. If $X : M \to \mathbb{R}^N$ is a map over some set M, then the observations are done at points $m_1, ...m_n$ with any $n \in \mathbb{N}$. The vectors $X(m_1), ..., X(m_n)$ are not related to a Hilbert space. As a consequence there is no relation between p in the definition of a primary observable, and n, the number of available data. We can guess that the larger p, the better it will be to fit a large number of observations with $Y_J(X)$. The "wave function" provides a general answer. And when the variables depend on a scalar variable, as in a system depending on time, we have important results. They are seen in the next sections.

[2] The definition is an extenstion based on the paper "Smooth structures on fiber jet spaces" by Jan Slovak.

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. The set 2^I of all subsets of I is the largest $\sigma-$algebra on I. The set $\left(I, 2^I\right)$ is measurable.

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

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

is a probability law on $\left(I, 2^I\right)$: it is positive, countably additive and $\widehat{\mu}_\psi(I) = 1$ (on Probability Maths.11.4).

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 :*

$$\Pr(\Phi(X)|X) = \frac{\left\| \widehat{\Phi}(\Upsilon(X)) \right\|_H^2}{\|\Upsilon(X)\|_H^2} \tag{2.4}$$

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 $\Pr(J_p|X) = \frac{\left\|\widehat{Y}_p \psi\right\|^2}{\|\psi\|^2} = \frac{\|\psi_p\|^2}{\|\psi\|^2}$

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:

$$\Pr(\Phi(X) = \lambda_p X_p | X) = \frac{\|\Upsilon(X_p)\|^2}{\|\Upsilon(X)\|^2} \tag{2.5}$$

Which can also be expressed 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. It gives a - theoretical - answer to the question of the quality of an observable, that is how close its result $\Phi(X)$ is from X. The quantity $\Pr(\Phi(X)|X) = \frac{\left\|\widehat{\Phi}(\Upsilon(X))\right\|_H^2}{\|\Upsilon(X)\|_H^2}$ can be seen as a proxy for such an indicator. Its value depend on $\widehat{\Phi}$ and X. The larger the set $\widehat{\Phi}(\Upsilon(O))$ the better : the quality of a primary observable increases with $card(J)$.

This result can be seen 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.

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 at 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.

The function W provides the mean to address the quality of an observable, when the data are measures done at a collection of points $m_1, ..m_n$.

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 $U : V \to V'$ 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 : 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 $U\left(\theta\right)$. So that the measures are $Y\left(\theta\right) = \widetilde{\Phi} \circ U\left(\theta\right) X$ and the relation $Y\left(\theta\right) = Q\left(\theta\right) Y\left(\theta_0\right)$ reads : $\widetilde{\Phi} \circ U\left(\theta\right)\left(X\right) = U\left(\theta\right) \circ \widetilde{\Phi}\left(X\right)$. 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 by 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[3]

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* $H_\Phi = \widehat{\Phi}(H)$

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 :

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

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

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

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 :.

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

$\Upsilon U = \varphi \circ \Upsilon$ (Maths.III.13.4)
$\langle UX_1, UX_2 \rangle_V = \langle \Upsilon U X_1, \Upsilon U X_2 \rangle_H = \langle \varphi \circ \Upsilon X_1, \varphi \circ \Upsilon X_2 \rangle_H$
$= \langle \Upsilon X_1, \Upsilon X_2 \rangle_H = \langle X_1, X_2 \rangle$
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 3 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 3) 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 $W' \subset V$ is not necessarily identical to W. It is associated to the operator : $\widehat{\Phi}' = \Upsilon \circ \Phi' \circ \Upsilon^{-1}$. W and $W\prime$ are finite dimensional vector subspaces of V.

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

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 :

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

it 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 relates 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.

Representation

The left action of a group G on a set E is a map :$\lambda : G \times E \to E :: \lambda(g, x)$ 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 $f : G \to G\mathcal{L}(E; E)$ (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 E is a Hilbert space (there is a definite positive scalar product on E) and $f(g)$ is unitary : $\forall u, v \in F, g \in G : \langle f(g)u, f(g)v\rangle = \langle u, v\rangle$

Any continuous representation (H, f) of a topological group can be decomposed in the sum of mutually orthogonal irreducible representations : $(H, f) = \oplus_k (H_k, f)$ where H_k are orthogonal subspaces of H. Moreover if G is compact then the H_k are finite dimensional.

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

$\phi : G' \to G :: \forall g, g' \in G' : \phi(g \cdot g') = \phi(g) \cdot \phi(g');$
$\phi(1_{G'}) = 1_G \Rightarrow \phi(g^{-1}) = \phi(g)^{-1}$
$f : G \to G\mathcal{L}(E; E)$
Define $f' : G' \to G\mathcal{L}(E; E) :: f'(g') = f(\phi(g'))$
$f'(g'_1 \cdot g'_2) = f(\phi(g'_1 \cdot g'_2)) = f(\phi(g'_2)) \circ f(\phi(g'_1)) = f'(g'_1) \circ f'(g'_2)$

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.

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

Lie groups

A **Lie group** is a group G endowed with the structure of a manifold (Maths.22). 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 T_1G, endowed with a bracket [] which is a bilinear antisymmetric map on T_1G.

The commutation on a group by an element g is the operation : $x \to g \cdot x \cdot g^{-1}$. Its derivative at $x = 1$ is the Adjoint map $Ad_g : T_1G \to T_1G$.

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.

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

A real finite dimensional, semi-simple, *compact* Lie group G has a Killing form which is a bilinear form, definite negative and is preserved by the adjoint map Ad, so its Lie algebra is a Hilbert space, and G has the unitary representation (T_1G, Ad)

Lie algebras of group of matrices are deduced from the standard representation by derivation.

Abelian groups

An abelian group is a commutative group (we will always assume that this is also a topological, finite dimensional Lie group, with unity denoted 0). Any n dimensional abelian Lie group over the field K is isomorphic to the product of groups (with addition) : $(K/\mathbb{Z})^p \times K^{n-p}$. So the representations of abelian groups are modelled on the representations of $((K/\mathbb{Z})^m, +)$ or of $(K^m, +)$. They are both vector spaces but $((K/\mathbb{Z})^m, +)$ is a compact group. Representations of abelian groups are linked to the Fourier transform.

Any irreducible, unitary or finite dimensional, representation (H, f) of an abelian group G is unidimensional. It can be written as :

$H = \{kU, k \in K\}$ for some fixed vector U

$f(g)u = \lambda(g)u$ with : $\lambda(g + g') = \lambda(g)\lambda(g')$

If the representation is unitary :

$\langle f(g)u, f(g)v\rangle = \langle u, v\rangle = \overline{\lambda(g)}\lambda(g)\langle u, v\rangle \Leftrightarrow \overline{\lambda(g)}\lambda(g) = 1 \Leftrightarrow \lambda(g) \in U(1) = \{\exp ix, x \in \mathbb{R}\}$

So a unitary irreducible representation is parametrized by a map : $\chi : G \to U(1)$ and the choice of a vector U

$\widehat{G}$ (called the Pontryagin dual) is the set of continuous morphisms

$\chi : G \to U(1)$ such that $\chi(g + g') = \chi(g)\chi(g'), \chi(0) = 1$

$\widehat{G}$ is fixed by G. This is a discrete group if G is compact, a finite group if G is finite.

This picture is generalized for any unitary representation, irreducible or not, by using spectral integrals.

Practically, for vector spaces of maps defined over $\mathbb{R}$,we have two cases.

i) Periodic maps : V is a vector space of periodic maps $X : \mathbb{R} \to E :: X(t + T) = X(t)$ where T is a fixed scalar, E a vector space endowed with a definite positive scalar product. Then X is defined by a Fourier series : it belongs to the Hilbert space H with basis : $\{\varepsilon_i \exp iz\varpi t\}_{i\in I}^{z\in\mathbb{Z}}$ where $(\varepsilon_i)_{i\in I}$ is an orthonormal basis of E, and

$X(t) = \sum_{z\in\mathbb{Z}} \widehat{X}(z) \exp iz\varpi t$ for a set $\left(\widehat{X}(z)\right)_{z\in\mathbb{Z}}$ of fixed vectors of E :

$\widehat{X}(z) = \frac{1}{T}\int_0^T X(t) \exp(-iz\varpi t)\, dt$

If $X(t)$ is real valued, then the formula still holds, with the additional condition :

$\forall n \in \mathbb{N} : \widehat{X}(-n) = \overline{\widehat{X}(n)}$

The scalar product is : $\langle X, Y\rangle_H = \frac{1}{T}\int_0^T \langle X(t), Y(t)\rangle_E\, dt = \left\langle \widehat{X}, \widehat{Y}\right\rangle = \sum_{z\in Z}\left\langle \widehat{X}(z), \widehat{Y}(z)\right\rangle_E$

ii) V is a vector space of maps $X : \mathbb{R} \to E$ where E is a vector space endowed with a definite positive scalar product, which is *globally* invariant by the operation : $f(\theta) : V \to V : f(\theta)(X)(t) = X(t + \theta)$ for *any* θ fixed. Then necessarily $V \subset L^2(\mathbb{R}, dt, E)$ which is a Hilbert space with scalar product :

$\langle X, Y\rangle = \int_{\mathbb{R}} \langle X(t), Y(t)\rangle_E\, dt$

and we can use all the properties of the Fourier transform.

If $X \in L^1(\mathbb{R}, dt, E)$ then $\mathcal{F}(X)(x) = \widehat{X}(x) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} X(t) \exp(-itx)\, dt$ [4].

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 value of g leads to redefine X as in the first way, but this requires an 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.

[4] The expression of the Fourier integral depends on the authors. On the properties of the Fourier transform see Maths.31.

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

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 $U\left(g^{-1}\right) = U(g)^{-1}$

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

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 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. 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 $\left(\widetilde{\varepsilon}_i\left(g\right)\right)_{i\in J}$ of H_Φ comprised of a finite number of vectors which are eigen vectors of $\widehat{\Phi}\circ\widehat{U}\left(g\right):\widehat{\Phi}\circ\widehat{U}\left(g\right)\left(\widetilde{\varepsilon}_j\left(g\right)\right)=\lambda_j\left(g\right)\widetilde{\varepsilon}_j\left(g\right)$

Any vector of H_Φ reads :

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

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

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

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

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

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

That we can write :

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

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

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

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

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

$$\left[\begin{array}{ccccccccccccc}
 & & U\left(g\right) & & & & \Phi & & & & & & \\
V & \rightarrow & \rightarrow & \rightarrow & V & \rightarrow & \rightarrow & \rightarrow & W & & & & \\
\downarrow & & & & \downarrow & & & & \downarrow & & & & \\
\downarrow & \Upsilon & & \Upsilon & \downarrow & & & \Upsilon & \downarrow & & & & \\
\downarrow & & \widehat{U}\left(g\right) & & \downarrow & & \widehat{\Phi} & & \downarrow & & \widehat{U}\left(g\right) & & \\
H & \rightarrow & \rightarrow & \rightarrow & H & \rightarrow & \rightarrow & \rightarrow & H_\Phi & \rightarrow & \rightarrow & \rightarrow & H_\Phi \\
 & & & & & & & & \downarrow & & & & \downarrow \\
 & & & & & & & \phi & \downarrow & & & \phi & \downarrow \\
 & & & & & & & & \downarrow & & f\left(g\right) & & \downarrow \\
 & & & & & & & & F & \rightarrow & \rightarrow & \rightarrow & F
\end{array}\right]$$

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

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

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

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

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

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

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

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

$X^j\left(g\right)=\sum_{k\in J}\left[f\left(g\right)\right]_j^kX^j\left(1\right)$

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

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

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

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

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

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

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.

i) We have seen that unitary representations of an abelian group are isomorphic to some special classes of maps, so usually the specifications of the variables can be deduced.

ii) An observable is the choice of a specification, that is the choice of a vector subspace of maps, depending on a finite number of parameters, which fixes the dimension of this vector space V_0. If we are in the conditions of the Theorem 22 then it makes sense to look for an irreducible representation. Indeed, if the representation is reducible, then, for all the possible values of g, the value of the observable belongs to a vector subspace of V_0, meaning that the specification of the variables requires fewer parameters.

So, in the conditions of the theorem, we can assume that **an observable belongs to an irreducible representation.**

iii) A continuous unitary representation $\left(H, \widehat{U}\right)$ can be decomposed on the sum $\oplus_k \left(H_k, \widehat{U}\right)$ of orthogonal irreducible representations so, in a continuous process, the system stays in states belonging to one of the irreducible representation H_k : a change $H_k \to H_j$ implies a discontinuous process, and this holds for X.

iv) If G is compact or finite then the H_k are finite dimensional. In each irreducible representation the variables are characterized by a finite number of parameters.

v) 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.

vi) In the v),vi) ot the theorem the nature of the space F in the equivalent representation (F, f) does not matter, only the matrices $[f(g)], [K]$.

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.

One parameter groups

An important case 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 variables must be defined over a domain which is invariant by the translation $t \to t+\theta$, so it must be the totality of $\mathbb{R}$, and not just an interval.

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 $\lim_{t\to 0} \left\| \widehat{U}(t) - Id \right\| = 0$), 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 t.

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.12.3.4). 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. One 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 is valued in a manifold N. 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 comprises :*

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, for each fixed value of Y the subset $[H]_y$ is invariant by the action of G, so it provides an irreducible representation, as well as $[F]_y$. **The observables belong to finite dimensional irreducible representations characterized by the value of** Y.

We have seen in the Theorem 6 that one can replace a discrete variable by a continuous function $f : V \to [0, 1]$ such that $f(X) = 1$ for $X \in O_1, f(X) = 0$ for $X \in O_2$. Then, in the conditions of the theorem above, the Hilbert space $H = H_1 \oplus H_2$ where H_1, H_2 are associated to each value of the discrete variable.

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) $d(X,0)_V \to 0 \Rightarrow \forall t \in R : \|X(t)\|_E \to 0$ and the condition ii) is met, but this is not a general result.

This condition is met if there is a solution which is not static : $\forall t \neq t' \in R, \exists X \in V : X(t) \neq X(t')$

Proof. The family of maps X is separating, the weak topology 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 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 $\Theta(t) \in \mathcal{L}(F;F)$.

$\widetilde{\Theta}(t)$ 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 the laws for the evolution of the system are such that the initial value $X(0)$ defines, up to a negligible set of points, uniquely $X(t)$.

When other arguments than t are involved this is more complicated. 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 :
$\mathcal{E}(t) = \mathcal{E}(0) \circ \exp tS$
$\forall t \in \mathbb{R} : X(t) = (\exp tS \circ X)(0) = \left(\sum_{n=0}^{\infty} \frac{t^n}{n!} S^n X\right)(0)$
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 $\widehat{U}:\mathbb{R}\rightarrow\mathcal{L}(H;H)$ 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 : $\widehat{U}:\mathbb{R}\rightarrow\mathcal{L}(H;H)$ 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.
$\Theta(t):F(0)\rightarrow F(t)::\Theta(t)u=\mathcal{E}(t)\circ\mathcal{E}(0)^{-1}u=\mathcal{E}(0)\circ\exp tS\circ\mathcal{E}(0)^{-1}u$
The map $\Theta(\theta):F\rightarrow F$ 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.

In the conditions of the last theorem It can be useful to introduce explicitly a complex structure. This can always be done in H. If $\widehat{S}$ is anti-hermitian, then $\widehat{S}' = i\widehat{S}$ is hermitian, and $\widehat{U}(\theta) = \exp\theta\widehat{S} = \exp\left(-i\theta\widehat{S}'\right)$ is the Fourier transform of $\widehat{S}'$. Then $\widehat{U}$ is the solution of the problem : find $\widehat{U}$ such that $-\frac{1}{i}\frac{d}{ds}\widehat{U}(s)|_{s=t} = \widehat{S}'\widehat{U}(t)$ with the initial condition : $\widehat{U}(0) = \widehat{S}'$. This is the usual formulation of the Schrödinger's law.

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 Periodic States

An important point to notice in the previous theorems is that $X, X' = X(t+\theta)$ are different variables, which does not necessarily take the same values. This is what we call equivariance : the variables represent the same state, their values for the same state are related, but not necessarily equal. This is the difference with a symmetry : in a symmetry the values are equal. In the present case two kinds of symmetry can be considered.

The system is in a static state if its state does not change with time. Then the variables do not depend of t, and actually this is of little interest here.

The system is in a periodic state. The variables are periodic with respect to the argument time. It implies that they are defined over $\mathbb{R}$ and $X_k(t+T) = X_k(t)$ for some fixed $T \in \mathbb{R}$. We have seen in the First Chapter that periodic states are the simplest generalization of static states, and the motivation for their study.

If the model is focused on looking for periodic states (that is finding their general properties) then :

i) the variables are defined over $\mathbb{R}$ and the choice of the origin of the time is arbitrary, however the conditions of the theorem 27 are not met : T is a fixed, given quantity.

ii) all variables can be considered as function of t only. Even if X_k depends on other arguments x, actually it is assumed that they take the same value with the periodicity T, so that their value itself if a function of t : $X_k(t,x) = X_k(t, x(t))$. For instance in the study of a star system, the gravitational field depends on the location and is defined everywhere, however what matters is its value at the locations of the planets, which depends only, for a given planet, on t.

Theorem 28 *The model representing the system meets the conditions 1. Moreover :*

i) the variables are maps : $X_k : \mathbb{R} \to E_k$ *where* E_k *is a normed vector space*

ii) the evolution of the system is periodic : $\exists T \in \mathbb{R} : \forall t \in \mathbb{R}, \forall k : X_k(t+T) = X_k(t)$

then :

i) there is a Hilbert space F *such that* $E \subset F$

ii) there is a sequence $\left(\widehat{\Theta}(z)\right) \in \mathcal{L}(F;F)^{\mathbb{Z}}$ *such that :*

$\widehat{\Theta}(0) = Id_F$

$\forall t \in \mathbb{R} : \Theta(t) = \sum_{z\in\mathbb{Z}} \widehat{\Theta}(z) \exp zi\varpi t \in \mathcal{L}(F;F) \Leftrightarrow \widehat{\Theta}(z) = \frac{1}{T}\int_0^T \Theta(t)\exp(-iz\varpi t)\,dt$

iv) for any periodic state the variables are given by :

$\forall t : X(t) = \Theta(t)(X(0))$

Proof. i) The theorem 26 holds. There is a Hilbert space F such that $X : \mathbb{R} \to F$.

Because X is periodic it can be written :

$X(t) = \sum_{z\in\mathbb{Z}} \widehat{X}(z) \exp zi\varpi t$

where $\widehat{X}(z) = \frac{1}{T}\int_0^T X(t)\exp(-iz\varpi t)\,dt$ are fixed vectors of F.

The scalar product on V is : $\langle X, Y\rangle_F = \frac{1}{T}\int_0^T \langle X(t), Y(t)\rangle_E\,dt = \sum_{z\in Z}\left\langle \widehat{X}(z), \widehat{Y}(z)\right\rangle_E$

ii) There is a map : $\Theta : R \to \mathcal{L}(F;F)$ such that $\Theta(t)$ is unitary and, for the admissible states $X \in O \subset V$:

$X(0) \in \widetilde{A} \subset F$

$\forall t : X(t) = \Theta(t)(X(0)) \in F$

$\Rightarrow X(t+T) = \Theta(t+T)(X(0)) \Rightarrow \Theta(t+T) = \Theta(t)$

Θ is a periodic map valued on the Hilbert space $\mathcal{L}(F;F)$. So :

$\Theta(t) = \sum_{z\in\mathbb{Z}} \widehat{\Theta}(z)\exp zi\varpi t$

$\Theta(0) = Id_F \Rightarrow \sum_{z\in\mathbb{Z}} \widehat{\Theta}(z) = Id_F$

where $\widehat{\Theta}(z) = \mathcal{F}(\Theta)(z) = \frac{1}{T}\int_0^T \Theta(t)\exp(-iz\varpi t)\,dt$ are fixed vectors of $\mathcal{L}(F;F)$

iii) $\forall t : X(t) = \Theta(t)(X(0)) \Rightarrow$

$X(t) = \sum_{z\in\mathbb{Z}} \widehat{\Theta}(z)(X(0))\exp zi\varpi t$

$\widehat{X}(z) = \frac{1}{T}\int_0^T \sum_{z\in\mathbb{Z}}\left(\widehat{\Theta}(z)(X(0))\exp zi\varpi t\right)\exp(-iz\varpi t)\,dt = \widehat{\Theta}(z)(X(0))$

iv) $\forall t, \Theta(t)$ is unitary on F

$\langle X(t), Y(t)\rangle_F = \langle X(0), Y(0)\rangle_F = \sum_{z\in Z}\left\langle \widehat{\Theta}(z)(X(0))\exp zi\varpi t, \widehat{\Theta}(z)(Y(0))\exp zi\varpi t\right\rangle_F$

$= \sum_{z\in Z} \left\langle \widehat{\Theta}(z)(X(0)), \widehat{\Theta}(z)(Y(0)) \right\rangle_F$
because : $\sum_{z\in\mathbb{Z}} \widehat{\Theta}(z) = Id_F \Rightarrow \widehat{\Theta}(z)(X(0)) = X(0)$ ■

It is clear that the model is focused on the search for periodic solutions. It can be a restriction of a more general model.

As a result the problem is well posed : the solutions $X(t)$ depend continuously, and even linearly, on the initial conditions. However we are here in the simple description, in the frame of a model, of the system. Usually the model includes some relations between the variables, which restrict the set of possible solutions. In particular, for a model involving the time, the derivatives with respect to the time are part of the description : $D = \frac{dX}{dt}$ is a separate variable. And a condition imposed to the solutions is that :

$D(t) = \sum_{z\in\mathbb{Z}} \widehat{D}_k(z) \exp iz\varpi t$
$X(t) = \sum_{z\in\mathbb{Z}} \widehat{X}_k(z) \exp iz\varpi t$
$D = \frac{dX}{dt} \Rightarrow \widehat{D}_k(z) = \frac{1}{T}\int_0^T \frac{dX}{dt}(t) \exp(-ik\varpi t)\, dt$
$= [X(t)\exp(-ik\varpi t)]_0^T + ik\varpi \frac{1}{T}\int_0^T X(t)\exp(-ik\varpi t)\, dt = ik\varpi \widehat{X}_k(z)$

A condition which is not necessarily met for any value of Θ. But we have another result, which can be seen as an illustration of the theorem 24. From

$\langle X(t), Y(t)\rangle_F = \langle X(0), Y(0)\rangle_F$
$\langle \Theta(t) X(0), \Theta(t) X(0)\rangle_F = \langle X(0), X(0)\rangle_F \Rightarrow \frac{d}{dt}\langle \Theta(t) X(0), \Theta(t) X(0)\rangle_F = 0$
$\left\langle \frac{d}{dt}\Theta(t) X(0), \Theta(t) X(0)\right\rangle_F + \left\langle \Theta(t) X(0), \frac{d}{dt}\Theta(t) X(0)\right\rangle_F = 0$
$\operatorname{Im}\left\langle \Theta(t) X(0), \frac{d}{dt}\Theta(t) X(0)\right\rangle_F = \frac{1}{i}\left\langle \Theta(t) X(0), \frac{d}{dt}\Theta(t) X(0)\right\rangle_F = \frac{1}{i}\left\langle X(t), \frac{d}{dt}X(t)\right\rangle_F$

The energy stored in a system is related to the state of the system and its rate of change, so the quantity $\frac{1}{i}\left\langle X(t), \frac{d}{dt}X(t)\right\rangle_F$ is a good candidate to represent it. And we see that, in a periodic state :

$\frac{1}{i}\left\langle X(t), \frac{d}{dt}X(t)\right\rangle_F = \frac{1}{T}\int_0^T \left\langle X(t), \frac{d}{dt}X(t)\right\rangle_E dt = \frac{1}{i}\sum_{z\in Z}\left\langle \widehat{X}(z), \widehat{D}(z)\right\rangle_E$
$= \frac{1}{i}\sum_{z\in Z}\left\langle \widehat{X}(z), ik\varpi\widehat{X}_k(z)\right\rangle_E = \varpi \sum_{z\in Z} k\left\langle \widehat{X}(z), \widehat{X}_k(z)\right\rangle_E$

$$E = \nu \sum_{z\in Z} 2\pi k \left\langle \widehat{X}(z), \widehat{X}_k(z)\right\rangle_E \tag{2.6}$$

For a given system the energy depends on the initial conditions and the general laws governing the system, but in a periodic state it is proportional to the frequency ν. Equivalently, to each level of energy is associated a specific frequency. Of course we cannot postulate that $\sum_{z\in Z} 2\pi k \left\langle \widehat{X}(z), \widehat{X}_k(z)\right\rangle_E$ is some universal constant, but it is fascinating that we retrieve, in the most general picture, a result which is reminiscent of the Planck's law.

2.5.3 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.

2.5.4 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 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 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 :

$$\begin{bmatrix} \ulcorner & S_1 & \urcorner & & & & \ulcorner & S_2 & \urcorner \\ X_1 & & Z_1 & & & & X_2 & & Z_2 \\ V_1 & \times & W_1 & & & & V_2 & \times & W_2 \\ & \downarrow \Upsilon_1 & & & & & & \downarrow \Upsilon_2 & \\ & \psi_1 & & & & & & \psi_2 & \\ & H_1 & & & & & & H_2 & \\ & & & \ulcorner & S_{1+2} & \urcorner & & & \\ & & & X_1 & & X_2 & & & \\ & & & V_1 & \times & V_2 & & & \\ & & & & & & & & \\ & & & \psi_1 & & \psi_2 & & & \\ & & & H_1 & \times & H_2 & & & \end{bmatrix}$$

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 latter 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 : $\Upsilon_k : V_k \rightarrow H_k :: \Upsilon_k (X_k) = \psi_k = \sum_{i \in I_k} \langle \phi_{ki}, \psi_k \rangle e_{ki}$

a positive kernel : $K_k : V_k \times V_k \rightarrow \mathbb{R}$

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 $\Upsilon' : V' \rightarrow H'$ 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 29 *i) The variable Y can be deduced from the value of X_1, X_2 : there must be a bilinear map : $\Phi : V_1 \times V_2 \rightarrow V'$*

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 :

$\forall X_1, X_1' \in V_1, \forall X_2, X_2' \in V_2 :$

$K' (\Phi (X_1, X_2), \Phi (X_1', X_2')) = K_1 (X_1, X_1') \times K_2 (X_2, X_2')$

We will prove the following :

Theorem 30 *Whenever two systems S_1, S_2 interact, there is a model S encompassing the two systems and meeting the conditions 29 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 : $\varphi : H_1 \times H_2 \rightarrow H' :: \varphi(\psi_1, \psi_2) = \Phi\left(\Upsilon_1^{-1}(\psi_1), \Upsilon_2^{-1}(\psi_2)\right)$ is bilinear. So, by the universal property of the tensorial product, there is a unique map $\widehat{\varphi} : H_1 \otimes H_2 \rightarrow H'$ such that : $\varphi = \widehat{\varphi} \circ \imath$ where $\imath : H_1 \times H_2 \rightarrow H_1 \otimes H_2$ is the tensorial product.

The condition iii) reads :

$\langle \Upsilon_1(X_1), \Upsilon_1(X_1') \rangle_{H_1} \times \langle \Upsilon_2(X_2), \Upsilon_2(X_2') \rangle_{H_2}$
$= \langle (\Upsilon' \circ \Phi(\Upsilon_1(X_1), \Upsilon_2(X_2)), \Upsilon' \circ \Phi(\Upsilon_1(X_1'), \Upsilon_2(X_2'))) \rangle_{H'}$
$\langle \psi_1, \psi_1' \rangle_{H_1} \times \langle \psi_2, \psi_2' \rangle_{H_2} = \langle \varphi(\psi_1, \psi_2), \varphi(\psi_1', \psi_2') \rangle_{H'}$
$= \langle \widehat{\varphi}(\psi_1 \otimes \psi_2), \widehat{\varphi}(\psi_1' \otimes \psi_2') \rangle_{H'}$

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 :
$\langle (\psi_1 \otimes \psi_2), (\psi_1' \otimes \psi_2') \rangle_{H_1 \otimes H_2} = \langle \psi_1, \psi_1' \rangle_{H_1} \langle \psi_2, \psi_2' \rangle_{H_2}$
and then the reproducing kernel is the product of the reproducing kernels.

So we must have : $\langle \widehat{\varphi}(\psi_1 \otimes \psi_2), \widehat{\varphi}(\psi_1' \otimes \psi_2') \rangle_{H'} = \langle \psi_1 \otimes \psi_2, \psi_1' \otimes \psi_2' \rangle_{H_1 \otimes H_2}$ and $\widehat{\varphi}$ must be an isometry : $H_1 \otimes H_2 \rightarrow H'$

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 : $V_1 \otimes V_2 \rightarrow V_1 \times V_2$. 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.[5]

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.13.5) :

$\Upsilon_N (X_1 \otimes ... \otimes X_N) = \Upsilon (X_1) \otimes ... \otimes \Upsilon (X_N)$

The state of the system is then totally defined by the value of tensors S, Ψ, with I^N components.

If $(\widetilde{\varepsilon}_i)_{i\in I}$ is a Hilbertian basis of H then $E_{i_1 ... i_N} = \widetilde{\varepsilon}_{i_1} \otimes ... \otimes \widetilde{\varepsilon}_{i_N}$ is a Hilbertian basis of $\otimes_{s=1}^N H$. The subspaces $\otimes_{s=1}^p H \otimes \widetilde{\varepsilon}_i \otimes_{s=p+2}^N H$ are orthogonal and $\otimes_{s=1}^N H \simeq \ell^2 (I^N)$. The scalar product is defined by linear extension of

$\langle \Psi, \Psi' \rangle_{\mathbf{H}_N} = \langle \psi_1, \psi_1' \rangle_H \times ... \times \langle \psi_N, \psi_N' \rangle_H$

for decomposable tensors : $\Psi = \psi_1 \otimes ... \otimes \psi_N, \Psi' = \psi_1' \otimes ... \otimes \psi_N'$.

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.

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 :

[5] On this point see Haag p.106

Definition 31 *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 :

Theorem 32 *The states Ψ of homogeneous systems belong to an open subset of a subspace $\mathbf{h}$ 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 $\mathbf{h}$ 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. 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 be a finite subset of I with $card(J) \geq N$, H_J the associated Hilbert space, $\widehat{Y}_J : H \to H_J$ 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.

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 **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 **h.** The dimension of **h** his given by the cardinality of $\mathfrak{S}(\lambda^c)$ that is : $\frac{N!}{n_1!...n_p!}$. All the vector spaces **h** of the same conjugacy class (but different vectors $\widetilde{\varepsilon}_i$) have the same dimension, thus they are isomorphic.

v) A basis of **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 **h** with respect to $\otimes_{s=1}^N H$. We have two non equivalent representation :

By symmetric tensors : **h** is then isomorphic to $\odot_{n_1} H_J \otimes \odot_{n_2} H_J ... \otimes \odot_{n_p} H_J$ with the symmetric tensorial product $\odot$ and the space of n order symmetric tensor on H_J is $\odot_n H_J$

By antisymmetric tensors : **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: $\mathbf{h} \to \mathbf{h}'$. 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 tensors $\odot_n H_J$ 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 : $\Psi \in \odot_n H_J \Leftrightarrow S_n (\Psi) = n! \Psi$.

The set of antisymmetric tensors $\Lambda_n H_J$ 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 : $\Psi \in \wedge_n H_J \Leftrightarrow A_n (\Psi) = n! \Psi$

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 33 *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 $\gamma = \widehat{\Phi} \left(\otimes_{n_1} \widetilde{\varepsilon}_{j_1} \otimes_{n_2} \widetilde{\varepsilon}_{j_2} ... \otimes_{n_p} \widetilde{\varepsilon}_{j_p} \right)$ which is measured, and from it $\lambda, (\widetilde{\varepsilon}_{j_k})_{k=1}^p$ are estimated.

No random behavior is assumed for the microsystems. However, formally, one can associate a probability π_j to the event that a microsystem were in the state ε_j. Then 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.

But is clear that no random behavior is assumed from the microsystem : the probability law is related to the - random - choice for a set of the states of the microsystems, under the constraint given by the value of the observable. This is similar to the - random - choice of a primary observable for a specification.

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 34 *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 32 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 $\lambda = \{0 \leq n_p \leq ... \leq n_1 \leq N, n_1 + ...n_p = N\}$ 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 λ : $\Pr(S(t) \in \mathbf{f}_\lambda)$ 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 QM

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 by 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 : $x : \mathbb{R} \rightarrow M$ 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 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. Because the operator is compact, it excludes the usual “observables” of location and position : they are actually (in the common framework) the infinitesimal generators of the translation operators.

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 subset $K \subset I$ 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 29. 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 eventually on measures of lengths and times. The concepts of space and time are at the foundation of theories about the geometry of the physical 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 the 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 an orthonormal frame, or in GR to a map with some coordinates ξ_α, often with additional provisions for "inertial frames", before a complicated discourse about light, and quite often trains for the Relativist picture. 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.

At 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, as long as only the measures done by a single observer are considered.

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 35 *The Universe can be represented by 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).

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 by 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 $T_m M$ does not depend on the chart. With this vector space structure one can define a dual space $T_m M^*$ and holonomic dual bases denoted $d\xi^\alpha$ with : $d\xi^\alpha \left(\partial \xi_\beta\right) = \delta^\alpha_\beta$, and any other tensorial structure (see Maths.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, others 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 two 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 : $V : M \rightarrow TM :: V(m) = \sum_{\alpha=0}^{3} v^\alpha (m) \partial\xi_\alpha$ 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 :

$v^\alpha (m) \rightarrow \widetilde{v}^\alpha (m) = \sum_{\beta=0}^{3} [J(m)]^\alpha_\beta v^\beta (m)$

where $[J(m)] = \left[\frac{\partial\eta^\alpha}{\partial\xi^\beta}(m)\right]$ is a 4×4 matrix called the jacobian

Similarly a one form on M is a map $\varpi : M \rightarrow TM^* :: \varpi(m) = \sum_{\alpha=0}^{3} \varpi_\alpha (m) d\xi^\alpha$ and the components change as :

$\varpi_\alpha (m) \rightarrow \widetilde{\varpi}_\alpha (m) = \sum_{\beta=0}^{3} [K(m)]^\beta_\alpha \varpi_\beta (m)$ 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 : $p : \mathbb{R} \rightarrow M :: m = p(\tau)$ 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 : $p'(\tau) : \mathbb{R} \rightarrow T_{p(\tau)}M :: \frac{dp}{d\tau} \in T_{p(\tau)}L_p$. 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 curve** which goes through a given point. The **flow** of a vector field V is the map :

$\Phi_V : \mathbb{R} \times M \rightarrow M :: \Phi_V (\tau, a)$ such that $\Phi_V (., a) : \mathbb{R} \rightarrow M :: m = \Phi_V (\tau, a)$ is the integral path going through a and $\Phi_V(.,a)$ is a local diffeomorphism :

$$\begin{gathered}\forall\theta\in\mathbb{R}:\frac{\partial}{\partial\tau}\Phi_V(\tau,a)\,|_{\tau=\theta}=V(\Phi_V(\theta,a))\\ \forall\tau,\tau'\in\mathbb{R}:\Phi_V(\tau+\tau',a)=\Phi_V(\tau,\Phi_V(\tau',a))\\ \Phi_V(0,a)=a\\ \forall\tau\in\mathbb{R}:\Phi_V(-\tau,\Phi_V(\tau,a))=a\end{gathered}\tag{3.1}$$

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. 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, 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 we assume 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, whose practical realization does not matter here.

Whenever there is, on a manifold, a map such that f_o, with $f_o'(m) \neq 0$, it defines on M a foliation : 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 foliation one can define a unique function f_o with these properties (Maths.1507[1]). The successions of present "spaces" for any observer is such a foliation, so our representation is consistent. And we state :

Proposition 36 *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 foliation 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 $\varphi(t', x)$.

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

[1] This theorem, which has far reaching consequences, is new and its proof, quite technical is given in my book.

by the matrix of the transformation of a frame), fluids (the speed of material points are given by a vector field). These relations are expressed in phenomenological laws, they are essential in practical applications. The generalization to Relativity of the concept of solids, or material bodies which have a spatial extension, is an important issue that we address in a following section.

We will consider first in this section material bodies which have no internal structures, or whose internal structure can be neglected, that we will call **particles**. The only property that we will consider here for a particle is its location, given by a geometrical point in the universe. A particle then can be an electron, a nucleus, a molecule, or even a star system, according to the scale of the study. We will add other properties to particles in the following chapters.

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 :

Proposition 37 *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 :*

$$\begin{array}{l} u(\theta) = \frac{dp}{d\tau}|_{\tau=\theta} \in T_{p(\theta)}M \\ p(\tau) = \Phi_u(\tau, a) \;\; with \; a = \Phi_u(0, a) = p(0) \end{array} \tag{3.4}$$

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 38 *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.

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 foliation $\Omega_3(t)$. 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 of any particle. 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 locations described previously, and

$m = \varphi_o\left(t, \xi^1, \xi^2, \xi^3\right)$ is a chart of the area $\Omega \subset M$ spanned by the $\Omega_3(t)$. Each point $m(t) = \varphi_o\left(t, \xi^1, \xi^2, \xi^3\right)$ 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 a 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.

An observer is not necessarily spatially immobile. But to know his new location he has to proceed to measures which are similar to setting up a chart, with similar protocols, so actually this is a change of chart and it is managed by the relations between old and new coordinates. In order to keep it simple, in this book we assume that the standard chart is a chart for an observer who is spatially immobile, and the motion of an observer is a change of observer.

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 their 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 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 relation 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).

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 $\alpha u + (1-\alpha)v$ for $\alpha \in]0,1[$ 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 consisting 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 comprises 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 :

$$\begin{aligned} &\text{signature } (3,1) : \langle u, v \rangle = u^1v^1 + u^2v^2 + u^3v^3 - u^0v^0 \\ &\text{signature } (1,3) : \langle u, v \rangle = -u^1v^1 - u^2v^2 - u^3v^3 + u^0v^0 \end{aligned} \tag{3.6}$$

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 39 *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. 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 40 $\langle \varepsilon_i, \varepsilon_j \rangle = \eta_{ij}$ *for the signature (3,1) and* $\langle \varepsilon_i, \varepsilon_j \rangle = -\eta_{ij}$ *for the signature (1,3)*

with the matrix $[\eta]$

Notation 41 *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 42 $[\eta] = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ *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**. These maps are represented by a matrix $[\chi]$ such that :

$$\boxed{[\chi]^t [\eta] [\chi] = [\eta]} \tag{3.7}$$

The group denoted equivalently $O(3,1)$ or $O(1,3)$, does not depend on the signature (replace $[\eta]$ by -$[\eta]$). (Maths.24.5). $O(3,1)$ is a 6 dimensional Lie group with Lie algebra $o(3,1)$ whose matrices $[h]$ are such that :

$$[h]^t [\eta] + [\eta] [h] = 0 \tag{3.8}$$

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 :

$$\exp[K(w)] = I_4 + \frac{\sinh\sqrt{w^t w}}{\sqrt{w^t w}} K(w) + \frac{\cosh\sqrt{w^t w} - 1}{w^t w} K(w)K(w)$$

$$\exp[K(w)] = \begin{bmatrix} \cosh\sqrt{w^t w} & w^t \frac{\sinh\sqrt{w^t w}}{\sqrt{w^t w}} \\ w\frac{\sinh\sqrt{w^t w}}{\sqrt{w^t w}} & I_3 + \frac{\cosh\sqrt{w^t w}-1}{w^t w} w w^t \end{bmatrix}$$

$$\exp[J(r)] = I_4 + \frac{\sin\sqrt{r^t r}}{\sqrt{r^t r}} J(r) + \frac{1-\cos\sqrt{r^t r}}{r^r r} J(r)J(r) = \begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix}$$

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 : $\chi : R \to SO(3,1)$ such that at each point of the path $p_o : R \to M$ defined on a interval R 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 $\{[\chi(\tau)], t \in R\}$, image of the connected interval R 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 43 ***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"). 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 44 *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 $SO(3,1)$ 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.

Killing vector fields are vector fields V such that their flow, which is always a diffeomorphism, preserves the metric : it is an isometry. We will use Killing vector fields in the Chapter 5.

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 45 *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 46 *For any observer there is a vector field $\varepsilon_0 \in \mathfrak{X}(TM)$ which is future oriented, with length $\langle \varepsilon_0(m), \varepsilon_0(m) \rangle = -1$, normal to $\Omega_3(t)$ and such that : $\varepsilon_0(p_o(t)) = \frac{1}{c}\frac{dp_o}{dt}$ 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 : $\forall v \in T_m\Omega_3(t) : f_o'(m)v = 0$, and for any future oriented vector $v \in T_mM : f_o'(m)v > 0$

Using the metric, it is possible to associate to $f_o'(m)$ a vector : $gradf_o : \langle gradf_o, v \rangle = f_o'(m)v$ which is unique up to a scalar. Thus $gradf_o$ is normal to $\Omega_3(t)$. It is time like. At each point m it is possible to define a vector $\varepsilon_0(m) = \lambda(m)\, gradf_o(m)$ such that $\varepsilon_0(m)$ is future oriented and with length $\langle \varepsilon_0(m), \varepsilon_0(m) \rangle = -1$

$\frac{dp_o}{dt}$ is orthogonal to $\Omega_3(t)$ thus $\frac{dp_o}{dt} = \mu\varepsilon_0(p_o(t))$ and $\left\langle \frac{dp_o}{dt}, \frac{dp_o}{dt} \right\rangle = -c^2 \Rightarrow \frac{dp_o}{dt} = c\varepsilon_0(p_o(t))$ ■

As a consequence :

Theorem 47 $\Omega_3(t)$ *are space like hypersurfaces, with unitary, future oriented, normal* $\varepsilon_0 \in \mathfrak{X}(TM)$

The vector field ε_0, deduced from the function f_o, characterizes the observer. His worldline is $p_o(t) = \Phi_{c\varepsilon_0}(t, p_o(0))$ and the flow of the vector field $c\varepsilon_0$ is a diffeomorphism $\Omega_3(0) \to \Omega_3(t)$:

$\forall x \in \Omega_3(0) : \Phi_{c\varepsilon_0}(t, x)$

Let $\varphi_\Omega : \mathbb{R}^3 \to \Omega_3(0) :: x = \varphi_\Omega(\xi^1, \xi^2, \xi^3)$ be a chart of $\Omega_3(0)$

Denote : $\xi^0 = ct$

Then the map : $\varphi_o : \mathbb{R}^4 \to \Omega :: \varphi_o(\xi^0, \xi^1, \xi^2, \xi^3) = \Phi_{c\varepsilon_0}(t, x)$ is a chart of Ω that we will call the **standard chart** associated to the observer ε_0.

For any fixed point $y = \varphi_\Omega(\eta^1, \eta^2, \eta^3) \in \Omega_3(0)$

$q(t) = \varphi_o(ct, \eta^1, \eta^2, \eta^3)$

$\frac{dq}{dt} = \sum_{\alpha=0}^{3} \frac{\partial \eta^\alpha}{\partial t} \partial\xi_\alpha = \frac{dct}{dt} \partial\xi_0(q(t)) = c\partial\xi_0(q(t))$

$\frac{dq}{dt} = \frac{\partial}{\partial t} \Phi_{c\varepsilon_0}(t, y) = c\varepsilon_0(q(t))$

thus $\partial\xi_0(m) = \varepsilon_0(m)$

And we will write often $m = \varphi_o(ct, \xi)$ where $\xi = (\xi^1, \xi^2, \xi^3)$ are the coordinates in $\Omega_3(0)$

$$\begin{gathered} \varphi_\Omega : \mathbb{R}^3 \to \Omega_3(0) :: x = \varphi_\Omega(\xi^1, \xi^2, \xi^3) \\ \varphi_o : \mathbb{R}^4 \to \Omega :: \varphi_o(ct, \xi^1, \xi^2, \xi^3) = \Phi_{c\varepsilon_0}(t, x) \\ \partial\xi_0(m) = \varepsilon_0(m) \\ \xi^0 = ct \end{gathered} \tag{3.10}$$

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 ε_0 defines a special chart, but also the system itself. Two observers who do not share the vector field ε_0 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 ε_0 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 :

$q(t) = \varphi_o(\xi^0(t), \xi^1(t), \xi^2(t), \xi^3(t))$

$q(t) \in \Omega_3(t) \Leftrightarrow \xi^0(t) = ct$

By differentiation with respect to t :

$\frac{dq}{dt} = \sum_{\alpha=0}^{3} \frac{\partial \xi^\alpha}{\partial t} \partial\xi_\alpha = c\partial\xi_0(q(t)) + \sum_{\alpha=1}^{3} \frac{\partial \xi^\alpha}{\partial t} \partial\xi_\alpha = c\varepsilon_0(q(t)) + \overrightarrow{v}$

$\overrightarrow{v} = \sum_{\alpha=1}^{3} \frac{d\xi_\alpha}{dt} \partial\xi_\alpha \in T_{q(t)}\Omega_3(t)$ so is orthogonal to $\varepsilon_0(q(t))$

Definition 48 *The* ***spatial speed*** *of a particle on its trajectory $q(t) = \varphi_o(ct, x(t))$ with respect to an observer is the vector of $T_{q(t)}\Omega_3(t)$:*

$\overrightarrow{v} = \varphi_o'\left(\sum_{\alpha=1}^{3} \frac{d\xi^\alpha}{dt} \partial\xi_\alpha\right)$

Thus for any particle in the standard chart of an observer :

$$V(t) = \frac{dq}{dt} = c\varepsilon_0 + \overrightarrow{v} \tag{3.11}$$

For the observer in the standard chart we have : $x = Ct \Leftrightarrow \overrightarrow{v} = 0$

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 :*

$$\frac{d\tau}{dt} = \sqrt{1 - \frac{\|\overrightarrow{v}\|^2}{c^2}} \tag{3.12}$$

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 :

$$\frac{dp}{d\tau} = \frac{1}{\sqrt{1 - \frac{\|\overrightarrow{v}\|^2}{c^2}}} \left(\overrightarrow{v} + c\varepsilon_0(m)\right) \tag{3.13}$$

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 $\Phi_{c\varepsilon_0}(t, x)$ of the observer O its trajectory is :

$q : \mathbb{R} \to M :: m = q(t) = \Phi_{c\varepsilon_0}(t, x(t))$

So there is a relation between t, τ :

$m = p(\tau) = \Phi_u(\tau, a) = q(t) = \Phi_{c\varepsilon_0}(t, x(t))$

By differentiation with respect to t :

$\frac{d}{dt}q(t) = c\varepsilon_0(q(t)) + \overrightarrow{v}$

$\frac{dq}{dt} = \frac{dp}{d\tau}\frac{d\tau}{dt} = u\frac{d\tau}{dt}$

$\left\langle \frac{dp}{d\tau}, \frac{dp}{d\tau} \right\rangle = -c^2$

$\left\langle \frac{dq}{dt}, \frac{dq}{dt} \right\rangle = -c^2 \left(\frac{d\tau}{dt}\right)^2$

$\left\langle \frac{dq}{dt}, \frac{dq}{dt} \right\rangle = \langle \overrightarrow{v}, \overrightarrow{v} \rangle_3 - c^2$ because $\varepsilon_0(m) \perp \Omega_3(t)$

$\|\overrightarrow{v}\|^2 - c^2 = -c^2 \left(\frac{d\tau}{dt}\right)^2$

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 :

$\frac{dp}{d\tau} = \frac{dq}{dt}\frac{dt}{d\tau} = \frac{1}{\sqrt{1 - \frac{\|\overrightarrow{v}\|^2}{c^2}}}\left(\overrightarrow{v} + c\varepsilon_0(m)\right)$ ■

As a consequence :

$$\|\overrightarrow{v}\|_3 < c \tag{3.14}$$

$V(t) = \frac{dq}{dt}$ 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 $V(t) = \frac{dq}{dt} = \frac{dp}{d\tau}\frac{d\tau}{dt} = u\sqrt{1 - \frac{\|\overrightarrow{v}\|^2}{c^2}}$ by a scalar.

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. 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. It is a vector, which is an intrinsic property of material bodies and particles. It is measured in bases, and the value of its components vary according to precise mathematical rules when one goes from one basis to another. The physical quantity is absolute, but its measure is relative. And this holds for all physical quantities : a measure in itself has no meaning if one does not know how it has been done, the units and the standards used. It is specially important in Relativity because the observers are not interchangeable. Meanwhile in Classic Physics we are used to some universal frame there is nothing equivalent in General Relativity, starting with the vectorial basis $(\partial \xi_\alpha)$ which varies with the location.

The most general mathematical tool to deal with this problem is the fiber bundle, which is a generalization of the concept of vector space tangent to a manifold.

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 : $\pi_P : P \to M$ 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)[2]. A **trivialization** is a map :

$\varphi_P : M \times F \to P :: p = \varphi_P(m, v)$

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 : $v : M \to F$ and $\mathbf{p}(m) = \varphi_P(m, v(m))$. 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.

Vector bundle

If $F = V$ is a vector space then P is a **vector bundle :**

$\varphi_P : M \times F \to P :: \mathbf{X}(m) = (m, \sum_{i=1}^n X_i(m)\varepsilon_i)$

This is a vector of V located at m. The rules in a change of trivialization are such that P has at each point the structure of a vector space :

$w_m = \varphi_P(m, w), w'_m = \varphi_P(m, w'), \alpha, \beta \in \mathbb{R}$:

$\alpha w_m + \beta w'_m = \varphi_P(m, \alpha w + \beta w')$

and one can define a holonomic basis : it is defined by a basis $(\varepsilon_i)_{i \in I}$ of V :

$\varepsilon_i(m) = \varphi_P(m, \varepsilon_i)$

and write :

[2] 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.

$\mathbf{X}(m) = (m, \sum_{i=1}^{n} X_i(m)\varepsilon_i) = \sum_{i=1}^{n} X_i(m)\varepsilon_i(m)$

$(\varepsilon_i)_{i\in I}$ plays the same role as the holonomic basis $(\partial\xi_\alpha)_{\alpha=0}^{3}$ of the tangent bundle TM.

Usually one requires some property of the basis ε_i, for instance it must be orthonormal. The mean to go from one basis to another is provided usually by the action of a group. So the vector bundles that we will meet are defined as associated to a principal 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) : the **standard gauge,** the element of the fiber bundle such that : $\mathbf{p}(m) = \varphi_P(m, 1)$. *This is not a section* : 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 :

$\cdot : P \times G \to P :: p \cdot g' = \varphi_P(m, g) \cdot g' = \varphi_P(m, g.g')$

such that $(p \cdot g') \cdot g" = p \cdot (g'.g")$

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 $\mathbf{p}(m) = \varphi_P(m, 1)$.

A change of trivialization (that we call a change of gauge) is induced by a map : $\chi : M \to G$ that is by a section $\chi \in \mathfrak{X}(P)$ and :

$\widetilde{\mathbf{p}}(m) = \widetilde{\varphi}_P(m, 1) = \mathbf{p}(m) \cdot \chi(m) = \varphi_P(m, \chi(m))$

$p = \varphi_P(m, g) = \mathbf{p}(m) \cdot g = \widetilde{\mathbf{p}}(m) \cdot \widetilde{g} = (\mathbf{p}(m) \cdot \chi(m)) \cdot \widetilde{g} = \mathbf{p}(m) \cdot (\chi(m).\widetilde{g})$

$\Rightarrow g = \chi(m).\widetilde{g} \Leftrightarrow \widetilde{g} = \chi(m)^{-1}.g$

$\chi(m)$ 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 :

$\sigma \in \mathfrak{X}(P) :: \sigma = \varphi_P(m, \sigma(m)) = \widetilde{\varphi}_P(m, \widetilde{\sigma}(m)) \Leftrightarrow \widetilde{\sigma}(m) = \chi(m)^{-1} \cdot \sigma(m)$

It will be more convenient to define a change of gauge by $\chi(m)^{-1}$ (and not $\chi(m)$) with the obvious adjustments.

$$\begin{gathered} \mathbf{p}(m) = \varphi_P(m, 1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m) \cdot \chi(m)^{-1} : \\ \sigma(m) = \varphi_P(m, g) = \widetilde{\varphi}_P(m, \chi(m) \cdot g) \\ \sigma(m) \to \widetilde{\sigma}(m) = \chi(m) \cdot \sigma(m) \end{gathered} \tag{3.15}$$

So changes of trivialization and change of gauge are the same operations, and we will consider usually a change of gauge.

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 $P[F, \lambda]$ which consists 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:

$$(\mathbf{p}(m), v) = (\varphi_P(m, 1), v) \sim \left(\varphi_P(m, g), \lambda\left(g^{-1}, v\right)\right) \tag{3.16}$$

A standard gauge is nothing more than the use of an arbitrary standard, represented by 1, with respect to which the measure is done. A change of standard gauge in the principal bundle impacts all associated fiber bundles (this is similar to the change of units) :

$$\begin{array}{c} \mathbf{p}(m)=\varphi_P(m,1)\rightarrow\widetilde{\mathbf{p}}(m)=\mathbf{p}(m)\cdot\chi(m)^{-1} \\ v_p=(\mathbf{p}(m),v)=(\widetilde{\mathbf{p}}(m),\widetilde{v}):\widetilde{v}=\lambda(\chi(m),v) \end{array} \tag{3.17}$$

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 :

$$v_m=(\mathbf{p}(m),v)=\left(\mathbf{p}(m),\sum_{i=1}^n v^i\varepsilon_i\right)=\sum_{i=1}^n v^i\varepsilon_i(m) \tag{3.18}$$

A change of standard gauge $\mathbf{p}(m)=\varphi_P(m,1)\rightarrow\widetilde{\mathbf{p}}(m)=\mathbf{p}(m)\cdot\chi(m)^{-1}$ in the principal bundle impacts all associated vector bundles.

For any vector :
$v_m=(\mathbf{p}(m),v)\sim\left(\mathbf{p}(m)\cdot\chi(m)^{-1},\rho(\chi(m))(v)\right)$

Meanwhile the holonomic basis of a vector bundle changes as :
$\varepsilon_i(m)=(\mathbf{p}(m),\varepsilon_i)\rightarrow$
$\widetilde{\varepsilon}_i(m)=(\widetilde{\mathbf{p}}(m),\varepsilon_i)=\left(\mathbf{p}(m)\cdot\chi(m)^{-1},\varepsilon_i\right)$
$\sim\left(\left(\mathbf{p}(m)\cdot\chi(m)^{-1}\right)\cdot\chi(m),\rho\left(\chi(m)^{-1}\right)\varepsilon_i\right)$
$=\left(\mathbf{p}(m),\rho\left(\chi(m)^{-1}\right)(\varepsilon_i)\right)=\rho\left(\chi(m)^{-1}\right)\varepsilon_i(m)$

so that the components of a vector in the holonomic basis change as :
$v_m=\sum_{i=1}^n v^i\varepsilon_i(m)=\sum_{i=1}^n\widetilde{v}^i\widetilde{\varepsilon}_i(m)=\sum_{i=1}^n\widetilde{v}^i\rho(\chi(m))^{-1}\varepsilon_i(m)$
$\Rightarrow\widetilde{v}^i=\sum_j[\rho(\chi(m))]_j^i v^j$

$$\begin{array}{c} \mathbf{p}(m)=\varphi_P(m,1)\rightarrow\widetilde{\mathbf{p}}(m)=\mathbf{p}(m)\cdot\chi(m)^{-1}: \\ v_m=(\mathbf{p}(m),v)\sim\left(\mathbf{p}(m)\cdot\chi(m)^{-1},\chi(m)(v)\right) \\ \varepsilon_i(m)=(\mathbf{p}(m),\varepsilon_i)\rightarrow\widetilde{\varepsilon}_i(m)=\rho(\chi(m))^{-1}\varepsilon_i(m) \\ v^i\rightarrow\widetilde{v}^i=\sum_j[\rho(\chi(m))]_j^i v^j \end{array} \tag{3.19}$$

The set of sections on $P[V,\rho]$, denoted $\mathfrak{X}(P[V,\rho])$, is an infinite dimensional vector space. $(\mathfrak{X}(P[V,\rho]),\rho)$ is an infinite dimensional representations of the group G. 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.

An important point : even if one denotes $\mathbf{v}(m)=\sum_{i=1}^n v^i(m)\varepsilon_i(m)$ actually the vector is measured in a fixed vector space : $\mathbf{v}(m)=(\varphi_P(m,1),v(m))$ where $v(m)=\sum_{i=1}^n v^i(m)\varepsilon_i\in V$. So that the derivatives : $\partial_\alpha\mathbf{v}(m)=(\varphi_P(m,1),\partial_\alpha v(m))$ with $\partial_\alpha v(m)=\sum_{i=1}^n\left(\partial_\alpha v^i(m)\right)\varepsilon_i$. The *fiber bundle formalism enables to consider the components independently from the basis.* This is possible because the gauge $\mathbf{p}(m)=\varphi_P(m,1)$ is not a section.

Any Lie Group G has a representation (T_1G,Ad) on its Lie algebra with its adjoint map Ad. So for any principal bundle there is the adjoint bundle $P_G[T_1G,Ad]$ which is a vector bundle, whose holonomic bases are given by bases of T_1G.

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.

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 $P[V,\rho]$:

$$\langle (\mathbf{p}(m), v), (\mathbf{p}(m), v') \rangle_{P[V,\rho]} = \langle v, v' \rangle_W \tag{3.20}$$

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.*

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 $SL(\mathbb{R},4)$). A standard chart defines a fiber bundle :
base $\mathbb{R}$
projection $\pi_o(m) = f_0(m) = t$
fiber $\Omega_3(0) : x = \varphi_\Omega(\xi^1,\xi^2,\xi^3)$
trivialization $m = \varphi(t,x)$
A change of trivialization is a change of chart on $\Omega_3(0)$.

3.2.2 Standard gauge 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 the chart is fixed for a given observer.

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 ε_0 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 arbitrary, 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 $P_o[\mathbb{R}^4,\imath]$ where $(\mathbb{R}^4,\imath)$ 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 50 *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 identity $SO_0(3,1)$, which defines at each point a **standard gauge** : $\mathbf{p}(m) = \varphi_P(m,1)$*

*an **associated vector bundle** structure $P_o[\mathbb{R}^4, \imath]$ where $(\mathbb{R}^4, \imath)$ 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.*

3.2.3 Formulas for a change of observer

Theorem 51 *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 following 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 $SO(3)$:*

$$[\chi] = \begin{bmatrix} \frac{1}{\sqrt{1-\frac{\|v\|^2}{c^2}}} & \frac{\frac{v^t}{c}}{\sqrt{1-\frac{\|v\|^2}{c^2}}} \\ \frac{\frac{v}{c}}{\sqrt{1-\frac{\|v\|^2}{c^2}}} & I_3 + \left(\frac{1}{\sqrt{1-\frac{\|v\|^2}{c^2}}} - 1\right)\frac{vv^t}{\|v\|^2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix} \tag{3.21}$$

Proof. Let be :

O be an observer (this will be main observer) with associated vector field ε_0 , proper time t and world line $p_o(t)$

A be another observer with associated vector field $\mathbf{e}_0$, 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 :

$\frac{dp_A}{d\tau} = c\mathbf{e}_0(m)$ 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 $(\mathbf{e}_i(m))_{i=0}^3$ 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 $\mathbf{e}_0(m)$, 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

$w = \frac{v}{\|v\|}\arg\tanh\left\|\frac{v}{c}\right\| = \frac{1}{2}\frac{v}{\|v\|}\ln\left(\frac{c+\|\overrightarrow{v}\|}{c-\|\overrightarrow{v}\|}\right) \sim \frac{1}{2}\frac{v}{\|v\|}\ln\left(1+2\frac{\|\overrightarrow{v}\|}{c}\right) \simeq \frac{v}{c}$ ∎

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 $(\mathbf{e}_i(m))_{i=0}^3$ 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.

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.

Proposition 52 *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* $\varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i)$..

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. We have defined standard charts and their holonomic bases, which depend on the observer, and are fixed for this observer. And physically one cannot conceive other charts : a physical chart is always the standard chart of some observer.

Standard basis, or standard gauges, are orthonormal, and chosen at any point by the observer. They comprise 4 vectors, called a **tetrad**. The time vector is imposed by the velocity of the observer, but the components of the spatial vectors can be measured in the holonomic basis of a chart.

With the structure of fiber bundle it is possible to compute the impact of a change of gauge We will always assume that a change of ε_0 is a change of observer. 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.

The condition for 4 vectors to be orthogonal depends on the metric, which changes with the location. It is proven in Differential Geometry that there is no chart such that its holonomic basis can be 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 [3]. As a consequence *an orthonormal basis cannot have fixed components in any chart*, even if the observer strives to keep them as fixed as possible. And the components of the tetrad in the - fixed - holonomic chart must change in order to keep the basis orthonormal. For the time being we do not make any assumption about the factors which can explain this varying metric, this will be seen in the Chapter 5.

Tetrad

Definition

Define a vector bundle associated to a principal bundle sums up to define, at each point, a basis associated to the standard gauge $\mathbf{p}_G(m) = \varphi_G(m,1)$:

$\varepsilon_i \in \mathbb{R}^4 \to \varepsilon_i(m) = (\mathbf{p}_G(m), \varepsilon_i)$

that is, practically, to choose an orthonormal basis with vectors $\varepsilon_i(m), i = 0...3$ at the point m. These vectors, called vierbein, constitute the tetrad, and can be expressed in the holonomic basis of any chart :

$$\varepsilon_i(m) = \sum_{\alpha=0}^{3} P_i^\alpha(m)\,\partial\xi_\alpha \Leftrightarrow \partial\xi_\alpha = \sum_{i=0}^{3} P_\alpha'^i(m)\,\varepsilon_i(m) \tag{3.22}$$

where $[P]$ is a real invertible matrix (which has no other specific property, it does not belong to $SO(3,1)$) and we denote

Notation 53 $[P'] = [P]^{-1} = [P_\alpha'^i]$.

The dual of $(\partial\xi_\alpha)_{\alpha=0}^3$ is $(d\xi^\alpha)_{\alpha=0}^3$ with the defining relation :

$d\xi^\alpha(\partial\xi_\beta) = \delta_\beta^\alpha$.

The dual $\left(\varepsilon^i(m)\right)_{i=0}^3$ is :

$$\varepsilon^i(m) = \sum_{i=0}^{3} P_\alpha'^i(m)\,d\xi^\alpha \Leftrightarrow d\xi^\alpha = \sum_{i=0}^{3} P_i^\alpha(m)\,\varepsilon^i(m) \tag{3.23}$$

The quantities $(P_i^\alpha(m))_{i=1}^3$ and $\left(P_\alpha'^i(m)\right)_{i=1}^3$ are one of the variables in any model in GR : as such they replace the metric g.

Change of gauge :

A change of observer is a change of gauge on the principal bundle P_G : $\mathbf{p}(m) = \varphi_P(m,1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m)\cdot\chi(m)^{-1}$ with $[\chi(m)] \in SO_0(3,1)$

The tetrad of the new observer is :

$\varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i) \to \widetilde{\varepsilon}_i(m) = (\widetilde{\mathbf{p}}(m), \varepsilon_i) \sim \left(\mathbf{p}(m), [\chi(m)]^{-1}\varepsilon_i\right) = \sum_{j=0}^3 \left[\chi(m)^{-1}\right]_i^j \varepsilon_j(m)$

$\sum_{\alpha=0}^3 \widetilde{P}_i^\alpha(m)\,\partial\xi_\alpha = \sum_{j=0}^3 \left[\chi(m)^{-1}\right]_i^j P_j^\alpha(m)\,\partial\xi_\alpha$

$\left[\widetilde{P}(m)\right] = [P(m)][\chi(m)]^{-1}$

$$\begin{gathered} \mathbf{p}(m) = \varphi_P(m,1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m)\cdot\chi(m)^{-1} \\ \varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i) \to \widetilde{\varepsilon}_i(m) = \sum_{j=0}^3 \left[\chi(m)^{-1}\right]_i^j \varepsilon_j(m) \\ \left[\widetilde{P}(m)\right] = [P(m)][\chi(m)]^{-1} \end{gathered} \tag{3.24}$$

[3] 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.

Standard chart :

In the Standard Chart the 4th vector is always in the direction of the velocity of the observer. So we have :

$\varepsilon_o(m) = \partial\xi^0(m) \Rightarrow P'^i_0 = \delta^i_0$

$\alpha = 1,2,3 : \frac{\partial}{\partial\xi^\alpha}\varphi_o\left(\xi^0,\xi^1,\xi^2,\xi^3\right) = \partial\xi_\alpha \in T_m\Omega_3(t) \Rightarrow P'^0_\alpha = 0$

and the matrix $[P]$ takes the simpler form :

$[P] = \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix}; [Q] = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix}$

$[P'] = \begin{bmatrix} 1 & 0 \\ 0 & Q' \end{bmatrix}; [Q'] = \begin{bmatrix} P'_{11} & P'_{12} & P'_{13} \\ P'_{21} & P'_{22} & P'_{23} \\ P'_{31} & P'_{32} & P'_{33} \end{bmatrix}$

$[Q][Q'] = I_3$

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 is denoted with upper indexes :

$g^* = \sum_{\alpha\beta} g^{\alpha\beta}\partial\xi_\alpha \otimes \partial\xi_\beta$

and its matrix is $[g]^{-1}$:

$g^{\alpha\beta}(m) = \sum_{ij=0}^3 \eta^{ij}[P]^\alpha_i[P]^\beta_j$

$$[g]^{-1} = [P][\eta][P]^t \Leftrightarrow [g] = [P']^t[\eta][P'] \tag{3.25}$$

It does not depend on the gauge on P_G :

$[\widetilde{g}] = \left[\widetilde{P'}\right]^t[\eta]\left[\widetilde{P'}\right] = [P']^t\left[\chi(m)^{-1}\right]^t[\eta]\left[\chi(m)^{-1}\right][P'] = [P']^t[\eta][P']$

The metric is the physical part of the Geometry of the universe. It imposes a constraint to the choice of the vectors of the tetrad, as the relations above show. The tetrad is defined up to a matrix of $SO(3,1)$ and we have seen how to extend this equivalence to different observer. The tetrad and its matrix $[P]$ makes the link between the abstract mathematical structures defined in orthonormal bases, and the physical world defined in a chart.

In the standard chart of the observer : $g^{00} = -1$.

$[g] = \begin{bmatrix} -1 & 0 \\ 0 & [g]_3 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & [Q']^t[Q'] \end{bmatrix}$

$[g]^{-1} = \begin{bmatrix} -1 & 0 \\ 0 & [g]_3^{-1} \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & [Q][Q]^t \end{bmatrix}$

and $[g]_3$ is definite positive.

The metric defines a **volume form** on M. Its expression in any chart is, by definition :

$\varpi_4(m) = \varepsilon_0 \wedge \varepsilon_1 \wedge \varepsilon_2 \wedge \varepsilon_3 = \sqrt{|\det[g]|}d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$[g] = [P']^t[\eta][P'] \Rightarrow \det[g] = -(\det[P'])^2 \Rightarrow \sqrt{|\det[g]|} = \det[P']$

$$\varpi_4 = \det[P']\,d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3 \tag{3.26}$$

Divergence of a vector field

By definition the divergence of a vector field, with respect to a volume form ϖ_4 is the function such that :

$\pounds_V\varpi_4 = div(V)\varpi_4$

So it is related to the metric :

$divV = \frac{1}{\sqrt{-\det g}}\sum_{\alpha=0}^3\partial_\alpha\left(V^\alpha\sqrt{-\det g}\right) = \sum_{\alpha=0}^3\frac{\partial V^\alpha}{\partial\xi^\alpha} + \frac{1}{2}V^\alpha\sum_{\beta\gamma=0}^3 g^{\beta\gamma}\partial_\alpha g_{\beta\gamma}$

Induced metric

The metric on M induces a metric on any submanifold but it can be degenerated.

On hypersurfaces the metric g_3 is non degenerated if the unitary normal n is such that $\langle n, n\rangle \neq 0$. The induced volume form is :

$\mu_3 = i_n \varpi_4 = \det[P'] \, d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3 (n)$

For $\Omega_3(t)$ the unitary normal n is ε_0 , the induced metric is Riemannian and the volume form ϖ_3 is :

$\varpi_3 = i_{\varepsilon_0} \varpi_4 = \det[P'] \, d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3 (\varepsilon_0)$

$= \det[P'] \, d\xi^0 (\varepsilon_0) \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$= \det[P'] \, d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$$\varpi_3 = \det[P'] \, d\xi^1 \wedge d\xi^2 \wedge d\xi^3 \tag{3.27}$$

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 $\xi^0 = ct$.

For a curve C, represented by any path : $p : \mathbb{R} \to C :: m = p(\theta)$ 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 : $q : \mathbb{R} \to M$ 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

$$\varpi_1(t) = \sqrt{-\langle V, V\rangle} dt \tag{3.28}$$

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 :

$\ell_p = \int_{\tau_A}^{\tau_B} \sqrt{-\left\langle \frac{dp}{d\tau}, \frac{dp}{d\tau} \right\rangle} d\tau = \int_{\tau_A}^{\tau_B} c d\tau = c(\tau_B - \tau_A)$

This is an illustration of the idea that all world lines correspond to a travel at the same speed.

Remarks

i) For an observer using his standard chart, the time vector ε_0 is necessarily in the direction of his velocity. However he can choose another vector, using the fiber bundle structure, which sums up to take as tetrad a basis which is linked to another observer. For instance a spatially immobile observer on Earth can choose a chart which is Sun centered, and the associated tetrad. The computations can be done using the formula for a change of chart and gauge, but they require to know the motion of the two observers with respect to each other. So we will usually consider a spatially immobile observer, using his standard chart and standard tetrad.

ii) One can be not comfortable with the tetrad : how do we know what is the tetrad at a given point ? Actually the issue is the same with any holonomic basis $(\partial\xi_\alpha)_{\alpha=0}^{3}$ given by a chart φ. The mathematical definition is clear, and there is a procedure which tells how to build, physically, a basis (by small translations along the coordinates) at any point, but this procedure is accessible only to a physicist located at this point. A tetrad is any orthonormal basis, and it is possible to check that it is orthonormal, but this can be done only locally. There is no way to compare two bases, holonomic or orthogonal, at different points without a special tool. Because we have freedom of gauge one can choose such a tool, and say that the tetrad at m is the image of the tetrad at some point O by a map. But, because the metric is not constant on the manifold, it requires that the map preserves the metric. One can prove that there are such maps (called isometries), and we will even see in the Chapter 4 how to define, physically, convenient isometries, which could be necessary in practical problems. But to establish the more fundamental results, which is the purpose of this book, it is more important to keep the freedom of gauge.

iii) For the same reason the derivatives of the matrix $[P]$ are, a priori, not defined, but the derivatives of $\det P, \det P'$ which come from the metric are well defined. One can also compute the derivatives of the components of tensors expressed in the tetrad, which is then considered as fixed, as we will see.

3.2.5 Symmetries

For any variable X defined over M a symmetry is defined with respect to a map $F : M \to M$. The variable is said to be symmetric if $X(F(m)) = X(m)$.

Notice the difference with the change of variables seen in the 2nd Chapter : m and $F(m)$ represent different *physical* locations, and we must account for the fact that the Universe is not uniform. Its geometry has some specific properties which lead to consider different symmetries.

Spatial and dynamic symmetries

The first property of the Universe is, for any observer, the breakdown of M in 3 dimensional hypersurfaces.

Spatial symmetries :

A spatial symmetry is defined with respect to a map F such that the function $f_o : M \to \mathbb{R}$ of an observer is itself symmetric with respect to the map $F : f_o(F(m)) = f_o(m)$. Then, at any time t, $F(m)$ belongs to the same hypersurface $\Omega_3(t)$. In the standard chart the map F is expressed as : $F(\varphi_o(ct,\xi)) = \varphi_o(ct, f(\xi))$ for some map $f : \mathbb{R}^3 \to \mathbb{R}^3$. And a variable $X : M \to E$ is symmetric with respect to F if $X(\varphi_o(ct,\xi)) = X(\varphi_o(ct, f(\xi)))$. It takes the same value at two different spatial locations, at any time. Because f is a map over $\mathbb{R}^3$ all the usual symmetries (translation, rotations,...) can be implemented easily.

Dynamic symmetries :

Let us denote $\pi_o : M \to \Omega_3(t_0)$ the projection on some fixed hypersurface $\Omega_3(t_0)$ for a given observer. A dynamic symmetry is defined with respect to a map F such that $\pi_o(F(m)) = \pi_o(m)$. In the standard chart the map F is expressed as : $F(\varphi_o(ct,\xi)) = \varphi_o(f(ct_o), \xi)$ for some map $f : \mathbb{R} \to \mathbb{R}$. Then a variable $X : M \to E$ is symmetric with respect to F if $X(\varphi_o(ct,\xi)) = X(\varphi_o(f(ct), \xi))$. Whatever the spatial location, the variable X takes the same value at the times ct and $f(ct)$. The most important dynamic symmetries are periodic maps : $f(t) = t + T$.

Symmetries defined by a group :

In both cases the symmetries are preserved in a change of chart *for the same observer.*

The map F is generally the continuous action of a Lie group $G : F(g) \circ F(g') = F(gg'); F(1) = Id$. Then there are mathematical tools available (Maths.5.7). The action is necessarily free and actually the study of the variable X sums up to the study of its value on the orbits of the action (the sets $G_m = \{F(g)(m), g \in G\}$). Moreover the variables $\widehat{X}(\xi_0, \xi_1, \xi_2, \xi_3)$ and $\widehat{X}'(\eta_0, \eta_1, \eta_2, \eta_3)$ defined with respect to the coordinates in a standard chart represent the same state. If they meet the conditions 20 one can implement the Theorem 22. The observables of X belong to an irreducible representation of the group G. In particular translations, either spatial or periodic, are abelian groups and the variable X can then be studied by Fourier series or integrals (see 2.4.4 in the 2nd Chapter).

Isometries

The most important physical property of the Universe is the existence of the metric g. The question is then the existence of symmetries for the metric. The metric is a tensor, it is defined in the tangent space and to compare the values of g at m and $F(m)$ one uses a special tool : the pull back of g by F. For any differentiable map : $F : M \to M$ the value of its derivative is a linear map : $F'(m) : T_m M \to T_{F(m)} M$. So we can define the pull back $F^* g$ of g by F by :

$\forall u, v \in T_m (F(m)) : F^* g(m)(u, v) = g(F(m))(F'(m) u, F'(m) v)$

and we say that F preserve the metric g if $F^* g = g$. Then the image of an orthonormal basis is an orthonormal basis, and F is an **isometry**.

The question which arises is then : are there isometries for the physical Universe ? Isometries are generated by the flow of Killing vector fields, defined by PDE, and one can say that, usually, for a given metric, isometries do exist. Moreover they constitute a Lie Group of dimension at most 10 (for a manifold of dimension 4).

A variable X is then symmetric with respect to an isometry if $X(F(m)) = X(m)$. And we see that it has a physical meaning only for variables which are defined at any point. This is not the case of variables related to particles, but it holds for force fields, and one can guess that they have a special role in the definition of physical isometries, as we will see in the Chapter 4.

If F is an isometry it is always possible to choose the tetrads such that : $[P(F(m))] = [P(m)]$, all the geometric quantities can then be expressed in the same, fixed, basis.

With the principal bundle P_G we can define more general symmetries, beyond isometries. P_G is not linked to the metric : a section $\sigma \in \mathfrak{X}(P_G)$ does not tell if a basis is orthonormal, it provides just the rules in a change of orthonormal basis (with the Lorentz scalar product) at each point, whatever the metric. σ can be symmetric with respect to a map F even if F is not an isometry. The same action (such as the rotation of the tetrad) is done by σ at m and $F(m)$, whatever the tetrad at m and $F(m)$, or the value of the metric. The symmetry will be independent on the gauge but the quantities are not necessarily symmetric when expressed in a chart.

In SR (or Euclidean Geometry) the metric is constant, but the difference between the 2 kinds of symmetries still holds : an isometry must be such that its jacobian $[F'(m)]$ is an orthogonal matrix : $[F'(m)]^t [\eta] [F'(m)] = [\eta]$.

3.2.6 Spherical charts

This is a frequent case, which can be implemented easily in our framework.

We single out a fixed point $O \in \Omega_3(0)$, and $O(t)$ is just O at the time t of the spatially fixed observer.

We assume the following :

There is a family $\mathcal{P} \subset C_1(\mathbb{R}; \Omega_3(0))$ of spatial paths : $p : \mathbb{R} \to \Omega_3(0)$ such that :

$\forall p \in \mathcal{P}, \forall \rho \neq \rho' : p(\rho) \neq p(\rho')$ there is no loop and each p is a bijection

$p(0) = O(0)$

$\forall x \in \Omega_3(0)$ there is a unique $p \in \mathcal{P}$ such that : $\exists \rho \in \mathbb{R} : p(\rho) = x$

$\frac{dp}{d\rho} = u(\rho) : \langle u(\rho), u(\rho) \rangle_3 = 1$

Thus the paths constitute a grid, centered in $O(0)$, to locate any point in $\Omega_3(0)$. This is what is done practically by an observer.

Then each path can be identified by the value of $u(0) = v$ and we denote $p(v, \rho) = p(\rho) \in \Omega_3(0)$ which is a chart of $\Omega_3(0)$. Each vector v can be identified by its components in any orthonormal basis at $O(0)$. Let us say :

$v = (\cos\phi\cos\theta, \cos\phi\sin\theta, \sin\phi)$ and one can take as coordinates in $\Omega_3(0)$:

$\xi^1 = \rho\cos\phi\cos\theta, \xi^2 = \rho\cos\phi\sin\theta, \xi^3 = \rho\sin\phi$.

The holonomic basis at $x = p(v, \rho)$ is the image of the basis at $O(0)$ by the derivative $p(v, \rho)'|_x$.

The standard chart is given by

$\varphi_o(ct, x) = \varphi_M(ct, \rho\cos\phi\cos\theta, \rho\cos\phi\sin\theta, \rho\sin\phi)$.

A path from $A = \varphi_o(c\tau_0, x_0)$ to $B = \varphi_o(c\tau_1, x_1)$ can be represented by :

$q(\tau) = \varphi_o(c\tau, x(\tau))$

$\frac{dq}{d\tau} = c\varepsilon_0 + \frac{dx}{d\tau}$

and its length is :

$\ell(A,B)=\int_{\tau_0}^{\tau_1}\sqrt{-\left\langle \frac{dq}{d\tau},\frac{dq}{d\tau}\right\rangle_M}d\tau=\int_{\tau_0}^{\tau_1}\sqrt{\left(c^2\tau^2-g_3(\tau,x(\tau))\left(\frac{dx}{d\tau},\frac{dx}{d\tau}\right)\right)}d\tau$

The volume measure ϖ_3 reads :

$\varpi_3(x)=\det[P'(\varphi_o(t,x))]\rho^2|\cos\phi|\,d\rho d\theta d\phi$

thus it still depends on t, but acts on variables whose arguments are defined through ρ,θ,ϕ.

No assumption has been made about the "shape" of $\Omega_3(0)$, just that this is a 3 dimensional manifold defined by the chart.

One can assume more, that there is a *physical* spherical symmetry. The physical part of the Geometry is the metric. So we assume that the metric has a symmetry in the following meaning.

There is an action of $SO(3)$ on the vectors v at $O(0)$ (they are defined in an orthonormal basis at $O(0)$) :

$v\rightarrow[h][v]$

which induces an action on $\mathcal{P}$ and $\Omega_3(0)$: $R_0 : SO(3)\times\Omega_3(0)\rightarrow\Omega_3(0) :: R_0([h])\,p(v,\rho)=p([h][v],\rho)$

which can be extended to M :

$R : SO(3)\times M\rightarrow M :: R([h])\,\varphi_o(t,x)=\varphi_o(t,R_0(g)(x))$

Notice that $SO(3)$ acts only on v and ρ is unchanged.

The geometry is said to be spherically symmetric if R_0 is an isometry. The metric is invariant by R_0 :

$R_0([h])_*\,g_3=g_3$

with the push forward[4] $R_0([h])_*\,g_3$:

$g_3(R_0([h])x)(R_0'([h](x))|_x u_x, R_0'([h](x))|_x v_x)=g_3(x)(u_x,v_x)$

The metric on $\Omega_3(0)$ does not depend on θ,ϕ but still depends on ρ.

Because ε_0 is invariant by this action, it can be extended to M if R_0 is an isometry for any t on $\Omega_3(t)$ .Then the metric, as well as $[P]$, depends only on t,ρ.

A cylindric symmetry can be represented in the same framework : the action is then that of a subgroup of $SO(3)$ with a definite axis, which can be taken as one of the vector of the orthonormal basis in $O(0)$.

If this symmetry applies to the whole system (the symmetry of the metric is a prerequisite) then the variables X which have the coordinates as arguments belong to a unitary representation of $SO(3)$ and the simplest is the trivial one : they depend only on t,ρ.

We are free to choose our charts and gauges. So in a problem one can take a particle as the observer, apply the rules above, then the results can be translated for any observer by applying the rules for a change of observer, using the Principle of Relativity. This is the simplest, and most rigorous, way to compute the EM field created by a charged particle.

3.2.7 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

[4] See Formulas for the definitions of push forward and pull back.

- 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, which consist 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 general results hold and observers can define a standard chart as seen in RG. However this chart is usually not defined by a frame $\left(O, (\varepsilon_i)_{i=0}^3\right)$. Observers can label points which are in their present with their proper time. The role of the function $f_o(m) = t$ is crucial, because it defines the 3 dimensional hypersurfaces $\Omega_3(t)$. 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 $\Omega_3(t)$ 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 : inertial observers are such that their velocity is a constant vector. *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 $\Omega_3(t)$. 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 $\overrightarrow{O(t)\,m} = \sum_{i=1}^3 \xi^i \varepsilon_i$

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 product of a fixed translation and a fixed rotation in the Minkovski space (an element of the Poincaré group) :

$OM = \sum_{i=0}^3 x_i \varepsilon_i$
$AM = \sum_{i=0}^3 \widetilde{x}_i \widetilde{\varepsilon}_i$
$OM = OA + AM = \sum_{i=0}^3 L_i \varepsilon_i + \sum_{i=0}^3 \widetilde{x}_i \widetilde{\varepsilon}_i$
$\widetilde{\varepsilon}_i = \sum_{j=0}^3 [\chi]_i^j \varepsilon_i, [\chi] \in SO(3,1)$

This result holds only for two inertial observers. Usually they are characterized as 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.

A representation which is valid only for the study of bodies in uniform translation is 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 MOTION

So far we have considered only particles, with no internal structure. The concept of a "material point" which occupies a geometric point, that is with no spatial extension, used to be shocking for many physicists. Actually Mechanics is built around the concept of solids, which can be rigid or deformable, but have an extension, and a particle is seen as an infinitesimal small solid. Solids bring a feature additional to their location, they have an "arrangement", which is represented by an orthonormal basis. As a consequence the motion of a solid encompasses not only a change in its location, but also a rotational motion. Motion, translational and rotational, is a purely geometric concept which is measured by geometric protocols. And we are lead to extend these properties to material points, that is particles : they have a location and an attached orthonormal basis.

The Relativist framework requires a new formalism to represent the motion of a material body, but it is useful to remind how this is done in Galilean Geometry.

3.3.1 Motion of a solid in Galilean Geometry

Rotation in Galilean Geometry

The concept of rotation is well defined in Mathematics : this is the operation which transforms the *orthonormal basis* of a vector space into another. From a physical point of view the rotation is the operation which transforms the orthonormal basis of the observer to an orthonormal basis which is attached to the material body : it measures the *arrangement* of the body with respect to the observer.

The operation belongs to the orthogonal group, in Galilean Geometry to $SO(3)$ and is represented by a matrix R. This is a 3 dimensional Lie group of matrices such that $R^tR = I$. Because of this relation the Lie algebra $so(3) = T_1SO(3)$ 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

$$j : \mathbb{R}^3 \to L(\mathbb{R}, 3) :: [j(r)] = \begin{bmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{bmatrix} \tag{3.29}$$

The operator j is very convenient to represent quantities which are rotated [5]. It 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 group $SO(3)$ is compact, thus the exponential is onto and any matrix of $SO(3)$ can be written as :

$\exp[j(r)] = I_3 + \frac{\sin\sqrt{r^tr}}{\sqrt{r^tr}}[j(r)] + \frac{1-\cos\sqrt{r^tr}}{r^tr}[j(r)][j(r)]$

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 is by definition the unique eigen vector of $[g]$ with eigen value 1 and norm 1 in the standard representation of $SO(3)$, it has for components $\begin{bmatrix} r^1 \\ r^2 \\ r^3 \end{bmatrix} / \sqrt{r^tr}$

Similarly one can define the angle θ of the rotation resulting from a given matrix, and $\theta = \sqrt{r^tr}$

[5] It is similar to the Levi-Civitta tensor ϵ but, in my opinion, much easier to use.

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 (the arrangement of a basis with respect to another), and for the motion (the instantaneous rotation around an axis), but they are two distinct concepts and the distinction is essential.

If 2 orthonormal bases (with same origin) are in relative motion, at any time t we have some rotation $R(t) \in SO(3)$ and naturally the instantaneous rotation is defined through the derivative $\frac{dR}{dt}$.

The usual convention is to represent the instantaneous rotational motion by $R(t)^{-1}\frac{dR}{dt} \in so(3)$, which takes as starting point the frame rotated by $R(t)$. Then it can be represented by a single vector : $R(t)^{-1}\frac{dR}{dt} = j(r)$. This choice is not without consequence : in a change of observer, corresponding to $R \rightarrow \widetilde{R} = g \times R$: $R(t)^{-1}\frac{dR}{dt}$ does not change : in Galilean Geometry a rotational motion is observer independent. The instantaneous rotational motion can be assimilated to a rotation with constant axis r and 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 $SO(3)$. 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 $\overrightarrow{r}$ is significant [6]. 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.

Motion of a rigid solid in Galilean Geometry

One can choose any point G, a fixed orthonormal basis $(e_i)_{i=1}^3$ attached to the solid, and represent the arrangement of the rigid solid at a given time as the operation to go from a fixed orthogonal frame $\left(O,(\varepsilon_i)_{i=1}^3\right)$ to $\left(G,(e_i)_{i=1}^3\right)$. It combines a translation D, belonging to the abelian group

[6] In his book "The road to reality" Penrose gives a nice, simple trick with a belt and book to show this fact.

$\mathcal{T}\left(\mathbb{R}^3\right)$ and a rotation $R \in SO(3)$, and belongs to the group of displacement, which is the semi-direct product $\mathcal{T}\left(\mathbb{R}^3\right) \ltimes SO(3)$. The "semi" implies some relations which make the structure of the group of displacements more complicated than the direct product $\mathcal{T}\left(\mathbb{R}^3\right) \times SO(3)$.

The motion (translational and rotational) of a rigid solid is then represented by the derivative of the displacement, or more conveniently by the value $\left(\frac{dD}{dt}, R^{-1}\frac{dR}{dt}\right)$ of the corresponding elements in the Lie algebra $T_1\left(\mathcal{T}\left(\mathbb{R}^3\right) \ltimes SO(3)\right)$, which is not the direct product $\left(\mathcal{T}\left(\mathbb{R}^3\right) \times so(3)\right)$. This is convenient because we can represent the motion by two vectors : $\overrightarrow{v}_G = \frac{d\overrightarrow{OG}}{dt}, r$ such as $R^{-1}\frac{dR}{dt} = [j(r)]$, however the formulas are a bit complicated (as can be seen in the law for the composition of speeds for rotating bodies) because the displacement is not a direct product.

So the representation of the motion of a rigid solid in Galilean Geometry implies :

- the location of G and its speed $\overrightarrow{v}_G = \frac{d\overrightarrow{OG}}{dt}$
- the rotation R of $\left(G, (e_i)_{i=1}^3\right)$ and its instantaneous change $R^{-1}\frac{dR}{dt}$

The motion is defined by 6 scalar parameters, or two 3 dimensional vectors.

Deformable solid

A deformable solid is a material body which keeps some integrity : its material points stay close to each other. It can be conveniently represented as follows.

The body occupies at the time t a compact area $\omega(t) \subset \mathbb{R}^3$. Each material point is identified by its location q at a time $t = 0$. It is assumed that there is a differentiable map : $\phi : \omega(0) \times \mathbb{R} \to \omega(t) :: x = \phi(q,t)$ which gives the location of the material point q at t. The map ϕ is the representation of the continuity of the body.

The orthonormal basis $(\varepsilon_i)_{i=1}^3$ of $\mathbb{R}^3$ at $t = 0$ is transported as : $e_i(q,t) = \phi'_q(q,t)\,\varepsilon_i$ which is usually no longer orthonormal.

By derivation :

$\frac{\partial}{\partial t} e_i(q,t) = \phi''_{qt}(q,t)\,\varepsilon_i = \phi''_{qt}(q,t)\left(\phi'_q(q,t)\right)^{-1} e_i(q,t)$

and the matrix $\gamma = \left[\phi''_{qt}(q,t)\left(\phi'_q(q,t)\right)^{-1}\right]$ is the deformation tensor. It can be decomposed in a symmetric matrix $\frac{1}{2}\left(\gamma + \gamma^t\right) = s$ and an antisymmetric matrix $\frac{1}{2}\left(\gamma - \gamma^t\right) = j(\omega)$ which measures the torsion. s has real eigen values and represents similitudes in the 3 axes (a "dilation"). $j(\omega)$ can be seen as a rotation with vector ω (a "shear"), and the deformation tensor is the sum of a shear $j(\omega)$ and a dilation s.

ϕ defines the manifolds $\omega(t) = \phi(.,t)$ embedded in $\mathbb{R}^3$ endowed with the induced metric : $g_{ij} = \sum_{k=1}^3 \left[\phi'_q(q,t)\right]_i^k \left[\phi'_q(q,t)\right]_j^k$.

The distance between 2 close elements $\delta q \in T_q\omega(0)$ change as $\sqrt{\sum_{ij} g_{ij}(\delta q)^i (\delta q)^j}$

The volume form is $\varpi = \sqrt{\det g}\,\varepsilon_1 \wedge \varepsilon_2 \wedge \varepsilon_3 = \det\left[\phi'_q(q,t)\right] \varepsilon_1 \wedge \varepsilon_2 \wedge \varepsilon_3$ and the volume changes as $\det\left[\phi'_q(q,t)\right]$: the material points which occupy a volume $\varepsilon_1 \wedge \varepsilon_2 \wedge \varepsilon_3$ at $t = 0$ occupy a volume $\left(\det\left[\phi'_q(q,t)\right]\right) \varepsilon_1 \wedge \varepsilon_2 \wedge \varepsilon_3$ at t.

3.3.2 Motion in Relativist Geometry

The Poincaré's group

The usual concept of rigid solid, as material body whose material points are at a constant distance, does not hold any more in the Relativist framework. Experiments show that atoms and subatomic particles have kinematic characteristics which look like rotation, and can be measured by quantities which transform according to the rules of $SO(3)$, with some complications, and this leads to the concept of spin. So one needs to incorporate rotations in Relativity, in a way similar to what is done with solids in Newtonian Mechanics, and this leads naturally to look for 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. In Special Relativity (and also in QTF) a law is deemed covariant if it is equivariant in a change of frame by the Poincaré's group : this is the implementation of the Principle of Relativity in a representation based on orthogonal frames. Assuming that the 4 momentum p is an intrinsic characteristic of particles, it should be equivariant. With the addition of some of the axioms of QM, this leads, by a demonstration due to Wigner's (see Weinberg for a full proof), to a broad classification of particles.

However the use of the Poincaré's group raises several serious 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. It has been considered in GR to use the group of isometries, that is of maps : $f : M \to M$ such that $f'(m) \in \mathcal{L}(T_m M; T_m M)$ preserves the metric. However in Physics, to compare two bases located at different points one does not jump, one follows a path and the path matters : the relativist universe is not isotropic[7]. This amplifies the issue of the spin group and its 2 values ± 1.

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). A displacement introduces a variable (the translation of the origin O of the frame to go from O to A) which has nothing to do in the matter : in the formulas in a change of observers the spatial speed $\overrightarrow{v}$ is the relative speed with respect to a "copy" of the observer who would be at the same location as the body. Indeed an element of the Poincaré's group is defined by 10 parameters (6 for the Lorentz group and 4 for the translation of the origin), meanwhile 6 suffice in Newtonian Mechanics to define the motion of a solid, and there is no reason why Relativity should add 4 parameters.

A group of displacement is not a direct product of groups, but a semi-direct product, and similarly for the Lie algebras. This introduces complications in Newtonian Mechanics which are amplified in Relativity. 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.

The Spin Bundle

Our purpose is to find an efficient way to represent the motion, translational and rotational in the General Relativity framework.

We start from the assumption that, to any material body, whatever its size, is attached a tetrad, that is an orthonormal basis which represents its arrangement with respect to an observer. This is an extension of the concept of particle, with additional physical properties, which must be accounted for in their representation. In some way it gives relief to the Geometry. Many models in theoretical physics involve a universe with more than 4 dimensions, to account for their physical properties such as charges. One could consider to define a material body by 4 coordinates, corresponding to its location, and 6 additional coordinates for their arrangement. However the arrangement has a meaning only locally, and with respect to a special basis : an orthonormal one (this is the only sensible way to represent a rotation). So actually these properties are related to the metric, which is the physical part of the Geometry.

The motion, as it is commonly understood in Physics, is the instantaneous motion, that is with respect to the position (location and arrangement) of the object : this is clear in the rotational motion $[R]^{-1}\left[\frac{dR}{dt}\right]$. So, to be consistent, the definition of the motion should involve the derivative of

[7] All the more so that most of the measures are done through a signal, which propagates along special curves, as we will see in the Chapter5 .

the velocity, that is the spatial acceleration. And, indeed, an observer attached to a material body can measure both a rotational motion and a change in its transversal motion.

To avoid confusion I will strive in the following to stick to :

- location : this is the point in M where is the particle
- arrangement : this is the arrangement of the tetrad of the particle with respect to the tetrad of an observer located at the same point
- position : is the combination of a location and an arrangement
- motion : is the instantaneous change of position with respect to the previous one.

To represent these concepts, we start from 4 facts :

i) We do not need the Poincaré's group : it is defined only in SR and for inertial observers. The origin O of the frame has no physical meaning, the measures should be done at the same location.

ii) The only clear concept of rotation is done by comparing the arrangement of two orthonormal bases, located at the same point. And in the relativist context this requires to consider a group which preserves the Lorentz scalar product.

iii) The right group to consider is the spin group. This holds already in Galilean Geometry, and in Relativist Geometry any observer can distinguish the orientation of the axis of a spatial rotation with respect to his own velocity. The spin groups $Spin\left(3,1\right), Spin\left(1,3\right)$ are isomorphic so on this point the signature does not matter.

iv) The convenient tool to compare orthonormal bases at a point is a principal fiber bundle.

We have already assumed the existence of a principal bundle $P_G\left(M, SO\left(3,1\right), \pi\right)$, so we make the assumption :

Proposition 54 *There is a* ***principal bundle*** $P_G\left(M, Spin_0\left(3,1\right), \pi_G\right)$ *which has for fiber the connected component of the identity of the Spin group, for trivialization the map :*

$\varphi_G : M \times Spin_0\left(3,1\right) \to P_G :: p = \varphi_G\left(m,s\right).$

The ***standard gauge*** *used by observers is* $\mathbf{p}\left(m\right) = \varphi_G\left(m, \mathbf{1}\right)$

A section $\sigma \in \mathfrak{X}\left(P_G\right)$ is defined by a map: $\sigma : M \to Spin\left(3,1\right)$ such that : $\sigma\left(m\right) = \varphi_G\left(m, \sigma\left(m\right)\right)$ and in a change of gauge :

$$\begin{array}{l} \mathbf{p}\left(m\right) = \varphi_G\left(m,1\right) \to \widetilde{\mathbf{p}}\left(m\right) = \mathbf{p}\left(m\right) \cdot \chi\left(m\right)^{-1} : \\ \sigma\left(m\right) = \varphi_G\left(m,\sigma\right) = \widetilde{\varphi}_G\left(m, \chi\left(m\right) \cdot \widetilde{\sigma}\right) : \widetilde{\sigma} = \chi\left(m\right) \cdot \sigma \end{array} \tag{3.30}$$

Motion of two orthonormal bases

Orthonormal bases are defined in the vector bundle associated to P_G. The arrangement of an orthonormal basis $\left(e_i\left(m\right)\right)_{i=0..3}$ is measured with respect to the tetrad $\left(\varepsilon_i\left(m\right)\right)_{i=0..3}$ of an observer by an element $[\chi]$ of P_G located at m.

$e_i\left(m\right) = \left(\varphi_G\left(m,1\right), \sum_{j=0}^{3} \left[\chi\left(m\right)\right]_i^j \varepsilon_j\right)$

The vectors ε_j are fixed. The motion is given by the derivative

$\frac{d}{dt} e_i\left(m\right) = \left(\varphi_G\left(m,1\right), \sum_{j=0}^{3} \left[\frac{d}{dt}\chi\left(m\right)\right]_i^j \varepsilon_j\right)$

and represented by $\left[\frac{d}{dt}\chi\left(m\right)\right]$, that is by the derivatives of the components of $e_i\left(m\right)$ in the fixed basis ε_j : the change of the tetrad (that is $\frac{dP}{dt}$) is not involved. The time axis e_0 is related to the velocity, w is related to the spatial speed $\overrightarrow{v}$, and r to the rotation of the spatial axes. The time vector is necessarily oriented as the velocity. So with a map : $\mathbb{R} \to P_G :: \varphi_G\left(q\left(t\right), \chi\left(t\right)\right)$ the arrangement and the motion can be efficiently represented. The motion depends on two vectors r, w of $\mathbb{R}^3$ and their derivatives. However the relation $[\chi] = [\exp K\left(w\right)]\left[\exp J\left(r\right)\right]$ is not convenient, and the group which is involved is the Spin group and not $SO\left(3,1\right)$. In order to get a good understanding of this representation and more convenient tools, we need to learn more about Clifford Algebras, which are at the root of the Spin groups. This is the topic of the next section.

3.4 CLIFFORD ALGEBRAS

Clifford algebra is a fascinating algebraic structure on vector spaces which is seen in details in Maths.9. The results which will be used in this book are summarized in this section, the proofs are given in the Annex. This mathematical section is long, but it provides many practical tools which are very convenient for the computations in the GR context.

3.4.1 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 :

$$\forall u, v \in F : u \cdot v + v \cdot u = 2 \langle u, v\rangle \tag{3.31}$$

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 55 $(\varepsilon_i)_{i=0}^3$ *is an orthonormal basis with scalar product :* $\langle\varepsilon_i, \varepsilon_i\rangle = \eta_{ii}$

So we have the relation :

$$\varepsilon_i \cdot \varepsilon_j + \varepsilon_j \cdot \varepsilon_i = 2\eta_{ij} \tag{3.32}$$

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 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).

Spin group

Some elements of the Clifford algebra have an inverse for the product, and there are subsets which have a group structure.

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 with an even number of vectors :

$Spin(3,1) = \{u_1 \cdot u_2 ... \cdot u_{2k}, \langle u_p, u_p\rangle = \pm 1, u_p \in F\}$

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** :

$$\mathbf{Ad}_s : Cl(3,1) \to Cl(3,1) :: \mathbf{Ad}_s X = s \cdot X \cdot s^{-1} \tag{3.33}$$

is such that

$$\forall V \in F : \mathbf{Ad}_s V \in F \tag{3.34}$$

and it preserves the scalar product on F :

$$\forall u, v \in F, s \in Pin(3,1) : \langle \mathbf{Ad}_s u, \mathbf{Ad}_s v \rangle_F = \langle u, v \rangle_F \tag{3.35}$$

Moreover :

$$\forall s, s' \in Pin(3,1) : \mathbf{Ad}_s \circ \mathbf{Ad}_{s'} = \mathbf{Ad}_{s \cdot s'} \tag{3.36}$$

$\mathbf{Ad}$ is distributive with respect to the addition and the product :

$\mathbf{Ad}_s (X \cdot Y) = s \cdot X \cdot Y \cdot s^{-1} = s \cdot X \cdot s^{-1} \cdot s \cdot Y \cdot s^{-1} = \mathbf{Ad}_s X \cdot \mathbf{Ad}_s Y$

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 $[h(s)] \in SO(3,1)$.

$v = \sum_{i=0}^3 v^i \varepsilon_i \to \widetilde{v} = \mathbf{Ad}_s v = \sum_{i=0}^3 \widetilde{v}^i \varepsilon_i$

$\widetilde{v}^i = \sum_{j=0}^3 [h(s)]_j^i v^j$

To two elements $\pm s \in Spin(3,1)$ correspond a unique 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 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

As any algebra $Cl(F, \langle\rangle)$ is a Lie algebra with the bracket :

$\forall X, X' \in Cl(F, \langle\rangle) : [X, X'] = X \cdot X' - X' \cdot X$

which is a bilinear, antisymmetric operation (but not associative) with the Jacobi identity :

$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$

The group $Spin(3,1)$ has a Lie algebra $T_1 Spin(3,1)$ which is a subset of the Clifford algebra. Its elements can be written as linear combinations of pairs of elements $\varepsilon_i \cdot \varepsilon_j$.

The map : $\Pi : so(3,1) \to T_1 Spin(3,1)$ is an isomorphism of Lie algebras which reads with any orthonormal basis $(\varepsilon_i)_{i=0}^3$ of F : $\Pi([\kappa]) = \frac{1}{4} \sum_{i,j=0}^3 ([\kappa][\eta])_j^i \varepsilon_i \cdot \varepsilon_j$

so that any element of $T_1 Spin(3,1)$ 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 Π takes a simple form with an adequate ordering of the vectors :

$\Pi([\kappa_1]) = \overrightarrow{\kappa}_1 = \frac{1}{2} \varepsilon_3 \cdot \varepsilon_2,$

$\Pi([\kappa_2]) = \overrightarrow{\kappa}_2 = \frac{1}{2} \varepsilon_1 \cdot \varepsilon_3,$

$\Pi([\kappa_3]) = \overrightarrow{\kappa}_3 = \frac{1}{2} \varepsilon_2 \cdot \varepsilon_1,$

$\Pi([\kappa_4]) = \overrightarrow{\kappa}_4 = \frac{1}{2} \varepsilon_0 \cdot \varepsilon_1,$

$\Pi([\kappa_5]) = \overrightarrow{\kappa}_5 = \frac{1}{2} \varepsilon_0 \cdot \varepsilon_2,$

$\Pi([\kappa_6]) = \overrightarrow{\kappa}_6 = \frac{1}{2} \varepsilon_0 \cdot \varepsilon_3$

where $([\kappa_a])_{a=1}^6$ is the basis of $so(3,1)$ already noticed such that :

$[\kappa] = K(w) + J(r) = \sum_{a=1}^3 r^a [\kappa_a] + w^a [\kappa_{a+3}]$

We will use extensively the convenient (the order of the indices matters) :

Notation 56 *for both* $Cl(3,1), Cl(1,3)$:

$$\upsilon(r,w) = \frac{1}{2}\left(w^1\varepsilon_0\cdot\varepsilon_1 + w^2\varepsilon_0\cdot\varepsilon_2 + w^3\varepsilon_0\cdot\varepsilon_3 + r^3\varepsilon_2\cdot\varepsilon_1 + r^2\varepsilon_1\cdot\varepsilon_3 + r^1\varepsilon_3\cdot\varepsilon_2\right) \tag{3.37}$$

With this notation, whatever the orthonormal basis $(\varepsilon_i)_{i=0}^3$, any element X of the Lie algebras $T_1Spin(3,1)$ or $T_1Spin(1,3)$ reads :

$$X = \upsilon(r,w) = \sum_{a=1}^{3} r^a \overrightarrow{\kappa}_a + w^a \overrightarrow{\kappa}_{a+3} \tag{3.38}$$

The bracket on the Lie algebra reads :
$[\upsilon(r,w), \upsilon(r',w')]$
$= \upsilon(r,w)\cdot\upsilon(r',w') - \upsilon(r',w')\cdot\upsilon(r,w)$
$= \upsilon(j(r)r' - j(w)w', j(w)r' + j(r)w')$
With signature (1,3) :
$[\upsilon(r,w), \upsilon(r',w')] = -\upsilon(j(r)r' - j(w)w', j(w)r' + j(r)w')$

Expression of elements of the spin group

Theorem 57 *The elements of the Spin groups read in both signatures, with the related* $a, (w^j, r^j)_{j=1}^3, b$ *real scalars and* $\varepsilon_5 = \varepsilon_0\cdot\varepsilon_1\cdot\varepsilon_2\cdot\varepsilon_3$

$$\begin{gathered} s = a + \upsilon(r,w) + b\varepsilon_5 \\ a^2 - b^2 = 1 + \tfrac{1}{4}(w^t w - r^t r) \\ ab = -\tfrac{1}{4} r^t w \\ (a + \upsilon(r,w) + b\varepsilon_5)^{-1} = a - \upsilon(r,w) + b\varepsilon_5 \end{gathered} \tag{3.39}$$

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 [8]:
$\sigma_w(t) = \sqrt{1 + \frac{1}{4}w^t w \sinh^2 \frac{1}{2}t\sqrt{w^t w}} + \sinh\frac{1}{2}t\sqrt{w^t w}\,\upsilon(0,w)$
$\sigma_r(t) = \sqrt{1 - \frac{1}{4}r^t r \sin^2 t\frac{1}{2}\sqrt{r^t r}} + \sin t\frac{1}{2}\sqrt{r^t r}\,\upsilon(r,0)$
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 $Spin(1,3)$:
$r" = \frac{1}{2}(j(r)r' - j(w)w') + a'r + ar' + b'w + bw'$
$w" = -\frac{1}{2}(j(w)r' + j(r)w') + a'w + aw' + b'r + br'$

[8] These quite awful formulas show the interest to use the Clifford algebra representation and not the group SO(3,1) itself.

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 :

$\langle \varepsilon_{i_1}\cdot \varepsilon_{i_2}\cdot ...\cdot \varepsilon_{i_n}, \varepsilon_{j_1}\cdot \varepsilon_{j_2}\cdot ...\cdot \varepsilon_{j_n}\rangle = \eta_{i_1 j_1}...\eta_{i_n j_n}$

This scalar product on $Cl(3,1), Cl(1,3)$ has the signature $(8,8)$: it is non degenerate but neither definite positive or negative. It is invariant by **Ad**.

$$\forall w, w' \in Cl(F,\langle\rangle) : \langle \mathbf{Ad}_s w, \mathbf{Ad}_s w'\rangle_{Cl(E,\langle\rangle)} = \langle w, w'\rangle_{Cl(E,\langle\rangle)} \tag{3.40}$$

$(Cl(3,1), \mathbf{Ad})$ is a representation of $Spin(3,1)$ and $(Cl(1,3), \mathbf{Ad})$ a representation of $Spin(1,3)$.

The basis of the Lie algebra is orthogonal.

$$\begin{array}{c} T_1 Spin(3,1) : \langle v(r,w), v(r',w')\rangle_{Cl} = \frac{1}{4}(r^t r' - w^t w') \\ T_1 Spin(1,3) : \langle v(r,w), v(r',w')\rangle_{Cl} = -\frac{1}{4}(r^t r' - w^t w') \end{array} \tag{3.41}$$

Derivatives of translations

As in any Lie group, the translations on $Spin(3,1)$ are :

$L_g h = g\cdot h, R_g h = h\cdot g$

and their derivatives :

$L'_g h : T_h Spin(3,1) \to T_{g\cdot h} Spin(3,1)$

$R'_g h : T_h Spin(3,1) \to T_{h\cdot g} Spin(3,1)$

Because the Lie algebra and the group belong both to the Clifford algebra, these relations take a simple form :

$X_h \in T_h Spin(3,1) : L'_g h(X_h) = g\cdot X_h, R'_g h(X_h) = X_h\cdot g$

And the usual adjoint map of Lie groups :

$Ad_g : T_1 Spin(3,1) \to T_1 Spin(3,1) ::$

$Ad_g X = \frac{d}{dg}\left(g\cdot X\cdot g^{-1}\right)|_{g=1} = L'_g g^{-1} \circ R'_{g^{-1}} 1(X) = R'_{g^{-1}} g \circ L'_g 1(X)$

is just the map **Ad** :

$$Ad_g X = \frac{d}{dg}\left(g\cdot X\cdot g^{-1}\right)|_{g=1} = \mathbf{Ad}_g X \tag{3.42}$$

Moreover the product is well defined for any element of the Clifford algebra, so the identities hold for any X.

The map : $\mathbf{Ad} : Spin(3,1) \to \mathcal{L}(T_1 Spin(3,1); T_1 Spin(3,1))$ itself is differentiable with respect to g.

$$(\mathbf{Ad}_g X)' = \mathbf{Ad}_g\left(X' + \left[g^{-1}\cdot g', X\right]\right) \tag{3.43}$$

$\left(\mathbf{Ad}_{g^{-1}} X\right)' = \mathbf{Ad}_{g^{-1}}\left(\left[X, g'\cdot g^{-1}\right] + X'\right)$

where $g' = \frac{d}{dx} g(x)$ for x belonging to any manifold.

Theorem 58 $\forall X, Y, Z \in Cl(3,1) : \langle X, [Y,Z]\rangle_{Cl} = \langle [X,Y], Z\rangle_{Cl}$

Proof. $\left\langle \mathbf{Ad}_{g(x)} X, \mathbf{Ad}_{g(x)} Y\right\rangle_{Cl} = \langle X, Y\rangle_{Cl}$

Take the derivative with respect to x :

$\left\langle \mathbf{Ad}_g\left(\left[g^{-1}\cdot g', X\right]\right), \mathbf{Ad}_g Y\right\rangle_{Cl} + \left\langle \mathbf{Ad}_g X, \mathbf{Ad}_g\left(\left[g^{-1}\cdot g', Y\right]\right)\right\rangle_{Cl} = 0$

$\left\langle \left[g^{-1}\cdot g', X\right], Y\right\rangle_{Cl} + \left\langle X, \left[g^{-1}\cdot g', Y\right]\right\rangle_{Cl} = 0$

$g^{-1}\cdot g' = Z$

$\langle [Z,X], Y\rangle_{Cl} + \langle X, [Z,Y]\rangle_{Cl} = 0$

$\langle Y, [Z,X]\rangle_{Cl} - \langle X, [Y,Z]\rangle_{Cl} = 0$ ■

3.4.2 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 [9]. 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.

Decomposition of the elements of the Spin group

Let us choose an orthonormal basis of F. It contains one vector such that $\langle \varepsilon_i, \varepsilon_i\rangle = -1$ (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 :

$$Spin\,(3) = \left\{ s_r = \epsilon\sqrt{1 - \frac{1}{4}r^t r} + \upsilon\,(r,0)\,, r \in \mathbb{R}^3, r^t r \leq 4, \epsilon = \pm 1 \right\} \tag{3.44}$$

$Spin(3)$ is a compact group, with 2 connected components. The connected component of the identity consist of elements with $\epsilon = 1$ and can be assimilated to $SO\,(3)$. [10]

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 axes 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. It is characterized by the equivalence relation :

$\forall s, s' \in Spin\,(3,1) : s \sim s' \Leftrightarrow \exists s_r \in Spin(3) : s' = s \cdot s_r$

[9] 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 $SO(4)/SO(1) \simeq SO\,(3)$.

[10] It is formally SO(3) plus +1

Then, 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 $s_w \in SW, s_r \in Spin(3)$:

$$\forall s = a + \upsilon(r,w) + b\varepsilon_5 \in Spin(3,1) : s = \epsilon(a_w + \upsilon(0,w)) \cdot \epsilon(a_r + \upsilon(r,0)) \qquad (3.45)$$

with : $a_r = \sqrt{1 - \frac{1}{4}r^t r}; a_w = \sqrt{1 + \frac{1}{4}w^t w}$

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. The elements $s_r \in Spin(3)$ are coordinated by 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 $s_r \in Spin(3)$.

$\upsilon(r,0)$ is represented in $so(3,1)$ by a matrix $[J(r)]$ and $\upsilon(0,w)$ by a matrix $[K(w)]$. So we replace the cumbersome formula in a change of gauge $[\chi] = \exp[K(w)]\exp[J(r)]$ by $s = s_w \cdot s_r$.with two elements which are simply related to the velocity (by w) and the rotation (by r). *The decomposition depends on the choice of* ε_0.

Decomposition of the elements of the Lie algebra

Similarly we have the same decomposition in the Lie algebra (see Annex). In any orthonormal basis an element of $T_1 Spin(3,1)$ reads :

$X = \upsilon(r,0) + \upsilon(0,w)$ and $\upsilon(r,0) \in T_1 Spin(3), \upsilon(0,w) \in T_1 SW$

The vectors r, w depends on the basis (they are components), however the elements $\upsilon(r,0), \upsilon(0,w) \in T_1 Spin(3,1)$ depend only on the choice of ε_0

$T_1 Spin(3,1) = L_0 \oplus P_0$

L_0, P_0 and the decomposition depend only on the choice of ε_0 and $L_0 = T_1 Spin(3), P_0 \simeq T_1 SW$.

L_0, P_0 are globally invariant by $Spin(3)$, the scalar product is definite (positive or negative) and preserved by $\mathbf{Ad}$, so L_0, P_0 are 3 dimensional Hilbert spaces, and for each choice of ε_0 $(L_0, \mathbf{Ad}), (P_0, \mathbf{Ad})$ are 3 dimensional unitary representations of $Spin(3)$. Then there is a norm on $T_1 Spin(3,1)$.

3.4.3 Change of basis

The operator : $\mathbf{Ad} : Spin(3,1) \times Cl(3,1) \to Cl(3,1) :: \mathbf{Ad}_s X = s \cdot X \cdot s^{-1}$ takes a different matrix form depending on X. See Annex for the computations.

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 $s = a + \upsilon(r,w) + b\varepsilon_5$ the matrix $[h(s)]$ is :

$$[h(s)] = \begin{bmatrix} a^2 + b^2 + \frac{1}{4}(r^t r + w^t w) & aw^t - br^t + \frac{1}{2}w^t j(r) \\ aw - br + \frac{1}{2}j(r)w & a^2 + b^2 + \frac{1}{4}(r^t r + w^t w) + aj(r) + bj(w) + \frac{1}{2}(j(r)j(r) + j(w)j(w)) \end{bmatrix}$$

$[h(s)] \in SO(3,1) : [h(s)]^t [\eta][h(s)] = [\eta]$.

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 $s = a_w + \upsilon(0,w)$

$$[h(s)] = \begin{bmatrix} 2a_w^2 - 1 & a_w w^t \\ a_w w & 2a_w^2 - 1 + \frac{1}{2}j(w)j(w) \end{bmatrix}$$

If $s = a_r + \upsilon(r, 0)$

$[h(s)] = \begin{bmatrix} 1 & 0 \\ 0 & 1 + a_r j(r) + \frac{1}{2} j(r) j(r) \end{bmatrix}$

$[C(r)] = 1 + a_r j(r) + \frac{1}{2} j(r) j(r) \in SO(3)$ and we have :

$[C(r)] = \exp j(\rho) = I_3 + \frac{\sin\sqrt{\rho^t \rho}}{\sqrt{\rho^t \rho}} [j(\rho)] + \frac{1-\cos\sqrt{\rho^t \rho}}{\rho^t \rho} [j(\rho)] [j(\rho)]$

with : $\rho = r \frac{1}{\sqrt{r^t r}} \arccos\left(1 - \frac{1}{2} r^t r\right)$

Expression of the Action $\mathbf{Ad}_s$ on the Lie algebra

The action of $Spin(3,1)$ is :

$Z = \sum_{a=1}^{6} Z_a \overrightarrow{\kappa}_a \rightarrow$

$\widetilde{Z} = \sum_{a=1}^{6} Z_a \mathbf{Ad}_s(\overrightarrow{\kappa}_a) = \sum_{a=1}^{6} Z_a s \cdot (\overrightarrow{\kappa}_a) \cdot s^{-1} = \sum_{a=1}^{6} Z_a \widetilde{\overrightarrow{\kappa}_a} = \sum_{a=1}^{6} \widetilde{Z}_a \overrightarrow{\kappa}_a$

With :

$Z = \upsilon(X, Y) \rightarrow \widetilde{Z} = \upsilon\left(\widetilde{X}, \widetilde{Y}\right)$

$\begin{bmatrix} \widetilde{X} \\ \widetilde{Y} \end{bmatrix} = [\mathbf{Ad}_s] \begin{bmatrix} X \\ Y \end{bmatrix}$

where $[\mathbf{Ad}_s]$ is a 6×6 matrix with $s = a + \upsilon(r, w) + b\varepsilon_5$:

$[\mathbf{Ad}_s] =$

$$\begin{bmatrix} 1 + aj(r) - bj(w) + \frac{1}{2}(j(r)j(r) - j(w)j(w)) & -\left(aj(w) + bj(r) + \frac{1}{2}(j(r)j(w) + j(w)j(r))\right) \\ aj(w) + bj(r) + \frac{1}{2}(j(r)j(w) + j(w)j(r)) & 1 + aj(r) - bj(w) + \frac{1}{2}(j(r)j(r) - j(w)j(w)) \end{bmatrix}$$

$[\mathbf{Ad}_{s \cdot s'}] = [\mathbf{Ad}_s][\mathbf{Ad}_{s'}]$

With $s_w = a_w + \upsilon(0, w)$

$[\mathbf{Ad}_s] = \begin{bmatrix} \left[1 - \frac{1}{2} j(w) j(w)\right] & -[a_w j(w)] \\ [a_w j(w)] & \left[1 - \frac{1}{2} j(w) j(w)\right] \end{bmatrix} = \begin{bmatrix} A & -B \\ B & A \end{bmatrix}$

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 :

$CC^t = C^t C = I_3$

$[\mathbf{Ad}_{s_r}]^{-1} = \left[\mathbf{Ad}_{s_r^{-1}}\right] = \begin{bmatrix} C^t & 0 \\ 0 & C^t \end{bmatrix}$

Change of basis in F

A change of orthonormal basis of F can be expressed by an action of the Spin group :

$s = a + \upsilon(r, w) + b\varepsilon_5 \in Spin(3,1)$: $\varepsilon_i \rightarrow \widetilde{\varepsilon}_i = \mathbf{Ad}_{s^{-1}} \varepsilon_i$

$\widetilde{\varepsilon}_i = \sum_{j=0}^{3} \left[h\left(s^{-1}\right)\right]_i^j \varepsilon_j$

Then the vectors of F and $T_1 Spin(3,1)$ stay the same, but their components change according to the inverse of the operations see above (as it is usual in any vector space).

$v = \sum_{i=0}^{3} v^i \varepsilon_i = \sum_{i=0}^{3} \widetilde{v}^i \widetilde{\varepsilon}_i$ with $\widetilde{v}^i = \sum_{j=0}^{3} [h(s)]_j^i v^j$

$\upsilon(X, Y) \rightarrow \widetilde{\upsilon}\left(\widetilde{X}, \widetilde{Y}\right)$ with $\begin{bmatrix} \widetilde{X} \\ \widetilde{Y} \end{bmatrix} = [\mathbf{Ad}_s] \begin{bmatrix} X \\ Y \end{bmatrix}$

3.4.4 Complex structure on the Clifford algebra

The subspaces L_0, P_0 are crucial in the properties of $T_1Spin(3,1)$, as seen in the notation $\upsilon(r,w)$. The computations can be made easier by defining on $Cl(3,1)$ and $Cl(1,3)$ a complex structure : the set does not change but it is split in a real and an imaginary part. It is convenient to make computations in the Clifford Algebra.

Complex structure

This is done by a linear map such that : $J^2 = -Id$. Then the product iX is defined as $iX = Xi = J(X)$.

Take $J(X) = X\cdot\varepsilon_5$ with $\varepsilon_5 = \varepsilon_0\cdot\varepsilon_1\cdot\varepsilon_2\cdot\varepsilon_3$ then $J^2(X) = X\cdot\varepsilon_5\cdot\varepsilon_5 = -X$. It holds on $Cl(1,3)$ and $Cl(3,1)$.

The distinction between the real and imaginary vector subspaces is done by splitting any orthonormal basis as follows.

$Cl(3,1)$:

$$\begin{bmatrix} \text{real} & \text{imaginary} & \\ E_j & E_j\cdot\varepsilon_5 = iE_j = E_j i & \varepsilon_5\cdot E_j \\ 1 & \varepsilon_5 & \varepsilon_5 \\ \varepsilon_1 & \varepsilon_0\cdot\varepsilon_3\cdot\varepsilon_2 & -\varepsilon_0\cdot\varepsilon_3\cdot\varepsilon_2 \\ \varepsilon_2 & \varepsilon_0\cdot\varepsilon_1\cdot\varepsilon_3 & -\varepsilon_0\cdot\varepsilon_1\cdot\varepsilon_3 \\ \varepsilon_3 & \varepsilon_0\cdot\varepsilon_2\cdot\varepsilon_1 & -\varepsilon_0\cdot\varepsilon_2\cdot\varepsilon_1 \\ \varepsilon_3\cdot\varepsilon_2 & \varepsilon_0\cdot\varepsilon_1 & \varepsilon_0\cdot\varepsilon_1 \\ \varepsilon_1\cdot\varepsilon_3 & \varepsilon_0\cdot\varepsilon_2 & \varepsilon_0\cdot\varepsilon_2 \\ \varepsilon_2\cdot\varepsilon_1 & \varepsilon_0\cdot\varepsilon_3 & \varepsilon_0\cdot\varepsilon_3 \\ \varepsilon_1\cdot\varepsilon_2\cdot\varepsilon_3 & \varepsilon_0 & -\varepsilon_0 \end{bmatrix}$$

$Cl(1,3)$:

$$\begin{bmatrix} \text{real} & \text{imaginary} & \\ E'_j & E'_j\cdot\varepsilon_5 = iE'_j & \varepsilon_5\cdot E_j \\ 1 & \varepsilon_5 & \varepsilon_5 \\ \varepsilon_0\cdot\varepsilon_3\cdot\varepsilon_2 & \varepsilon_1 & -\varepsilon_1 \\ \varepsilon_0\cdot\varepsilon_1\cdot\varepsilon_3 & \varepsilon_2 & -\varepsilon_2 \\ \varepsilon_0\cdot\varepsilon_2\cdot\varepsilon_1 & \varepsilon_3 & -\varepsilon_3 \\ \varepsilon_0\cdot\varepsilon_1 & \varepsilon_2\cdot\varepsilon_3 & \varepsilon_2\cdot\varepsilon_3 \\ \varepsilon_0\cdot\varepsilon_2 & \varepsilon_3\cdot\varepsilon_1 & \varepsilon_3\cdot\varepsilon_1 \\ \varepsilon_0\cdot\varepsilon_3 & \varepsilon_1\cdot\varepsilon_2 & \varepsilon_1\cdot\varepsilon_2 \\ \varepsilon_0 & \varepsilon_1\cdot\varepsilon_2\cdot\varepsilon_3 & -\varepsilon_1\cdot\varepsilon_2\cdot\varepsilon_3 \end{bmatrix}$$

So we have for any vector of the Clifford algebra :

$X = \sum_{j=1}^{8}\left(X^jE_j + Y^jiE_j\right) = \sum_{j=1}^{8}\left(X^j + iY^j\right)E_j = \sum_{j=1}^{8} Z^jE_j$

and the Clifford algebra becomes a 8 dimensional complex vector space $Cl(3,1)_C$. The complex structure does not depend on the choice of a basis : a change of basis is the application of $\mathbf{Ad}_s$ and the operation commutes with the product by ε_5 :

$\varepsilon_5 \to \widetilde{\varepsilon}_5 = \mathbf{Ad}_s\varepsilon_5 = \varepsilon_5$.

$E_j = \varepsilon_p\cdot\varepsilon_q \to \widetilde{E}_j = \mathbf{Ad}_s\varepsilon_p\cdot\mathbf{Ad}_s\varepsilon_q = \mathbf{Ad}_s(\varepsilon_p\cdot\varepsilon_q) = \mathbf{Ad}_sE_j$

$\widetilde{E}_j\cdot\widetilde{\varepsilon}_5 = \mathbf{Ad}_sE_j\cdot\varepsilon_5 = \mathbf{Ad}_s(E_j\cdot\varepsilon_5)$

The conjugate is $\overline{\operatorname{Re}X + i\operatorname{Im}X} = \operatorname{Re}X - i\operatorname{Im}X$.

The complex formalism can be used to represent any element of the Clifford algebra, however we will use it essentially for the elements of the Spin group and the Lie algebra.

Lie algebra

The basis of $T_1Spin(3,1)$ is :

$\overrightarrow{\kappa}_1 = \frac{1}{2}\varepsilon_3 \cdot \varepsilon_2,$
$\overrightarrow{\kappa}_2 = \frac{1}{2}\varepsilon_1 \cdot \varepsilon_3,$
$\overrightarrow{\kappa}_3 = \frac{1}{2}\varepsilon_2 \cdot \varepsilon_1,$
and
$\overrightarrow{\kappa}_4 = \frac{1}{2}\varepsilon_0 \cdot \varepsilon_1 = i\overrightarrow{\kappa}_1$
$\overrightarrow{\kappa}_5 = \frac{1}{2}\varepsilon_0 \cdot \varepsilon_2 = i\overrightarrow{\kappa}_2$
$\overrightarrow{\kappa}_6 = \frac{1}{2}\varepsilon_0 \cdot \varepsilon_3 = i\overrightarrow{\kappa}_3$
So we can write :

$$\upsilon(r,w) = \sum_{a=1}^{3}(r_a + iw_a)\overrightarrow{\kappa}_a = \sum_{a=1}^{3} Z^a \overrightarrow{\kappa}_a = Z \tag{3.46}$$

The product by i commutes with the product of vectors κ_a :
$a = 1,2,3 : \varepsilon_5 \cdot \kappa_a = \kappa_a \cdot \varepsilon_5$
$\Rightarrow i(\kappa_a \cdot \kappa_b) = \kappa_a \cdot \kappa_b \cdot \varepsilon_5 = \kappa_a \cdot \varepsilon_5 \cdot \kappa_b = (i\kappa_a) \cdot \kappa_b = \kappa_a \cdot (i\kappa_b)$
$i(\upsilon(r,w)) = \left(\sum_{a=1}^{3} r_a\kappa_a + w_a\kappa_a \cdot \varepsilon_5\right) \cdot \varepsilon_5$
$= \left(\sum_{a=1}^{3} r_a\kappa_a \cdot \varepsilon_5 - w_a\kappa_a\right) = \sum_{a=1}^{3}(ir_a - w_a)\kappa_a = \sum_{a=1}^{3} i(r_a + iw_a)\kappa_a = \upsilon(-w,r)$
With this complex notation :

$$Z' \cdot Z = -\frac{1}{4}Z^t Z' + \frac{1}{2}j(Z')Z \tag{3.47}$$

Proof. $\upsilon(r',w') \cdot \upsilon(r,w)$
$= \frac{1}{4}(w^t w' - r^t r') + \frac{1}{2}\upsilon(-j(r)r' + j(w)w', -j(w)r' - j(r)w') - \frac{1}{4}(w^t r' + r^t w')\varepsilon_5$
$= \frac{1}{4}\left((\operatorname{Im} Z)^t \operatorname{Im} Z' - (\operatorname{Re} Z)^t \operatorname{Re} Z'\right)$
$+\frac{1}{2}\upsilon(-j(\operatorname{Re} Z)\operatorname{Re} Z' + j(\operatorname{Im} Z)\operatorname{Im} Z', -j(\operatorname{Im} Z)\operatorname{Re} Z' - j(\operatorname{Re} Z)\operatorname{Im} Z')$
$-\frac{1}{4}i\left((\operatorname{Im} Z)^t \operatorname{Re} Z' + (\operatorname{Re} Z)^t \operatorname{Im} Z'\right)$
$= -\frac{1}{4}Z^t Z' + \frac{1}{2}((-j(\operatorname{Re} Z)\operatorname{Re} Z' + j(\operatorname{Im} Z)\operatorname{Im} Z') - j(i\operatorname{Im} Z)\operatorname{Re} Z' - j(\operatorname{Re} Z)i\operatorname{Im} Z')$
$= -\frac{1}{4}Z^t Z' + \frac{1}{2}(-j(Z)\operatorname{Re} Z' - j(Z)i\operatorname{Im} Z')$ ■
The bracket reads in $Cl(3,1)$:

$$[\upsilon(r,w), \upsilon(r',w')] = j(Z)Z' \tag{3.48}$$

Proof. $[\upsilon(r,w), \upsilon(r',w')] = \upsilon(j(r)r' - j(w)w', j(w)r' + j(r)w')$
$= j(r)r' - j(w)w' + i(j(w)r' + j(r)w')$
$= j(r)r' + j(iw)iw' + j(iw)r' + j(r)iw'$
$= j(r)(r' + iw') + j(iw)(iw' + r')$ ■
and in $Cl(1,3) : [\upsilon(r,w), \upsilon(r',w')] = -j(Z)Z'$
In any Lie algebra the bracket is an antisymmetric bilinear map, which reads :
$[X,Y] = \sum_{a,b,c} C^a_{bc} X^b Y^c \overrightarrow{\kappa}_a$
where the scalars C^a_{bc} are the structure constants, with $C^a_{bc} = -C^a_{cb}$.
From the expression above, we have in complex notation :
$[Z,Z']^a = [j(Z)Z']^a = \sum_{b=1}^{3}[j(Z)]^a_c Z'^c = -\sum_{b=1}^{3}\epsilon(a,c,b)Z^b Z^c = \sum_{b=1}^{3}\epsilon(a,b,c)Z^b Z^c$

$$C^a_{bc} = \epsilon(a,b,c) \tag{3.49}$$

with

Notation 59 $\epsilon(j,k,l)$ *= the signature of the permutation of the three different integers* $i,j,k,0$ *if two integers are equal*

The scalar product reads :
$\langle \upsilon (r,w), \upsilon (r',w') \rangle = Z^t Z' = \frac{1}{4} (r+iw)^t (r'+iw')$
We can define the hermitian form :
$(\upsilon (r,w), \upsilon (r',w')) = \left\langle \overline{\upsilon (r,w)}, \upsilon (r',w') \right\rangle_{Cl} = \langle \upsilon (r,-w), \upsilon (r',w') \rangle = \frac{1}{4} (r-iw)^t (r'+iw')$
and, with the complex structure $T_1 Spin(3,1)_{\mathbb{C}}$ and $T_1 Spin(1,3)_{\mathbb{C}}$ are both 3 dimensional complex Hilbert space.

With a choice of ε_0 the vector space $T_1 Spin(3) = \operatorname{Re} T_1 Spin(3,1)_{\mathbb{C}}$ is a 3 dimensional real Hilbert space.

Spin group

For $g = a + \upsilon (r,w) + b\varepsilon_5 \in Spin(3,1)$

$$g = a + \upsilon (r,w) + b\varepsilon_5 = A + Z \tag{3.50}$$

with
$A = a + ib$
$Z = \upsilon (r,w)$
The identities
$a^2 - b^2 = 1 + \frac{1}{4} (w^t w - r^t r)$
$ab = -\frac{1}{4} r^t w$
read :
$A^2 = a^2 - b^2 + 2iab$
$Z^t Z = (r+iw)^t (r+iw) = r^t r - w^t w + 2ir^t w = 4(1 - a^2 + b^2) + 2i(-4ab) = 4(1 - a^2 + b^2 - 2iab) = 4(1 - A^2)$
$\Leftrightarrow$

$$A^2 = 1 - \frac{1}{4} Z^t Z \tag{3.51}$$

$Z \in \mathbb{R}^3 \Leftrightarrow A + Z \in Spin(3)$
$g^{-1} = a - \upsilon (r,w) + b\varepsilon_5 = A - Z$
$g \cdot g' = (A+Z)\cdot(A'+Z') = AA' + A'Z + AZ' + Z \cdot Z' = AA' + A'Z + AZ' - \frac{1}{4} Z'^t Z + \frac{1}{2} j(Z) Z'$

Expression of the derivatives on the Spin group

Theorem 60 *Let : $\sigma : F \to Spin(3,1) :: \sigma(x)$ be a differentiable map with any argument x. Then $\sigma^{-1} \cdot \frac{\partial \sigma}{\partial x}, \frac{\partial \sigma}{\partial x} \cdot \sigma^{-1} \in T_1 Spin(3,1)$ and we have with $\sigma = A + Z$:*

$$\begin{array}{c} \frac{\partial \sigma}{\partial x} \cdot \sigma^{-1} = [D(Z)] \frac{\partial Z}{\partial x} \\ \sigma^{-1} \cdot \frac{\partial \sigma}{\partial x} = [D(-Z)] \frac{\partial Z}{\partial x} \\ D(Z) = \frac{1}{A} + \frac{1}{2} j(Z) + \frac{1}{4A} j(Z) j(Z) \end{array} \tag{3.52}$$

Proof. In the complex formalism :
$\sigma = A + Z$ and $A^2 = 1 - \frac{1}{4} Z^t Z$
$\sigma^{-1} \cdot \frac{\partial \sigma}{\partial x} = (A - Z) \cdot \left(\frac{\partial A}{\partial x} + \frac{\partial Z}{\partial x} \right)$
$= A \frac{\partial A}{\partial x} + A \frac{\partial Z}{\partial x} - \frac{\partial A}{\partial x} Z - Z \cdot \frac{\partial Z}{\partial x}$
$= A \frac{\partial A}{\partial x} + A \frac{\partial Z}{\partial x} - \frac{\partial A}{\partial x} Z + \frac{1}{4} Z^t \frac{\partial Z}{\partial x} - \frac{1}{2} j(Z) \frac{\partial Z}{\partial x}$
But : $A \frac{\partial A}{\partial x} = -\frac{1}{4} Z^t \frac{\partial Z}{\partial x}$
$\sigma^{-1} \cdot \frac{\partial \sigma}{\partial x} = A \frac{\partial Z}{\partial x} + \frac{1}{4A} Z Z^t \frac{\partial Z}{\partial x} - \frac{1}{2} j(Z) \frac{\partial Z}{\partial x}$
$= \left(A + \frac{1}{4A} Z Z^t - \frac{1}{2} j(Z) \right) \frac{\partial Z}{\partial x}$
$= \left(A + \frac{1}{4A} \left(j(Z) j(Z) + Z^t Z \right) - \frac{1}{2} j(Z) \right) \frac{\partial Z}{\partial x}$

$= \left(\frac{1}{A} - \frac{1}{2}j(Z) + \frac{1}{4A}j(Z)j(Z)\right)\frac{\partial Z}{\partial x} = D(-Z)\frac{\partial Z}{\partial x}$

Similarly $\frac{\partial \sigma}{\partial x}\cdot\sigma^{-1} \in T_1 Spin(3,1)$ and one can check that :

$\frac{\partial \sigma}{\partial x}\cdot\sigma^{-1} = \left(\frac{1}{A} + \frac{1}{2}j(Z) + \frac{1}{4A}j(Z)j(Z)\right)\frac{\partial Z}{\partial x} = D(Z)\frac{\partial Z}{\partial x}$

Moreover $\det D(Z) = \frac{1}{A}, [D(Z)]^{-1} = A - \frac{1}{2}j(Z)$ ■

$[D(Z)]$ is a 3×3 matrix on $\mathbb{C}^3$.

$\frac{\partial \sigma}{\partial x}\cdot\sigma^{-1} = \sum_{a=1}^{3}[D(Z)]_b^a \frac{\partial Z^b}{\partial x}\overrightarrow{\kappa}_a$

The formulas are useful, because they relate easily the derivatives in the Spin group to the derivatives of their scalar components : $\frac{\partial Z}{\partial x} = \left(\frac{\partial r}{\partial x} + i\frac{\partial w}{\partial x}\right)$ which reads

$\left(\frac{\partial r}{\partial x} + i\frac{\partial w}{\partial x}\right) = \upsilon\left(\frac{\partial r}{\partial x}, \frac{\partial w}{\partial x}\right) \in T_1 Spin(3,1)$.

Moreover we have the identities :

$\frac{\partial A}{\partial x} = -\frac{1}{4A}Z^t\frac{\partial Z}{\partial x}$

$\frac{\partial Z}{\partial x} = \left(A - \frac{1}{2}j(Z)\right)\frac{\partial \sigma}{\partial x}\cdot\sigma^{-1}$

$D(Z)\frac{\partial Z}{\partial x} = \frac{1}{A}\frac{\partial Z}{\partial x} + \frac{1}{2}\left[Z, \frac{\partial Z}{\partial x}\right] + \frac{1}{4A}\left[Z, \left[Z, \frac{\partial Z}{\partial x}\right]\right]$

$\frac{\partial}{\partial y}\left(\frac{\partial \sigma}{\partial x}\cdot\sigma^{-1}\right) = \left(\frac{\partial}{\partial y}D(Z)\right)\frac{\partial Z}{\partial x} + D(Z)\frac{\partial}{\partial y}\frac{\partial Z}{\partial x}$

The adjoint map

Ad is a complex linear map :

Proof. $\forall s \in Spin(3,1) : \mathbf{Ad}_s\varepsilon_5 = \varepsilon_5 \Leftrightarrow s\cdot\varepsilon_5 = \varepsilon_5\cdot s$

$\mathbf{Ad}_s X\cdot\varepsilon_5 = s\cdot X\cdot\varepsilon_5\cdot s^{-1} = s\cdot X\cdot s^{-1}\cdot\varepsilon_5 = \mathbf{Ad}_s\cdot\varepsilon_5 \Leftrightarrow \mathbf{Ad}_s iX = i\mathbf{Ad}_s X$ ■

The conjugate $\overline{\mathbf{Ad}_g}$ is defined as the complex linear map :

$\forall X \in Cl(\mathbb{C},3) : \overline{\mathbf{Ad}_g}(X) = \overline{\mathbf{Ad}_g\left(\overline{X}\right)} = \overline{g\cdot\overline{X}\cdot g^{-1}} = \overline{g}\cdot X\cdot\overline{g}^{-1} = \mathbf{Ad}_{\overline{g}}X$

so, because $g \in Spin(3) \Rightarrow g = \overline{g}$ the hermitian product $(X,X')_C = \left\langle\overline{X},X'\right\rangle_{Cl}$ is preserved by $Spin(3)$:

$g \in Spin(3) : (\mathbf{Ad}_{\overline{g}}X, \mathbf{Ad}_{\overline{g}}X')_C = \left\langle\overline{\mathbf{Ad}_{\overline{g}}X}, \mathbf{Ad}_{\overline{g}}X'\right\rangle_{Cl} = \left\langle\mathbf{Ad}_{\overline{g}}\overline{X}, \mathbf{Ad}_{\overline{g}}X'\right\rangle_{Cl} = \left\langle\overline{X},X'\right\rangle_{Cl}$

Theorem 61 *Over the Lie algebra, the map :* $\mathbf{Ad}_s\upsilon(r,w) = s\cdot X\cdot s^{-1}$ *reads in matrix form :*

$\mathbf{Ad}_s\upsilon(r,w) = (A+Z)\cdot X\cdot(A-Z) = Ad(Z)[X] = \left(1 + Aj(Z) + \frac{1}{2}j(Z)j(Z)\right)[X]$

Proof. $s = A+Z, \upsilon(r,w) = X$

$\mathbf{Ad}_s\upsilon(r,w) = (A+Z)\cdot X\cdot(A-Z)$

$= A^2X + \frac{1}{4}AX^tZ - \frac{1}{2}Aj(X)Z + AZ\cdot X + \frac{1}{4}(X^tZ)Z - \frac{1}{2}Z\cdot j(X)Z$

$= \frac{1}{4}A(X^tZ) - \frac{1}{4}A(Z^tX) + \frac{1}{4}(X^tZ)Z + A^2X + Aj(Z)X + \frac{1}{4}j(Z)j(Z)X$

$= \frac{1}{4}Z(Z^tX) + A^2X + Aj(Z)X + \frac{1}{4}j(Z)j(Z)X$

$= \frac{1}{4}j(Z)j(Z)X + \frac{1}{4}(Z^tZ)X + A^2X + Aj(Z)X + \frac{1}{4}j(Z)j(Z)X$

$= \left(1-A^2\right)X + A^2X + Aj(Z)X + \frac{1}{2}j(Z)j(Z)X$

$= X + Aj(Z)X + \frac{1}{2}j(Z)j(Z)X$ ■

that we can write :

$$[\mathbf{Ad}_s]_C[X]_C = [Ad(Z)][X]_C = \left(1 + Aj(Z) + \frac{1}{2}j(Z)j(Z)\right)[X]_C \tag{3.53}$$

with : $\left[1 + Aj(Z) + \frac{1}{2}j(Z)j(Z)\right]^{-1} = [\mathbf{Ad}_{s^{-1}}]_C = \left[1 - Aj(Z) + \frac{1}{2}j(Z)j(Z)\right]$

And we have the identity, with the matrix $D(Z)$:

$$[Ad(Z)][D(Z)] = \left(1 - Aj(Z) + \frac{1}{2}j(Z)j(Z)\right)D(Z) = D(-Z) = \frac{1}{A} - \frac{1}{2}j(Z) + \frac{1}{4A}j(Z)j(Z) \tag{3.54}$$

$\Leftrightarrow [\mathbf{Ad}_{\sigma^{-1}}]\frac{\partial \sigma}{\partial x}\cdot\sigma^{-1} = \sigma^{-1}\cdot\frac{\partial \sigma}{\partial x}$

$\Rightarrow Ad(Z) = D(-Z)D(Z)^{-1} = D(-Z)\left(A - \frac{1}{2}j(Z)\right)$

3.4.5 Coordinates on the Clifford Algebra

The Clifford Algebra is a vector space, and any element can be represented as a vector with its components in the canonic basis.

The Lie Algebra is a vector subspace, and we have the choice between :

$\upsilon(X_r, X_w) = \sum_{a=1}^{3} X_r^a \overrightarrow{\kappa}_a + \sum_{a=4}^{6} X_w^{a-3} \overrightarrow{\kappa}_a$

and the complex representation : $Z = \sum_{a=1}^{3} Z^a \overrightarrow{\kappa}_a$

The Spin Group is not a vector space, but a 6 dimensional manifold embedded in the Clifford Algebra. Its elements depend on 2 vectors of $\mathbb{R}^3$: r, w but their meaning depend on the chart used.

i) The simplest chart is :

$\sigma : \mathbb{R}^3 \times \mathbb{R}^3 \to Spin(3,1) :: \sigma = a + \upsilon(r, w) + b\varepsilon_5$

with $a^2 - b^2 = 1 + \frac{1}{4}(w^t w - r^t r)$

$ab = -\frac{1}{4} r^t w$

ii) The decomposition :

$\sigma : \mathbb{R}^3 \times \mathbb{R}^3 \to Spin(3,1) :: \sigma = \sigma_w \cdot \sigma_r = (a_w + \upsilon(0, w)) \cdot (a_r + \upsilon(r, 0))$

with :

$a_w^2 = 1 + \frac{1}{4} w^t w$

$a_r^2 = 1 - \frac{1}{4} r^t r$

Then σ_w, σ_r are defined up to the sign.

iii) The complex representation :

$\sigma : \mathbb{C}^3 \to Spin(3,1) :: \sigma = A + \sum_{a=1}^{3} Z^a \overrightarrow{\kappa}_a$

with : $A^2 = 1 - \frac{1}{4} Z^t Z, Z = r + iw$

We go from one to the other by :

$\sigma_w \cdot \sigma_r = (a_w + \upsilon(0, w)) \cdot (a_r + \upsilon(r, 0)) = a_w a_r + \upsilon\left(a_w r, \frac{1}{2} j(w) r + a_r w\right) - \frac{1}{4}(w^t r)\varepsilon_5 = A + Z$

The choice of the chart can be fitted to the problem at hand. And we will also write $\sigma(r, w)$ when no choice has been done.

Notice that w corresponds to a pure translational motion only in the second chart : $\sigma_w \cdot \sigma_r = (a_w + \upsilon(0, w)) \cdot (a_r + \upsilon(r, 0))$.

3.5 REPRESENTATION OF THE MOTION IN GENERAL RELATIVITY

We have now the mathematical tools to build a representation of the motion of material bodies in the Geometry of GR. It is based upon the existence :
- of a tetrad attached to each particle, or material point,
- of a tetrad for any observer at each point,
- of a relation between the velocity of the particle and its tetrad.

3.5.1 Description of the fiber bundles

Associated vector bundles

From the principal bundle is $P_G(M, Spin(3,1), \pi_G)$ other fiber bundles can be defined.

Definition 62 *The vector bundle TM defined through the tetrad of an observer is $P_G[\mathbb{R}^4, \mathbf{Ad}]$: $\varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i)$*

In a change of observer :

$$\begin{gathered}\mathbf{p}(m) = \varphi_G(m,1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m) \cdot \chi(m)^{-1} : \\ (\mathbf{p}(m), u) \sim \left(\widetilde{\mathbf{p}}(m), \mathbf{Ad}_{\chi(m)} u\right) \\ \varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i) \to \widetilde{\varepsilon}_i(m) = \mathbf{Ad}_{\chi(m)^{-1}} \varepsilon_i(m) = \sum_{j=0}^3 \left[h\left(\chi(m)^{-1}\right)\right]_i^j \varepsilon_j(m)\end{gathered} \tag{3.55}$$

The formulas are the same as previously, the relation between $\varepsilon_i(m), \widetilde{\varepsilon}_i(m)$ is just explicit with $\mathbf{Ad}$. In $P_G[\mathbb{R}^4, \mathbf{Ad}]$ the components of vectors are measured in orthonormal bases.

$\varepsilon_0(m) = (\mathbf{p}(m), \varepsilon_0)$ is the 4th vector both in the Clifford algebra and in the tangent space T_mM. It corresponds to the velocity of the observer : $\varepsilon_0(q_o(t)) = \frac{1}{c}\frac{dq_o}{dt}$ 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 $P_G[\mathbb{R}^4, \mathbf{Ad}]$.

The gauge of an observer is defined by his tetrad : it is the physical link between the abstract fiber bundle P_G and the measures involving P_G.

Definition 63 *The **adjoint bundle** is the associated vector bundle $P_G[T_1Spin(3,1), \mathbf{Ad}]$*

Because M is endowed with the structure of the principal bundle P_G, there is a structure of **Clifford bundle** $Cl(TM)$: 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 it is isomorphic to $Cl(3,1)$ (Maths.2106). Pointwise the Clifford product holds with the usual properties, and with the vectors defined in the tetrad.

Definition 64 *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 $Cl(TM)$ transforms as :

$$\begin{gathered}\mathbf{p}(m) = \varphi_G(m,1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m) \cdot \chi(m)^{-1} \\ (\mathbf{p}(m), X) \sim \left(\widetilde{\mathbf{p}}(m), \mathbf{Ad}_{\chi(m)} X\right)\end{gathered} \tag{3.56}$$

A basis of $Cl(T_mM)$ is given by 1 and ordered products of $\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3$. It changes as $\widetilde{\varepsilon}_i(m) = \mathbf{Ad}_{\chi(m)^{-1}} \varepsilon_i(m)$ and the components change as $\left[\mathbf{Ad}_{\chi(m)}\right] X$, the matrix $\left[\mathbf{Ad}_{\chi(m)}\right]$ depending on X.

Fundamental symmetry breakdown

The observer uses the frame $\left(O,(\varepsilon_i)_{i=0}^3\right)$ to measure the components of vectors of TM. 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

$P_G(P_W, Spin(3), \pi_R)$ 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 $s \in Spin(3,1)$.

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 $\varepsilon_0(m)$.

3.5.2 Motion of a particle

The motion is defined as the change of position with respect to a present position, as in the instantaneous rotation $R^{-1}\frac{dR}{dt}$, so we need first to define the position, and this is done through the tetrad attached to the particle.

Arrangement of the particle

The fundamental assumption is the existence of an orthonormal basis $(e_i)_{i=0}^3$ attached to the particle. At each point it is measured in the vector bundle $P_G\left[\mathbb{R}^4, \mathbf{Ad}\right]$. The basis $(e_i)_{i=0}^3$ is deduced from the tetrad $(\varepsilon_i)_{i=0}^3$ of the observer by an element $\sigma \in Spin(3,1)$ such that :

$e_i = \mathbf{Ad}_\sigma \varepsilon_i \Leftrightarrow e_i(q(t)) = \left(\mathbf{p}(q(t)), \mathbf{Ad}_{\sigma(t)}\varepsilon_i\right)$

and we define the *arrangement* of the particle with respect to the observer O by σ.

The velocity $\frac{dq}{dt}$ of the particle reads for an observer :

- in the basis of the standard chart :$V = \frac{dq}{dt} = \sum_{\alpha=0}^3 V^\alpha \partial\xi_\alpha = c\varepsilon_0 + \overrightarrow{v}$

- in the tetrad at each point : $U = \sum_{j=0}^3 U^j \varepsilon_j = c\left(\varepsilon_0 + \overrightarrow{u}\right)$

(Notice that we have $\overrightarrow{v}$ but $c\overrightarrow{u}$ for convenience in further computations)

and $V = U = \frac{dq}{dt}$ as vectors.

Because the velocity V of the particle is proportional to e_0 we have :

$V = \sqrt{-\langle V, V\rangle}e_0 \Leftrightarrow U = \sqrt{-\langle U, U\rangle}\mathbf{Ad}_\sigma\varepsilon_0$

and

$\langle U, \varepsilon_0\rangle_{Cl} = \left\langle \sqrt{-\langle U,U\rangle}\mathbf{Ad}_\sigma\varepsilon_0, \varepsilon_0\right\rangle_{Cl} = \left\langle c\left(\varepsilon_0 + \overrightarrow{u}\right), \varepsilon_0\right\rangle_{Cl} = \langle c\varepsilon_0, \varepsilon_0\rangle_{Cl} = -c$

$\sqrt{-\langle U, U\rangle} = -c\frac{1}{\langle \mathbf{Ad}_\sigma\varepsilon_0, \varepsilon_0\rangle_{Cl}}$

$U = -c\frac{1}{\langle \mathbf{Ad}_\sigma\varepsilon_0, \varepsilon_0\rangle_{Cl}}\mathbf{Ad}_\sigma\varepsilon_0$

$\langle \mathbf{Ad}_\sigma\varepsilon_0, \varepsilon_0\rangle_{Cl}$ is the scalar product in the tetrad, so $\langle \mathbf{Ad}_\sigma\varepsilon_0, \varepsilon_0\rangle_{TM} = \langle \mathbf{Ad}_\sigma\varepsilon_0, \varepsilon_0\rangle_{Cl}$ and does not depend on the metric. Notice that $\mathbf{Ad}_\sigma\varepsilon_0, \varepsilon_0$ are both vectors in the fixed vector space $\mathbb{R}^4$

$\Rightarrow \sqrt{-\langle V, V\rangle} = -\frac{c}{\langle \mathbf{Ad}_\sigma\varepsilon_0, \varepsilon_0\rangle_{Cl}}$

In a change of gauge : $\mathbf{p}(m) = \varphi_G(m,1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m)\cdot\chi(m)^{-1}$ the vector e_i does not change, but its components change, and thus σ changes as a section of P_G :

$e_i(m) = (\mathbf{p}(m), \mathbf{Ad}_\sigma\varepsilon_i) \sim \left(\mathbf{p}(m)\cdot\chi(m)^{-1}, \mathbf{Ad}_{\chi(m)}\mathbf{Ad}_\sigma\varepsilon_i\right) = \left(\widetilde{\mathbf{p}}(m), \mathbf{Ad}_{\chi(m)\cdot\sigma}\varepsilon_i\right)$

Notice that the vector $\varepsilon_i \in \mathbb{R}^4$ and never changes.

These formulas are just the expression, in the Clifford algebra, of the classic relations in a change of basis between two observers (51).

$$\begin{gathered} e_i = \mathbf{Ad}_\sigma\varepsilon_i \\ V = \frac{dq}{dt} = c\varepsilon_0 + \overrightarrow{v} = U = -\frac{c}{\langle\mathbf{Ad}_\sigma\varepsilon_0,\varepsilon_0\rangle_{Cl}}\mathbf{Ad}_\sigma\varepsilon_0 \end{gathered} \tag{3.57}$$

With the chart : $\sigma = \sigma_w\cdot\sigma_r = \epsilon(a_w + \upsilon(0,w))\cdot\epsilon(a_r + \upsilon(r,0))$ with $\epsilon = \pm 1$

$\frac{U}{\sqrt{-\langle V,V\rangle}} = e_0 = \mathbf{Ad}_{\sigma_w\cdot\sigma_r}\varepsilon_0 = \mathbf{Ad}_{\sigma_w}\mathbf{Ad}_{\sigma_r}\varepsilon_0 = \mathbf{Ad}_{\sigma_w}\varepsilon_0$ because $\sigma_r \in T_1 Spin(3)$ so :

$\frac{U}{\sqrt{-\langle V,V\rangle}} = \mathbf{Ad}_{\sigma_w}\varepsilon_0$

The matrix of $\mathbf{Ad}_{\sigma_w}$ is :

$[h(\sigma_w)] = \begin{bmatrix} 2a_w^2 - 1 & a_w w^t \\ a_w w & 2a_w^2 - 1 + \frac{1}{2}j(w)j(w) \end{bmatrix}$

$\mathbf{Ad}_\sigma\varepsilon_0 = \left(2a_w^2 - 1\right)\varepsilon_0 + a_w\sum_{i=1}^3 w_i\varepsilon_i$

$\langle\mathbf{Ad}_\sigma\varepsilon_0, \varepsilon_0\rangle = -\left(2a_w^2 - 1\right)$

$U = -\frac{c}{\langle\mathbf{Ad}_\sigma\varepsilon_0,\varepsilon_0\rangle_{Cl}}\mathbf{Ad}_\sigma\varepsilon_0 = c\left(\varepsilon_0 + \frac{a_w}{2a_w^2-1}\sum_{i=1}^3 w_i\varepsilon_i\right)$

$\sqrt{-\langle U,U\rangle} = \frac{c}{2a_w^2-1}$

$u^t u = \left(\frac{2a_w}{2a_w^2-1}\right)^2\left(a_w^2 - 1\right)$

$V = c\sum_{\alpha=0}^3\left(\mathbf{P}_0^\alpha(q(t)) + \frac{a_w}{2a_w^2-1}\sum_{i=1}^3 w_i(t)\mathbf{P}_i^\alpha(q(t))\right)\partial\xi_\alpha$

$\overrightarrow{v} = 0 \Leftrightarrow w = 0$

U is determined by σ_w only. Meanwhile σ is uniquely defined by $(e_i)_{i=0}^3$, σ_w is defined up to the sign. In all cases we have $a_w\sum_{i=1}^3 w_i\varepsilon_i = a_w\overrightarrow{w}$ in the same direction as the spatial velocity, but this can be achieved either by $\overrightarrow{w}$ in the same direction as the spatial velocity and $a_w > 0$ or by $-\overrightarrow{w}$ and $-a_w$. σ_r is similarly defined up to the sign.

$$U = c\left(\varepsilon_0 + \sum_{a=1}^3\frac{a_w}{2a_w^2-1}w_a\varepsilon_a\right) \tag{3.58}$$

With the complex chart :

$\sigma = A + \sum_{a=1}^3 Z^a\overrightarrow{\kappa}_a = a + \upsilon(r,w) + b\varepsilon_5$

$A = a + ib$

$A^2 = 1 - \frac{1}{4}Z^t Z$

The matrix $[h(\sigma)]$ has been given previously and :

$\mathbf{Ad}_\sigma\varepsilon_0 = [h(s)] = \left(a^2 + b^2 + \frac{1}{4}\left(r^t r + w^t w\right)\right)\varepsilon_0 + \left(aw - br + \frac{1}{2}j(r)w\right)$

$\langle\mathbf{Ad}_\sigma\varepsilon_0, \varepsilon_0\rangle_{Cl} = -\left(a^2 + b^2 + \frac{1}{4}\left(r^t r + w^t w\right)\right)$

$U = -\frac{c}{\langle\mathbf{Ad}_\sigma\varepsilon_0,\varepsilon_0\rangle_{Cl}}\mathbf{Ad}_\sigma\varepsilon_0 = c\varepsilon_0 + \frac{c}{\left(a^2+b^2+\frac{1}{4}(r^t r+w^t w)\right)}\left(aw - br + \frac{1}{2}j(r)w\right)$

$u^t u = 1 - \frac{1}{\left(a^2+b^2+\frac{1}{4}(r^t r+w^t w)\right)^2} = 1 - \frac{1}{A\overline{A}+\frac{1}{4}Z^t\overline{Z}}$

$A\overline{A} + \frac{1}{4}Z^t\overline{Z} = a^2 + b^2 + \frac{1}{4}\left(r^t r + w^t w\right)$

$aw - br + \frac{1}{2}j(r)w = \operatorname{Re}A\operatorname{Im}Z - \operatorname{Im}A\operatorname{Re}Z + \frac{1}{2}j(\operatorname{Re}Z)\operatorname{Im}Z$

$j(Z)\overline{Z} = -2ij(\operatorname{Re}Z)(\operatorname{Im}Z)$

$A\overline{Z} = (\operatorname{Re}A)\operatorname{Re}Z - i(\operatorname{Re}A)\operatorname{Im}Z + i\operatorname{Im}A\operatorname{Re}Z + \operatorname{Im}A\operatorname{Im}Z$

$aw - br + \frac{1}{2}j(r)w = -\operatorname{Im}\left(A\overline{Z} + \frac{1}{4}j(Z)\overline{Z}\right)$

$U = c\varepsilon_0 - \frac{c}{A\overline{A}+\frac{1}{4}Z^t\overline{Z}} \operatorname{Im}\left(A\overline{Z} + \frac{1}{4}j(Z)\overline{Z}\right)$

$$U = c\left(\varepsilon_0 - \frac{1}{A\overline{A}+\frac{1}{4}Z^t\overline{Z}} \operatorname{Im}\left\{\left(A + \frac{1}{4}j(Z)\right)\overline{Z}\right\}\right) \tag{3.59}$$

Motion

The tetrad attached to the particle is defined in the tetrad of the observer, and the motion is defined by derivation with respect to a fixed observer, that is *with respect to a fixed tetrad*. To underline this fact it is useful to use the vector U, defined in the tetrad. Which is equivalent to consider the vector $U(m) = (\mathbf{p}(m), U)$ in the associated vector bundle with $U \in \mathbb{R}^4$.

A continuous motion is such that the map : $\sigma : \mathbb{R} \to Spin(3,1)$ with respect to the time t of the observer is smooth. From the definitions above :

$\forall i = 0..3 : e_i = \mathbf{Ad}_\sigma \varepsilon_i$

$\frac{de_i}{dt} = \frac{d}{dt}\mathbf{Ad}_\sigma \varepsilon_i = \mathbf{Ad}_\sigma\left[\sigma^{-1}\cdot\frac{d\sigma}{dt}, \varepsilon_i\right] = \left[\frac{d\sigma}{dt}\cdot\sigma^{-1}, \mathbf{Ad}_\sigma\varepsilon_i\right] = \left[\frac{d\sigma}{dt}\cdot\sigma^{-1}, e_i\right]$

$\forall i = 0..3 : \frac{de_i}{dt} = \left[\frac{d\sigma}{dt}\cdot\sigma^{-1}, e_i\right]$

$U = \sqrt{-\langle U, U\rangle}\mathbf{Ad}_\sigma\varepsilon_0 = \sqrt{-\langle U, U\rangle}e_0$

$\frac{dU}{dt} = \frac{d}{dt}\sqrt{-\langle U, U\rangle}e_0 + \sqrt{-\langle U, U\rangle}\frac{de_0}{dt}$

$= \left(\frac{1}{\sqrt{-\langle U,U\rangle}}\frac{d}{dt}\sqrt{-\langle U, U\rangle}\right)\sqrt{-\langle U, U\rangle}e_0 + \sqrt{-\langle U, U\rangle}\left[\frac{d\sigma}{dt}\cdot\sigma^{-1}, e_0\right]$

$\frac{dU}{dt} = \left(\frac{1}{\sqrt{-\langle U,U\rangle}}\frac{d}{dt}\sqrt{-\langle U, U\rangle}\right)U + \left[\frac{d\sigma}{dt}\cdot\sigma^{-1}, U\right]$

$\frac{d}{dt}\sqrt{-\langle U, U\rangle} = \frac{c}{\left(\langle\mathbf{Ad}_\sigma\varepsilon_0,\varepsilon_0\rangle_{Cl}\right)^2}\frac{d}{dt}\langle\mathbf{Ad}_\sigma\varepsilon_0, \varepsilon_0\rangle_{Cl}$

$\frac{1}{\sqrt{-\langle U,U\rangle}}\frac{d}{dt}\sqrt{-\langle U, U\rangle}$

$= -\frac{1}{\langle\mathbf{Ad}_\sigma\varepsilon_0,\varepsilon_0\rangle_{Cl}}\left\langle\frac{d}{dt}\mathbf{Ad}_\sigma\varepsilon_0, \varepsilon_0\right\rangle_{Cl}$

$= -\frac{1}{\langle\mathbf{Ad}_\sigma\varepsilon_0,\varepsilon_0\rangle_{Cl}}\left\langle\left[\frac{d\sigma}{dt}\cdot\sigma^{-1}, \mathbf{Ad}_\sigma\varepsilon_0\right], \varepsilon_0\right\rangle_{Cl}$

$= \frac{1}{c}\left\langle\left[\frac{d\sigma}{dt}\cdot\sigma^{-1}, U\right], \varepsilon_0\right\rangle_{Cl}$

$\frac{dU}{dt} = \frac{U}{c}\left\langle\left[\frac{d\sigma}{dt}\cdot\sigma^{-1}, U\right], \varepsilon_0\right\rangle_{Cl} + \left[\frac{d\sigma}{dt}\cdot\sigma^{-1}, U\right]$

And we define the motion (both translational and rotational) of the particle by $\frac{d\sigma}{dt}\cdot\sigma^{-1} \in T_1Spin(3,1)$.

$\delta_R\sigma = \frac{d\sigma}{dt}\cdot\sigma^{-1}$ is the right logarithmic derivative, and $\delta_L\sigma = \sigma^{-1}\cdot\frac{d\sigma}{dt}$ is the left logarithmic derivative. They both belong to $T_1Spin(3,1)$ and are related by $\mathbf{Ad}_\sigma : \delta_R\sigma = \mathbf{Ad}_\sigma\delta_L\sigma \Leftrightarrow \delta_L\sigma = \mathbf{Ad}_{\sigma^{-1}}\delta_R\sigma$.

$$\begin{gathered}\frac{d\sigma}{dt}\cdot\sigma^{-1} = \upsilon(X_r, X_w) \in T_1Spin(3,1)\\ \forall i = 0..3 : \frac{de_i}{dt} = [\upsilon(X_r, X_w), e_i]\\ \frac{dU}{dt} = \frac{U}{c}\langle[\upsilon(X_r, X_w), U], \varepsilon_0\rangle_{Cl} + [\upsilon(X_r, X_w), U]\end{gathered} \tag{3.60}$$

With $\sigma = \sigma_w\cdot\sigma_r = \epsilon(a_w + \upsilon(0, w))\cdot\epsilon(a_r + \upsilon(r, 0))$

$\frac{d\sigma}{dt}\cdot\sigma^{-1} = \upsilon(X_r, X_w)$ with

$X_r = -\frac{1}{2}j(w)\frac{dw}{dt} + \left[1 - \frac{1}{2}j(w)j(w)\right]\left(\frac{1}{a_r} + \frac{1}{2}j(r) + \frac{1}{4a_r}j(r)j(r)\right)\frac{dr}{dt}$

$X_w = \frac{1}{a_w}\left(1 - \frac{1}{4}j(w)j(w)\right)\frac{dw}{dt} + [a_w j(w)]\left(\frac{1}{a_r} + \frac{1}{2}j(r) + \frac{1}{4a_r}j(r)j(r)\right)\frac{dr}{dt}$

and the inverse relation reads, with some computation :

$\frac{dr}{dt} = \left(\frac{1}{a_r} - \frac{1}{2}j(r) + \frac{1}{4a_r}j(r)j(r)\right)\left(X_r + \frac{1}{2}\frac{1}{a_w}j(w)X_w\right)$

$\frac{dw}{dt} = -j(w)X_r + \left(a_w - \frac{1}{4a_w}j(w)j(w)\right)X_w$

$\frac{dV}{dt} = cX_w + \left(j\left(X_r\right) - \left(X_w^t v\right)\frac{1}{c}\right)v$ with : $V = c\varepsilon_0 + v$

$$\begin{array}{c} \frac{d\sigma}{dt}\cdot\sigma^{-1} = \upsilon\left(X_r, X_w\right) \\ X_r = -\frac{1}{2}j\left(w\right)\frac{dw}{dt} + \left[1 - \frac{1}{2}j\left(w\right)j\left(w\right)\right]\left(\frac{1}{a_r} + \frac{1}{2}j\left(r\right) + \frac{1}{4a_r}j\left(r\right)j\left(r\right)\right)\frac{dr}{dt} \\ X_w = \frac{1}{a_w}\left(1 - \frac{1}{4}j\left(w\right)j\left(w\right)\right)\frac{dw}{dt} + \left[a_w j\left(w\right)\right]\left(\frac{1}{a_r} + \frac{1}{2}j\left(r\right) + \frac{1}{4a_r}j\left(r\right)j\left(r\right)\right)\frac{dr}{dt} \\ \frac{dU}{dt} = cX_w + \left(j\left(X_r\right) - \left(X_w^t v\right)\frac{1}{c}\right)v \end{array} \tag{3.61}$$

With $\sigma = A + \sum_{a=1}^{3} Z^a \overrightarrow{\kappa}_a = A + Z$
$\frac{d\sigma}{dt}\cdot\sigma^{-1} = D\left(Z\right)\frac{dZ}{dt} = Y_r + iY_w$
$\left[D\left(Z\right)\right]^{-1} = A - \frac{1}{2}j\left(Z\right)$
$\frac{dZ}{dt} = \left[D\left(Z\right)\right]^{-1}\left(Y_r + iY_w\right) = \left(A - \frac{1}{2}j\left(Z\right)\right)\left(Y_r + iY_w\right)$

$$\begin{array}{c} \frac{d\sigma}{dt}\cdot\sigma^{-1} = D\left(Z\right)\frac{dZ}{dt} = Y_r + iY_w \\ \frac{dZ}{dt} = \left(A - \frac{1}{2}j\left(Z\right)\right)\left(Y_r + iY_w\right) \\ \frac{dU}{dt} = cY_w + \left(j\left(Y_r\right) - \left(\left[Y_w\right]^t\left[v\right]\right)\frac{1}{c}\right)v \end{array} \tag{3.62}$$

To sum up

The geometric characteristics of a particle can be represented by a map :

$\sigma : \mathbb{R} \to P_G :: \sigma\left(t\right) = \varphi_G\left(q\left(t\right), \sigma\left(t\right)\right)$

It defines :

- its trajectory : $q\left(t\right) = \pi_G\left(\sigma\left(t\right)\right)$
- its arrangement in the tetrad of the observer : $e_i\left(t\right) = \mathbf{Ad}_{\sigma(t)}\varepsilon_0\left(q\left(t\right)\right) = \left(\mathbf{p}\left(q\left(t\right)\right), \mathbf{Ad}_{\sigma(t)}\varepsilon_0\right)$
- its velocity, with respect to the observer in his standard chart :

$U\left(t\right) = -\frac{c}{\left\langle \mathbf{Ad}_{\sigma(t)}\varepsilon_0, \varepsilon_0\right\rangle_{Cl}}\mathbf{Ad}_{\sigma(t)}\varepsilon_0\left(q\left(t\right)\right) = \left(\mathbf{p}\left(q\left(t\right)\right), \frac{c}{\left\langle \mathbf{Ad}_{\sigma(t)}\varepsilon_0, \varepsilon_0\right\rangle_{Cl}}\mathbf{Ad}_{\sigma(t)}\varepsilon_0\right)$

$V^\alpha\left(t\right) = \sum_{i=0}^{3} P_i^\alpha\left(q\left(t\right)\right)U^i\left(t\right)$

$V\left(t\right) = \frac{dq}{dt} = c\varepsilon_0 + \overrightarrow{v}$

and by construct $V\left(t\right)$ is a time vector, future oriented.

Both the arrangement and the trajectory are fully defined by σ : the derivatives are not involved. The tetrad is involved when one needs the components of the velocity or of the tetrad of the particle in the holonomic basis. The vectors V, U are the same geometric quantity, defined in 2 different bases : $V^\alpha = \sum_{i=0}^{3}\left[P'\right]_i^\alpha U^i$.

- its motion is then, with respect to the observer at $q\left(t\right)$:

$\forall i = 0..3 : \frac{de_i}{dt} = \left[\frac{d\sigma}{dt}\cdot\sigma^{-1}, e_i\right]$

$\frac{dU}{dt} = \frac{U}{c}\left\langle\left[\frac{d\sigma}{dt}\cdot\sigma^{-1}, U\right], \varepsilon_0\right\rangle_{Cl} + \left[\frac{d\sigma}{dt}\cdot\sigma^{-1}, U\right]$

$V^\alpha = \sum_{i=0}^{3}\left[P'\right]_i^\alpha U^i$ and because the tetrad is assumed to be fixed in the motion :

$$\frac{dV^\alpha}{dt} = \sum_{i=0}^{3}\left[P'\right]_i^\alpha \frac{dU^i}{dt} \tag{3.63}$$

so the relations above give the derivatives of the components V^α and not the derivative of the vectors $V \in TM$, which would be defined in the bitangent bundle TM^2.

- the motion, continuous or not, is represented by $\upsilon\left(X_r, X_w\right)$ in the Lie algebra $T_1 Spin\left(3,1\right)$, that is by 6 components in the basis $\overrightarrow{\kappa}_a$: 3 to represent the translational motion (the imaginary part) and 3 to represent the rotational motion (the real part), or equivalently by 2 vectors of $\mathbb{R}^3$ as in Galilean Geometry.

Spatial Speed

i) $U = c\left(\varepsilon_0 + \overrightarrow{u}\right) = -\frac{c}{\langle \mathbf{Ad}_\sigma \varepsilon_0, \varepsilon_0\rangle_{Cl}} \mathbf{Ad}_\sigma \varepsilon_0$

$\overrightarrow{u} = -c\varepsilon_0 - \frac{c}{\langle \mathbf{Ad}_\sigma \varepsilon_0, \varepsilon_0\rangle_{Cl}} \mathbf{Ad}_\sigma \varepsilon_0$

$$\langle \overrightarrow{u}, \overrightarrow{u}\rangle = 1 - \left(\frac{1}{\langle \mathbf{Ad}_\sigma \varepsilon_0, \varepsilon_0\rangle_{Cl}}\right)^2 \tag{3.64}$$

ii) In $\frac{dU}{dt} = \frac{U}{c}\left\langle \left[\frac{d\sigma}{dt}\cdot\sigma^{-1}, U\right], \varepsilon_0\right\rangle_{Cl} + \left[\frac{d\sigma}{dt}\cdot\sigma^{-1}, U\right]$ the value of $\left[\frac{d\sigma}{dt}\cdot\sigma^{-1}, U\right]$ is a vector, which is expressed (see Annex) :

$\left[\frac{d\sigma}{dt}\cdot\sigma^{-1}, U\right] = c\left(X_w + \{X_w^t u\}\varepsilon_0 + j(X_r)u\right)$ where $\frac{d\sigma}{dt}\cdot\sigma^{-1} = \upsilon(X_r, X_w), U = c\varepsilon_0 + cu$

$\frac{dU}{dt} = c\left(X_w\left(1-(u^t u)\right) + j(X_r)u - j(u)j(u)X_w\right) = c\frac{du}{dt}$

$$u^t\frac{du}{dt} = \left(1-(u^t u)\right)\left(u^t X_w\right) \tag{3.65}$$

The motion is with a constant spatial speed $u^t\frac{du}{dt} = 0$ iff $u^t X_w = 0$, so notably if

$X_w = \operatorname{Im}\frac{d\sigma}{dt}\cdot\sigma^{-1} = 0 \Leftrightarrow \frac{d\sigma}{dt}\cdot\sigma^{-1} \in T_1 Spin(3)$

Spin

The spin is a rotational motion. The spatial basis of the particle is deduced from the spatial tetrad by a rotation of $SO(3)$:

$[h(\sigma_r)] = \begin{bmatrix} 1 & 0 \\ 0 & 1 + a_r j(r(t)) + \frac{1}{2}j(r(t))j(r(t)) \end{bmatrix}$

and the rotational motion can be defined as : $\frac{d\sigma_r}{dt}\cdot\sigma_r^{-1} \in T_1 Spin(3)$.

In Galilean Geometry the rotation of the spatial basis is usually measured by a matrix $R(t) \in SO(3)$ based also on a vector $\rho \in \mathbb{R}^3$:

$R(t) = \exp j(\rho(t)) = I_3 + \frac{\sin\sqrt{\rho^t\rho}}{\sqrt{\rho^t\rho}}[j(\rho)] + \frac{1-\cos\sqrt{\rho^t\rho}}{\rho^t\rho}[j(\rho)][j(\rho)]$

so : $[h(\sigma_r)] = \begin{bmatrix} 1 & 0 \\ 0 & I_3 + a_r j(r) + \frac{1}{2}j(r)j(r) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & R(t) \end{bmatrix}$

$I_3 + \frac{\sin\sqrt{\rho^t\rho}}{\sqrt{\rho^t\rho}}[j(\rho)] + \frac{1-\cos\sqrt{\rho^t\rho}}{\rho^t\rho}[j(\rho)][j(\rho)] = 1 + a_r j(r) + \frac{1}{2}j(r)j(r)$

$I_3 + a_r j(r) + \frac{1}{2}j(r)j(r)$ has for eigen vector r with eigen value 1

$\exp j(\rho(t))$ has for eigen vector ρ with eigen value 1

thus $r = \lambda\rho$

The sign of a_r is fixed by ϵ, that is the orientation of w.

$\lambda = \epsilon\sqrt{2\frac{1-\cos\sqrt{\rho^t\rho}}{\rho^t\rho}}$

And : $\sigma_r = \epsilon\left(\frac{\sin\sqrt{\rho^t\rho}}{\sqrt{2\left(1-\cos\sqrt{\rho^t\rho}\right)}} + \upsilon\left(\sqrt{2\frac{1-\cos\sqrt{\rho^t\rho}}{\rho^t\rho}}, 0\right)\right)$

Meanwhile for the representation of the decomposition of σ we have the choice of ϵ and $(r, w) \sim (-r, -w)$, the rotational motion $\frac{d\sigma_r}{dt}\cdot\sigma_r^{-1}$ does not depend on ϵ, but introduces a new factor with the derivative. In Galilean Geometry the convention is that $-\rho$ represents the opposite spin, with the same axis. In the Relativist framework, one can distinguish the two rotations, because there is always a privileged direction (that of the velocity). One can distinguish the two spin elements $\pm\sigma_r$ (which correspond to the same matrix of $SO(3,1)$) and differentiate the rotational motion from its opposite by fixing ϵ. If we impose that $\overrightarrow{w}$ is in the direction of $\overrightarrow{v}$, then $+\rho$ and $-\rho$ represent spinning with the same axis, but opposite rotations, or equivalently, to keep the usual convention, rotations with opposite axis. These opposite rotational motions are usually called polarization (spin “up” or “down”).

In Galilean Geometry two opposite rotational motions are the image of each other in a space inversion (a symmetry with respect to a plan). In the Relativist Framework such an operation is a symmetry with respect to a *spatial* vector (and not the space inversion which is a symmetry with respect to $\Omega_3(t)$). And actually this is done through the choice of an orientation for $\overrightarrow{w}$.

The vector $r(t) \in \mathbb{R}^3$, however the characteristic of the spin is $\frac{d\sigma_r}{dt} \cdot \sigma_r^{-1} = \upsilon(X_r, 0) \in T_1 Spin(3)$ and we have seen that $\upsilon(X_r, 0)$ 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 ! Notice that this paradox exists already in Galilean Geometry : when the instantaneous rotation is represented by $R^{-1}\frac{dR}{dt}$ this quantity does not depend on the frame of the observer.

Units

We represent the motion by $\frac{d\sigma}{dt} \cdot \sigma^{-1} = \upsilon(X_r, X_w)$ with :

$$X_r = -\frac{1}{2}j(w)\frac{dw}{dt} + \left[1 - \frac{1}{2}j(w)j(w)\right]\left(\frac{1}{a_r} + \frac{1}{2}j(r) + \frac{1}{4a_r}j(r)j(r)\right)\frac{dr}{dt}$$

$$X_w = \frac{1}{a_w}\left(1 - \frac{1}{4}j(w)j(w)\right)\frac{dw}{dt} + [a_w j(w)]\left(\frac{1}{a_r} + \frac{1}{2}j(r) + \frac{1}{4a_r}j(r)j(r)\right)\frac{dr}{dt}$$

In the formula above : $V = c\left(\varepsilon_0 + \sum_{a=1}^3 \frac{a_w}{2a_w^2 - 1} w_a \varepsilon_a\right)$ 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. The rotational motion is represented also by a vector $r(t) \in \mathbb{R}^3$ which must also be unitless, and $\frac{dr}{dt}$ has the dimension $[T]^{-1}$. A rotational motion is measured in rad/s, however in Galilean Geometry it is conventionally represented by a rotation with constant axis ρ and rotational speed $\sqrt{\rho^t\rho}$: $R(t) = \exp tj(\rho) \Leftrightarrow R(t)^{-1}\frac{dR}{dt} \in so(3)$. The advantage is that ρ is observer independent. The angle θ of the rotation resulting from a given matrix is $\theta = \sqrt{\rho^t\rho}$. So a rotational motion of 2π rad/s (1 turn / s or 1 cycle / s that is 1 Hz) is represented by a vector ρ of one unit of length / s, or equivalently 2π $rad/s \sim 1m/s$.

The same rotation represented in Galilean Geometry by $R(t) = \exp j(\rho(t))$ is represented here by

$$\sigma_r = \epsilon\left(\frac{\sin\sqrt{\rho^t\rho}}{\sqrt{2\left(1-\cos\sqrt{\rho^t\rho}\right)}} + \upsilon\left(\sqrt{2\frac{1-\cos\sqrt{\rho^t\rho}}{\rho^t\rho}}, 0\right)\right)$$

and r can be measured in cycles, without unit, and $\frac{dr}{dt}$ in Hz.

Then r, w are unitless, and X_r, X_w have the dimension of $[T]^{-1}$. If we want to keep the usual and natural convention of measuring the motion by the units $[L][T]^{-1}$ we should introduce another constant in front of $\frac{d\sigma}{dt} \cdot \sigma^{-1} = \upsilon(X_r, X_w)$. We will come back on these issues in the Chapter 6.

Estimates

It is useful to have estimates for w, using the spatial speed.

Let us denote : $x = 1 - \frac{\|\overrightarrow{v}\|^2}{c^2}$

With the representation $\sigma = \sigma_w \cdot \sigma_r = (a_w + \upsilon(0, w)) \cdot (a_r + \upsilon(r, 0))$

$$a_w = \epsilon\sqrt{\frac{1}{2}\left(\frac{1}{\sqrt{x}} + 1\right)} = \epsilon\frac{1}{\sqrt{2}}x^{-1/4}\left(1 + \sqrt{x}\right)^{1/2}$$

$$w = \epsilon\sqrt{2}\left(1 - \frac{\|\overrightarrow{v}\|^2}{c^2} + \sqrt{1 - \frac{\|\overrightarrow{v}\|^2}{c^2}}\right)^{-1/2}\frac{\overrightarrow{v}}{c} = \epsilon\sqrt{2}x^{-1/4}\left(1 + \sqrt{x}\right)^{-1/2}\frac{\overrightarrow{v}}{c}$$

Usually $\frac{\|\overrightarrow{v}\|^2}{c^2} \ll 1$ and we have the estimates :

$$a_w \simeq \epsilon\left(1 + \frac{1}{8}\frac{\|\overrightarrow{v}\|^2}{c^2}\right)$$

$w \simeq \epsilon\left(1+\frac{3}{8}\frac{\left\|\overrightarrow{v}\right\|^2}{c^2}\right)\frac{\overrightarrow{v}}{c}$
$V \simeq c\left(\varepsilon_0+\epsilon\left(1-\frac{3}{8}\frac{\left\|\overrightarrow{v}\right\|^2}{c^2}\right)\overrightarrow{w}\right)$
$A \simeq \epsilon a_r\left(1+\frac{1}{8}\frac{\left\|\overrightarrow{v}\right\|^2}{c^2}\right)-i\frac{1}{4}\epsilon r^t\frac{\overrightarrow{v}}{c}$
$Z \simeq \epsilon\left(1+\frac{1}{8}\frac{\left\|\overrightarrow{v}\right\|^2}{c^2}\right)r+i\left(a_r-\frac{1}{2}j\left(r\right)\right)\epsilon\frac{\overrightarrow{v}}{c}$
The derivative of w is given by the formula :
$\frac{dw}{dt}=\left[\left(\frac{2a_w^2+1}{4a_w^3}\right)j\left(\frac{\overrightarrow{v}}{c}\right)j\left(\frac{\overrightarrow{v}}{c}\right)+\left(\frac{2a_w^2+1}{4a_w^3}\right)\frac{\left\|\overrightarrow{v}\right\|^2}{c^2}+\frac{2a_w^2-1}{a_w}\right]\left(\frac{d}{dt}\frac{\overrightarrow{v}}{c}\right)$
$\frac{dw}{dt}\simeq\left(1+\frac{9}{8}\frac{\left\|\overrightarrow{v}\right\|^2}{c^2}+\frac{3}{4}j\left(\frac{\overrightarrow{v}}{c}\right)j\left(\frac{\overrightarrow{v}}{c}\right)\right)\left(\frac{d}{dt}\frac{\overrightarrow{v}}{c}\right)$
$X_r\simeq-\frac{1}{2}\left(1+\frac{3}{4}\frac{\left\|\overrightarrow{v}\right\|^2}{c^2}\right)j\left(\frac{\overrightarrow{v}}{c}\right)\left(\frac{d}{dt}\frac{\overrightarrow{v}}{c}\right)+\left[1-\frac{1}{2}j\left(\frac{\overrightarrow{v}}{c}\right)j\left(\frac{\overrightarrow{v}}{c}\right)\right]\left(\frac{1}{a_r}+\frac{1}{2}j\left(r\right)+\frac{1}{4a_r}j\left(r\right)j\left(r\right)\right)\frac{dr}{dt}$
$X_w\simeq\left(1+\frac{\left\|\overrightarrow{v}\right\|^2}{c^2}-\frac{1}{2}j\left(\frac{\overrightarrow{v}}{c}\right)j\left(\frac{\overrightarrow{v}}{c}\right)\right)\left(\frac{d}{dt}\frac{\overrightarrow{v}}{c}\right)+\left(1+\frac{1}{2}\frac{\left\|\overrightarrow{v}\right\|^2}{c^2}\right)j\left(\frac{\overrightarrow{v}}{c}\right)\left(\frac{1}{a_r}+\frac{1}{2}j\left(r\right)+\frac{1}{4a_r}j\left(r\right)j\left(r\right)\right)\frac{dr}{dt}$

3.5.3 Motion of material bodies

It is possible to extend the concept of deformable solid to the framework of RG.

Representation of collection of particles by sections of the fiber bundle

In Physics it is common to have a collection of particles with identical characteristics following trajectories which do not cross, like in a beam. They can be represented by a section of P_G.

Let be $\sigma=\varphi_G\left(m,\sigma\left(m\right)\right)\in\mathfrak{X}\left(P_G\right)$ and a given observer who uses his standard gauge. Then the relations :
$m=\pi_G\left(\sigma\right)$
$U\left(m\right)=-\frac{c}{\left\langle\mathbf{Ad}_{\sigma(m)}\varepsilon_0,\varepsilon_0\right\rangle_{Cl}}\mathbf{Ad}_{\sigma(m)}\varepsilon_0\left(m\right)=\left(\mathbf{p}\left(m\right),-\frac{c}{\left\langle\mathbf{Ad}_{\sigma(m)}\varepsilon_0,\varepsilon_0\right\rangle_{Cl}}\mathbf{Ad}_{\sigma(m)}\varepsilon_0\right)$
$V^\alpha\left(m\right)=\sum_{i=0}^3P_i^\alpha\left(m\right)U^i\left(m\right)$
define a vector field $V\in\mathfrak{X}\left(TM\right)$.
$V\left(m\right)$ is a timelike, future oriented vector :
$\left\langle V,V\right\rangle=\left\langle U,U\right\rangle=\left\langle\frac{c}{\left\langle\mathbf{Ad}_{\sigma(m)}\varepsilon_0,\varepsilon_0\right\rangle_{Cl}}\mathbf{Ad}_{\sigma(m)}\varepsilon_0,\frac{c}{\left\langle\mathbf{Ad}_{\sigma(m)}\varepsilon_0,\varepsilon_0\right\rangle_{Cl}}\mathbf{Ad}_{\sigma(m)}\varepsilon_0\right\rangle_{Cl}$
$=-\left(\frac{c}{\left\langle\mathbf{Ad}_{\sigma(m)}\varepsilon_0,\varepsilon_0\right\rangle_{Cl}}\right)^2<0$
$\left\langle V,\varepsilon_0\left(m\right)\right\rangle=\left\langle U,\varepsilon_0\right\rangle=-\frac{c}{\left\langle\mathbf{Ad}_{\sigma(m)}\varepsilon_0,\varepsilon_0\right\rangle_{Cl}}\left\langle\mathbf{Ad}_{\sigma(m)}\varepsilon_0,\varepsilon_0\right\rangle_{Cl}=-c<0$

V define integral curves $q\left(\tau\right)=\Phi_V\left(\tau,a\right)$ with a parameter τ uniquely defined by a point $\Phi_V\left(0,a\right)=a$. The map $\tau\rightarrow q\left(\tau\right)=\varphi_o\left(\xi_0\left(\tau\right),\xi_1\left(\tau\right),\xi_2\left(\tau\right),\xi_3\left(\tau\right)\right)$ is defined by the differential equations :
$\frac{\partial\xi_\alpha}{\partial\tau}=V^\alpha\left(q\left(\tau\right)\right)\Rightarrow\frac{\partial\xi_0}{\partial\tau}=V^0\left(q\left(\tau\right)\right)=c\Rightarrow\xi_0\left(\tau\right)=c\tau+Ct$
$q\left(\tau\right)=\varphi_o\left(c\tau+Ct,\xi_1\left(\tau\right),\xi_2\left(\tau\right),\xi_3\left(\tau\right)\right)$
which is equivalent to say that the parameter on the integral curves is the time t of the observer
$V^\alpha\left(\Phi_V\left(t,a\right)\right)=\sum_{i=0}^3P_i^\alpha\left(\Phi_V\left(t,a\right)\right)U^i\left(\Phi_V\left(t,a\right)\right)$
From σ we have at each point an arrangement :
$e_i\left(m\right)=\mathbf{Ad}_{\sigma(m)}\varepsilon_i\left(m\right)=\left(\mathbf{p}\left(m\right),\mathbf{Ad}_{\sigma(m)}\varepsilon_i\right)$
and the motion along an integral curve, the derivative being with respect to the time of the observer :

$\forall i=0..3: \frac{de_i}{dt}=\left[\frac{d\sigma}{dt}\cdot\sigma^{-1},e_i\right]$
$\frac{dU}{dt}=\frac{U}{c}\left\langle\left[\frac{d\sigma}{dt}\cdot\sigma^{-1},U\right],\varepsilon_0\right\rangle_{Cl}+\left[\frac{d\sigma}{dt}\cdot\sigma^{-1},U\right]$
$\frac{dV^\alpha}{dt}=\sum_{i=0}^3\left[P'\right]_i^\alpha\frac{dU^i}{dt}$
So, a section $\sigma\in\mathfrak{X}(P_G)$ represents, in a given gauge, the motion of a collection of particles which follow trajectories belonging to the same vector field.

The same section $\sigma\in\mathfrak{X}(P_G)$ is represented, in a change of gauge $\mathbf{p}(m)\to\widetilde{\mathbf{p}}(m)=\mathbf{p}(m)\cdot\chi(m)^{-1}$ by $\sigma\to\widetilde{\sigma}=\chi(m)\cdot\sigma$ and defines the vector :

$\widetilde{e}_i(m)=\left(\widetilde{\mathbf{p}}(m),\mathbf{Ad}_{\widetilde{\sigma}(m)}\varepsilon_i\right)=\left(\mathbf{p}(m)\cdot\chi(m)^{-1},\mathbf{Ad}_{\chi(m)}\mathbf{Ad}_{\sigma(m)}\varepsilon_i\right)$
$\sim\left(\mathbf{p}(m),\mathbf{Ad}_{\sigma(m)}\varepsilon_i\right)=e_i(m)$
thus it represents the same motion.

Conversely, if we have such a collection of particles, then for a given observer, the tetrad of the particles are measured as :

$\forall i=0..3: e_i(q(t))=\mathbf{Ad}_{\sigma(q(t))}\varepsilon_i(q(t))$
and we can proceed to the same computations as in the first subsection :

$\sigma(q(t))$ is defined through $e_i(q(t))$ which are physical vectors, in a change of gauge : $\mathbf{p}(m)=\varphi_G(m,1)\to\widetilde{\mathbf{p}}(m)=\mathbf{p}(m)\cdot\chi(m)^{-1}:\sigma\to\widetilde{\sigma}=\chi(m)\cdot\sigma$

$e_i(q(t))=(\mathbf{p}(m),\mathbf{Ad}_\sigma\varepsilon_i)\sim\left(\widetilde{\mathbf{p}}(m),\mathbf{Ad}_{\chi(m)}\mathbf{Ad}_\sigma\varepsilon_i\right)=(\widetilde{\mathbf{p}}(m),\mathbf{Ad}_{\widetilde{\sigma}}\varepsilon_i)$
and we have a section of P_G.

Theorem 65 *The motion of a collection of particles whose trajectories belong to a vector field can uniquely be represented, for any observer, by a section of P_G.*

A section of P_G can represent, for any observer, the motion of particles whose trajectories belong to a vector field.

This representation is crucial : it is at the foundation of the concept of matter field that we will develop in the next chapter.

Derivatives

With this representation one can address the problem of variable trajectories. The tetrad of the observer is assumed to be fixed, so that the derivatives of $V^\alpha=\sum_{i=0}^3\left[P'\right]_i^\alpha U^i$ can be computed by

$\partial_\beta V^\alpha(m)=\sum_{i=0}^3\left[P'\right]_i^\alpha\partial_\beta U^i(m)$
With a computation as above :
$\partial_\beta U=\frac{U}{c}\left\langle\left[\partial_\beta\sigma\cdot\sigma^{-1},U\right],\varepsilon_0\right\rangle_{Cl}+\left[\partial_\beta\sigma\cdot\sigma^{-1},U\right]$
$\left[\partial_\beta\sigma\cdot\sigma^{-1},U\right]$ is computed in the Clifford Algebra. The result is a vector.
$\left[\partial_\beta\sigma\cdot\sigma^{-1},U\right]=\sum_{j=0}^3\left[\partial_\beta\sigma\cdot\sigma^{-1},U\right]^j\varepsilon_j$
$\partial_\beta V^\alpha=\sum_{j=0}^3P_j^\alpha\left\{\left[\partial_\beta\sigma\cdot\sigma^{-1},U\right]^j+\frac{U^j}{c}\left\langle\left[\partial_\beta\sigma\cdot\sigma^{-1},U\right],\varepsilon_0\right\rangle_{Cl}\right\}$
$=\frac{V^\alpha}{c}\left\langle\left[\partial_\beta\sigma\cdot\sigma^{-1},U\right],\varepsilon_0\right\rangle_{Cl}+\sum_{j=0}^3P_j^\alpha\left[\partial_\beta\sigma\cdot\sigma^{-1},U\right]^j$
Moreover : $V^0=c\Rightarrow\partial_\beta V^0=0\Rightarrow\left\langle\left[\partial_\beta\sigma\cdot\sigma^{-1},U\right],\varepsilon_0\right\rangle_{Cl}=-\sum_{j=0}^3P_j^0\left[\partial_\beta\sigma\cdot\sigma^{-1},U\right]^j$
$\partial_\beta V^\alpha=-\frac{V^\alpha}{c}\sum_{j=0}^3P_j^0\left[\partial_\beta\sigma\cdot\sigma^{-1},U\right]^j+\sum_{j=0}^3P_j^\alpha\left[\partial_\beta\sigma\cdot\sigma^{-1},U\right]^j$

$$\alpha>0:\partial_\beta V^\alpha=\sum_{j=0}^3\left(P_j^\alpha-\frac{1}{c}P_j^0V^\alpha\right)\left[\partial_\beta\sigma\cdot\sigma^{-1},U\right]^j \tag{3.66}$$

$\partial_\beta\sigma\cdot\sigma^{-1}=\upsilon(\delta_\beta X_r,\delta_\beta X_w)$ depends on the chart used in the Clifford algebra.

Integral curves of the section

The vector field associated to a section has for components in a standard chart $m = \varphi_o(t,\xi), \xi = (\xi_1,\xi_2,\xi_3)$:

$\beta = 1,2,3 : V^\beta(t,\xi) = \sum_{j=1}^3 Q_j^\beta(t,\xi) U^j(t,\xi)$

By construct, they do not depend on the gauge.

For instance with the Clifford chart :

$\sigma = \sigma_w(m) \cdot \sigma_r(m) = \epsilon(a_w + \upsilon(0, w(m))) \cdot \epsilon(a_r + \upsilon(r(m), 0))$

$\sigma_w(m)$ is defined by a map $w : \mathbb{R}^4 \to \mathbb{R}^3 :: w(t,\xi)$ and $U(t,\xi) = \frac{a_w}{2a_w^2-1}[w(t,\xi)] = \epsilon\frac{\sqrt{1+\frac{1}{4}w^t w}}{1+\frac{1}{2}w^t w}[w(t)]$.

The integral curves passing through a point $A = \varphi_o(ct, \zeta^1, \zeta^2, \zeta^3)$, have the equation :

$q(\tau) = \Phi_V(\tau, A) = \varphi_o(c(t+\tau), f(\tau,\zeta))$

with $\zeta = (\zeta^1, \zeta^2, \zeta^3)$

$f(\tau,\zeta) = (f_1(\tau,\zeta), f_2(\tau,\zeta), f_3(\tau,\zeta))$

where the functions $f_\beta(\tau,\zeta)$ are solution of the differential equation :

$\beta = 1,2,3 : V^\beta(c(t+\tau), f(\tau,\zeta)) = \frac{\partial f_\beta}{\partial \tau}$

$f_\beta(0,\zeta) = \zeta_\beta$

$f_\beta(\tau,\zeta) = \zeta_\beta + \sum_{j=1}^3 \int_0^\tau V^\beta(t+s, f(s,\zeta))\, ds$

The integral curves depend only on $\sigma_w(m)$. So the product of σ by any section $s \in Spin(3)$ gives the same curves.

Representation of material bodies in GR

In Mechanics a material body is made of "material points" that is elements of matter whose location is a single geometric point, and changes with time in a consistent way : their trajectories do not cross, so that the material body keeps its cohesion. By adding a frame to each material point one can models its deformation by a deformation tensor. The generalization to GR is immediate.

Definition 66 *A material body is a collection of particles whose trajectories are integral curves of a time like, future oriented vector field V, and which belong, at some point on their trajectories to a compact subset of a spatial hypersurface of M. It can be represented by a section $\sigma \in P_G$ with compact support.*

The vector field defines a parameter τ which is the proper time of the material body. If at τ_0 all the particles are in ω, then, because $\Phi_V(\tau, a), a \in \omega$ is a diffeomorphism, the image of ω at any other time is still compact. The definition is independent of any observer.

The set $\widehat{\omega} \subset M$ swept by the particles is the support of σ. It is such that its intersection by any spatial hypersurface is still a compact hypersurface, but all the observers do not see the same body[11] : they see $\widehat{\omega} \cap \Omega_3(t)$. They see always the same body if their velocity belongs to the vector field V.

Conversely, given a section $\sigma \in \mathfrak{X}(P_G)$ it defines a vector field of trajectories with respect to any observer : the parameter on the integral curves is then the time of the observer, and σ defines a solid whose physical body ω at any time belongs to the present of the observer : $\omega(t) \subset \Omega_3(t)$.

It is usually more convenient to define a deformable solid from the point of view of an observer $\mathcal{B}$ attached to the solid : $\varepsilon_0 = \frac{1}{\sqrt{-\langle V,V\rangle}}V$, and do the computations in a spherical system of coordinates. Then proceed to a change of gauge to represent the motion of the solid from the point of view of another observer. It requires just the map $\chi : \mathcal{B} \to O \in Spin(3,1)$. The composition of motions in the GR framework is thus easy.

Of course particles follow such trajectories only if they have similar physical properties, and are submitted to adequate forces. We will come back on these issues in the following.

[11] This point is a the origin of many misunderstandings in simplist experiments based on material objects (usually trains).

This general definition applies to solids, in the usual meaning, but also to fluids, which are composed of material points which travel along trajectories which do not cross.

Deformation tensor :

We can define a deformation tensor, similar to what is done in Newtonian Mechanics. In Galilean Geometry the deformation tensor is defined by the change $\frac{\partial}{\partial t} e_i(q,t)$ of $e_i(q,t)$ with respect to $e_i(q,t)$ .The equivalent in our framework is $\frac{d\sigma}{dt} \cdot \sigma^{-1} = \sum_{\alpha=0}^{3} V^\alpha \partial_\alpha \sigma \cdot \sigma^{-1}$ whose matrix is $[\upsilon(X_r, X_w)] = [K(X_w)] + [J(X_r)] \in so(3,1)$:

$[K(X_w)] = \begin{bmatrix} 0 & X_w^t \\ X_w & 0 \end{bmatrix}, [J(X_r)] = \begin{bmatrix} 0 & 0 \\ 0 & j(X_r) \end{bmatrix}$

The deformation tensor has a symmetric ($[K(X_w)]$) and an antisymmetric ($[J(X_r)]$) part, as the usual deformation tensor.

Rigid solid :

The arrangement of each individual particle, represented by σ, is not necessarily identical. A **rigid solid** can be defined as a solid such that the motion is identical at each point :

$\forall x \in \omega(0) : \frac{d\sigma}{dt} \cdot \sigma^{-1}(\Phi_V(t,x)) = \upsilon(Y_r(t), Y_w(t))$

$\Leftrightarrow \sigma(\Phi_V(t,x)) = s(t) \cdot \sigma(\Phi_V(0,x))$ with $s(t) \in Spin(3,1)$

and $s(t)$ represents the arrangement of the rigid solid with respect to the observer. Then the deformation tensor depends only on t.

3.5.4 Symmetries

Particle

The geometric characteristics of a particle can be represented by a map : $\sigma : \mathbb{R} \to P_G :: \sigma(t) = \varphi_G(q(t), \sigma(t))$

For a particle symmetric motions are essentially periodic motions, which can be understood in two different, complementary, ways :

- periodic instantaneous motions : $\sigma(t+T) = \sigma(t)$
- periodic trajectories : $q(t) = \varphi_0(ct, \xi(t)) : \xi(t+T) = \xi(t)$ (there is no 4 dimensional loop in a trajectory)

Periodic instantaneous motions

$\sigma(t+T) = \varphi_G(q(t+T), \sigma(t))$

The motion is periodic if the arrangement is the same at $q(t)$ and $q(t+T)$.

The arrangement $\sigma(t)$ can be equivalently defined by the tetrad $(e_i(t), i = 0..3)$ attached to the particle so : $e_i(q(t)) = \mathbf{Ad}_{\sigma(q(t))}\varepsilon_i = e_i(q(t+T)) = \mathbf{Ad}_{\sigma(q(t+T))}\varepsilon_i$

The vectors $e_i(q(t+T)), e_i(q(t))$ have not necessarily the same components in the holonomic chart, but they are defined by the same rotation with respect to the local tetrad.So the symmetry is with respect to the tetrad at each point, and from this definition, it does not depend on the choice of the tetrad.

The trajectory $q : \mathbb{R} \to M :: q(t) = \varphi_o(ct, \xi(t))$ is defined separately.

A bonded particle is such that $x(t) = Ct$. So $U = 0$. In the chart $\sigma = \sigma_w \cdot \sigma_r = \epsilon(a_w + \upsilon(0,w)) \cdot \epsilon(a_r + \upsilon(r,0))$ we have $w = 0$ and a periodic rotation if $r(t+T) = r(t)$,

$\frac{d\sigma}{dt} \cdot \sigma^{-1} = \upsilon\left(\frac{dr}{dt}, 0\right) \in so(3)$.

$\frac{de_i}{dt} = \left[\upsilon\left(\frac{dr}{dt}, 0\right), e_i\right]$

Periodic trajectories

Using the relations above :

$U = -\frac{c}{\langle \mathbf{Ad}_\sigma \varepsilon_0, \varepsilon_0 \rangle_{Cl}} \mathbf{Ad}_\sigma \varepsilon_0 \Rightarrow U(t+T) = U(t) \Rightarrow u(t+T) = u(t)$

and in the standard chart :

$V^0(t) = [U(t)]^0 \partial\xi_0(q(t)) = c = \frac{\partial \varphi_o}{\partial t}$

$\frac{dx}{dt}(t) = v(t) = \sum_{j,\beta=1}^3 Q_j^\beta(q(t)) u^j(t) \partial\xi_\beta$

$\xi_\beta(t+T) = \xi_\beta(t) + \int_t^{t+T} \sum_{j,\beta=1}^3 Q_j^\beta(q(s)) u^j(s) ds$

So we do not have necessarily $\xi(t+T) = \xi(t)$, that is a spatial periodic trajectory in $\Omega_3(0)$, but the spatial speed, $u(t)$ as measured with respect to the tetrad, is periodic. If the metric does not depend on t then the tetrad can be chosen such that it does not depend on t, and the spatial trajectory is periodic if

$\int_t^{t+T} \sum_{j,\beta=1}^3 Q_j^\beta(q(s)) u^j(s) ds = 0$

Phase velocity

In our representation of the motion of a particle the 2 components r, w which define σ are independent. The arrangement of the particle is only submitted to the orientation of its vector e_0. And a particle can have a periodic rotational motion without a periodic trajectory, as seen above. The **phase velocity** is then defined as :

$\omega(t) = \frac{1}{T}\int_t^{t+T} \sqrt{\langle v(q(\tau)), v(q(\tau))\rangle_3} d\tau$

with the spatial speed $v(q(\tau))$.

$\langle v(q(\tau)), v(q(\tau))\rangle_3 = \langle u(\tau), u(\tau)\rangle_3 = [u(\tau)]^t [g_3(q(\tau))][u(\tau)]$

In SR $[g_3(q(\tau))] = I$:

$\omega(t) = \frac{1}{T}\int_t^{t+T} \sqrt{[u(\tau)]^t [u(\tau)]} d\tau$

In a periodic instantaneous motion the phase velocity is constant

For any motion one can compute the integral :

$\widehat{\phi}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \sqrt{\langle v(q(\tau)), v(q(\tau))\rangle_3} e^{-i\omega\tau} d\tau$

which gives the decomposition of the motion in instantenous periodic motions of period $T = \frac{2\pi}{\omega}$.

Symmetries in motions defined by a section

For any observer, a section $\sigma \in \mathfrak{X}(P_G)$ defines a material body whose proper time is the same as the time of the observer. Thus symmetries are defined with respect to a map : $F : M \to M$, which can be an isometry or not.

Spatial symmetries

There is a spatial symmetry if there is a map $f : \mathbb{R}^3 \to \mathbb{R}^3$ such that :

$\forall t : \sigma(\varphi_o(ct, f(\xi))) = \sigma(\varphi_o(ct, \xi))$

The vectors $e_i(m)$ and the velocity $U(m)$ are identical at $\varphi_o(ct, \xi)$ and $\varphi_o(ct, f(\xi))$ with respect to the local tetrad. If moreover f is an isometry on $\Omega_3(0)$ the components of the vectors are identical.

If f is a continuous action of a group G on $\Omega_3(0)$ then the subsets $\Omega_3(0)/R$ has the structure of a manifold with dimension $3 - \dim G$. We have structures similar to crystals.

Dynamic symmetries

They involve a map F which is not restricted to $\Omega_3(0)$. There is no loop for the trajectory of a particle, so $f(m)$ must be in the future of any observer, or equivalently the matrix (in any chart) $[F'(m)]^t [g(m)][F'(m)]$ must be definite negative.

The action F can be defined by the flow of a vector field $W \in \mathfrak{X}(TM) : F(\tau)(m) = \Phi_W(\tau, m)$ then this is an action of $\mathbb{R}$ along the integral curves of W (and not on M). A symmetry is then

$\sigma(\Phi_W(\tau, m)) = \sigma(m)$ which implies that W is identical to the vector field V induced by σ, and $\sigma = Ct$.

More interesting are periodic motions along some integral curves defined by V. In the chart $\sigma(m) = \sigma_w \cdot \sigma_r = \epsilon(a_w + \upsilon(0, w)) \cdot \epsilon(a_r + \upsilon(r, 0))$ the integral curves have for equations :

$q(t) = \varphi_o(ct, \xi(t))$

$\left[\frac{d\xi}{dt}\right] = \epsilon \frac{\sqrt{1+\frac{1}{4}w^t w}}{1+\frac{1}{2}w^t w} \sum_{j=1}^3 \left[Q_j^\beta(t, x(t))\right][w(t, x(t))]$

where $w : \mathbb{R}^4 \to \mathbb{R}^3$ is a given map.

If there is a map : $\eta : \mathbb{R} \to \mathbb{R}^3$ which is periodic and solution of this differential equation, then the rotational motion given by $a_r + \upsilon(r(ct, \eta(t)), 0)$ is also periodic, with the same period. We have a periodic motion (in the meaning above) along an integral curve $q(t) = \varphi_o(t, \eta(t))$ passing through the point $O = \varphi_o(0, \eta(0))$.

There is not necessarily a solution, and each one is valid for curves passing through the point $\varphi_o(0, y(0))$. There can be several solutions, with different periods.

For any section σ such that $\sigma(m) \in Spin(3)$ the integral curves are $q(t) = \varphi_o(ct, \xi)$ with $\xi = Ct$. Then the motion is periodic along the integral curve $q(t) = \varphi_o(ct, \xi)$ iff $r(t + T, \xi) = r(t, \xi)$.

One can build a generic periodic section as follows. Let $w : \mathbb{R} \to \mathbb{R}^3$ be any map, and $y : \mathbb{R} \to \mathbb{R}^3$ a solution of the differential equation : $\left[\frac{dy}{dt}\right] = \epsilon \frac{\sqrt{1+\frac{1}{4}w^t w}}{1+\frac{1}{2}w^t w} w(t)$.

Define the path : $q : \mathbb{R} \to M$ by $q(t) = \varphi_o(ct, \xi(t))$ with :

$\xi(t) = \xi(0) + \int_0^t \sum_{j=1}^3 \left[Q_j^\beta(s, y(s))\right][y(s)]\, ds$

$\Rightarrow \frac{d\xi}{dt} = \epsilon \frac{\sqrt{1+\frac{1}{4}w^t w}}{1+\frac{1}{2}w^t w} \sum_{j=1}^3 \left[Q_j^\beta(s, y(t))\right][w(t)]$

Then the corresponding curves are integral curves of the section $s(\varphi_o(ct, \xi)) = a_w + \upsilon(0, w(t))$.

Take any map : $T : \Omega_3(0) \to \mathbb{R}$

and any map : $r : \mathbb{R}^4 \to \mathbb{R}^3$ periodic with respect to t :

$\forall x \in \Omega_3(0), x = \varphi_\Omega(\xi) : r(t + T(\xi), \xi) = r(t, \xi)$

and define the section $\sigma_r(\varphi_o(ct, \xi)) = a_r + \upsilon(r(t, \xi), 0),$. Then the section $\sigma = s \cdot \sigma_r$ has the path $q(t) = \varphi_o(ct, \xi(t)) = \Phi_V(t, \xi(0))$ as integral curves (they depend only on s), and for each $\varphi_o(ct, \xi(0))$ the instantaneous motion is periodic with period $T(\xi)$ depending on x. And if w is periodic then the full motion is periodic. Of course there could be many variant in this construct.

Fourier transform on $Cl(3,1)$

The complex vector subspace F of $Cl(3,1)$ generated by the vectors $(1, \overrightarrow{\kappa}_a, a = 1, 2, 3)$:

$F = \left\{A + \sum_{a=1}^3 Z^a \overrightarrow{\kappa}_a\right\}$

is a 4 dimensional complex Hilbert space with the hermitian scalar product :

$\langle \phi_1, \phi_2 \rangle = \overline{A_1} A_2 + \overline{Z_1^t} Z_2$

The set of maps :

$\phi : \mathbb{R} \to F :: \phi(t) = A(t) + Z(t)$ such that $\int_{\mathbb{R}} \langle \phi(t), \phi(t) \rangle dt < \infty$ is a complex Hilbert space, infinite dimensional and separable. So one can define Fourier integrals :

$\widehat{\phi}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \phi(t) e^{-i\omega t} dt$

with all the usual properties (derivation, ...).

Any map : $\sigma : \mathbb{R} \to Spin(3,1)$ can be written as a map : $\sigma : \mathbb{R} \to F$ with the condition $A(t)^2 = 1 - \frac{1}{4} Z(t)^t Z(t)$. However the Fourier transform $\widehat{\sigma}$ usually does not belong to $Spin(3,1)$.

Any periodic map : $\sigma : t \to F :: \sigma(t + T) = \sigma(t)$ can be expressed as a Fourier series.

$\sigma : \mathbb{R} \to Spin(3,1) :: \sigma(t) = A(t) + Z(t)$ where $Z(t + T) = Z(t)$ for some fixed period. Then $A^2(t + T) = 1 - \frac{1}{4} Z(t + T)^t Z(t + T) = A^2(t)$

Z can be written :

$Z(t) = \sum_{n \in \mathbb{Z}} \widehat{Z}(n) \exp in\omega t$ with $\widehat{Z}(n) = \frac{1}{T} \int_0^T Z(t) \exp(-in\omega t)\, dt$ and $\omega = \frac{2\pi}{T}$

$Z(0) = \sum_{n\in\mathbb{Z}} \widehat{Z}(n)$
$A(t) = \sum_{n\in\mathbb{Z}} \widehat{A}(n) \exp in\omega t$ with $\widehat{A}(n) = \frac{1}{T}\int_0^T A(t)\exp(-in\omega t)\,dt$ and $\omega = \frac{2\pi}{T}$
$A(t)^2 = 1 - \frac{1}{4}Z(t)^t Z(t)$
$Z(t)^t Z(t) = \sum_{n\in\mathbb{Z}}\sum_{p\in\mathbb{Z}} \widehat{Z}(n-p)^t \widehat{Z}(p)\exp in\omega t = 4\left(1 - \sum_{n\in\mathbb{Z}}\sum_{p\in\mathbb{Z}} A(n-p)A(p)\exp in\omega t\right)$
$n \neq 0 : \sum_{p\in\mathbb{Z}} \widehat{Z}(n-p)^t \widehat{Z}(p) - 4A(n-p)A(p) = 0$
$\sum_{p\in\mathbb{Z}} \widehat{Z}(-p)^t \widehat{Z}(p) - 4A(-p)A(p) = 4$
$\widehat{A}(n)+\widehat{Z}(n) \in F$, but we do not have necessarily $\widehat{Z}(n)^t \widehat{Z}(n) = 4\left(1-\widehat{A}(n)^2\right)$ so $\widehat{A}(n)+\widehat{Z}(n) \in Cl(3,1)$ but does not necessarily belong to $Spin(3,1)$.

We have similarly :
$\frac{d\sigma}{dt}\cdot\sigma^{-1} = \delta Z(t) = \sum_{n\in\mathbb{Z}} \widehat{\delta Z}(n)\exp in\omega t$
with $\widehat{\delta Z}(n) = \frac{1}{T}\int_0^T \delta Z(t)\exp(-in\omega t)\,dt$
In a continuous motion :
$\delta Z(t) = D(Z(t))\frac{dZ}{dt}$
$\sum_{n\in\mathbb{Z}} \widehat{\delta Z}(n)\exp in\omega t = i\omega D(Z(t))\sum_{n\in\mathbb{Z}} n\widehat{Z}(n)\exp in\omega t$
$\widehat{\delta Z}(n) = i\omega n D(Z(t))\widehat{Z}(n)$

3.5.5 Jet Bundles

The arrangement is represented by an element of the Spin Group and the motion by an element of the Lie Algebra and both are related by the derivatives. Moreover one goes from the associated vector bundle $P_G[\mathbb{R}^4, \mathbf{Ad}]$ to the holonomic basis by the tetrad. It is useful to combine all this in a formalism which underlines their relation. This is done by Jet Bundles which is a general Mathematical Theory with many applications. This is the classic formulation of lagrangians, the framework used in differential equations and variational derivatives, and it enables to represent non continuous processes.

Definition

In Differential Geometry one avoids as much as possible the coordinates expressions. But this is difficult when dealing with partial derivatives. The r-jet formalism provides a convenient solution, which goes beyond the computational issue. See Maths.26 for more.

For any r differentiable map $f \in C_r(M;N)$ between manifolds, the partial derivatives $\frac{\partial^s f}{\partial\xi^{\alpha_1}...\partial\xi^{\alpha_s}}$ at a point m are s symmetric linear maps from the tangent space T_mM to the tangent space T_pN. As any linear map their expression in holonomic bases is a set of scalars $f^i_{\alpha_1...\alpha_s}$, *symmetric* in the indices $\alpha_1,..\alpha_s$.

The relation of equivalence on $C_r(M;N)$:
$f \sim g \Leftrightarrow f(m) = g(m) = p, ...\frac{\partial^s f}{\partial\xi^{\alpha_1}...\partial\xi^{\alpha_s}}(m) = \frac{\partial^s g}{\partial\xi^{\alpha_1}...\partial\xi^{\alpha_s}}(m), s = 1...r, \alpha_k = 1..\dim M$
defines classes of equivalences of maps f, g which have the same value and partial derivative at m up to the order r. They are characterized by the set of scalars :
$j^r = \left(z^i_{\alpha_1...\alpha_s} \in \mathbb{R}, s = 1...r, \alpha_k = 1..\dim M, i = 1..\dim N\right) \in J^r_m(M,N)_p$
$z^i_{\alpha_1...\alpha_s}$ symmetric in the indices $\alpha_1,..\alpha_s$

The set $J^r_m(M,N)_p$ is a vector space. The $z^i_{\alpha_1...\alpha_s}$ are the components of symmetric tensors belonging to $\odot^s T_mM^* \otimes T_pN$.

A r jet with source m and target p is a set $j^r_{m,p} = (m,p,j^r)$ and more generally a r jet is a map $j^r(m) = (m, p(m), j^r(m))$

The r jet prolongation of f is the map :
$J^r f(m) = \left(m, f(m), \frac{\partial^s f^i}{\partial\xi^{\alpha_1}...\partial\xi^{\alpha_s}}(m), s = 1...r, \alpha_k = 1..\dim M, i = 1..\dim N\right)$

A key point is that any map f has a r jet prolongation, which is a r jet, but conversely in a r jet *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 f. The distinction between $\frac{\partial f^i}{\partial\xi^{\alpha_1}...\partial\xi^{\alpha_s}}$ and $z^i_{\alpha_1...\alpha_s}$ is useful : a differential equation is a relation between components of a r-jet : $L\left(m, z, z^i_{\alpha_1...\alpha_s}\right) = 0$ and a solution is a map f of $C_r(M;N)$ such that $L\left(m, f(m), \frac{\partial^s f}{\partial\xi^{\alpha_1}...\partial\xi^{\alpha_s}}\right) = 0$.

Fiber bundles $P(M,E,\pi)$ are manifolds, so we can implement the principle above by taking as maps sections on P. They are defined by :

$S : M \to P :: S(m) = \varphi_P(m, z(m))$

and r jets on P are defined by r jets prolongations of z.

The coordinates of $z(m) \in E$ are $z^i, i = 1...\dim E$ in a chart $\{z^i\} = \varphi_E(z)$ of E.

The partial derivatives $\frac{\partial^s z}{\partial\xi^{\alpha_1}...\partial\xi^{\alpha_s}}$ are linear maps whose *components in charts* of M, E are scalars : $z^i_{\alpha_1...\alpha_s}$ with the condition that they are symmetric in the indices $\alpha_1, ...\alpha_s$.

The r jet prolongation J^rP of the fiber bundle $P(M,E,\pi)$ is the vector bundle :

$J^rP\left(P, J^r_0\left(\mathbb{R}^{\dim M}, E\right)_0, \pi^r\right)$ with basis P, fiber the vector space :

$J^r_0\left(\mathbb{R}^{\dim M}, V\right)_0 = \left\{z^i_{\alpha_1...\alpha_s} \in \mathbb{R}, s = 1...r, \alpha_k = 1..\dim M, i = 1..\dim E\right\}$ and projection : π^r : $J^rP \to P$.

A section on J^rP is a map : $j^rp(m) = \left(p(m), z^i_{\alpha_1...\alpha_s}(m)\right)$

A section S on P gives a section on $J^rP : J^rS = \left(S(m), \frac{\partial^s z^i}{\partial\xi^{\alpha_1}...\partial\xi^{\alpha_s}}(m)\right)$

Two sections S, S' belong to the same r jet if the value of z, z' and their r derivatives are equal.

J^1P is an affine bundle with fiber $TM^* \otimes VE$ where VE is the vertical bundle (isomorphic to TE) :

$j^1p(m) = \left(p(m), z^i(m), z^i_\alpha(m), \alpha = 1..\dim M, i = 1..\dim E\right)$

It has an affine structure because the element $p(m)$ is common.

Jet prolongation of a vector bundle

The r-jet prolongation of a vector bundle (associated or not) $E(M;V;\pi)$ is a vector bundle. A section $\mathbf{X} \in \mathfrak{X}(E)$ is defined by a map :

$\mathbf{X} : M \to E :: X(m) = (m, \sum_{i=1}^n u_i(m) e_i)$

r-jets are defined from the components in the charts, so here the partial derivatives of the map : $X : M \to V$ where V is a fixed vector space. The r-jet prolongation of $\mathbf{X}$ is then defined by

$\left\{m, u^i, X^i_{\alpha_1...\alpha_s}, \alpha_p = 0...3, i = 1...\dim V, s = 0...r\right\}$

The holonomic basis of the vector bundle J^rE is the set of vectors

$\{e_i, e_i^{\alpha_1...\alpha_s}, \alpha_p = 0...3, i = 1...\dim V, s = 0...r\}$ localized at $m \in M$. For the 1st jet prolongation this is just : $\{e_i, e_i^\alpha, \alpha = 0...3, i = 1...\dim V\}$ and one can write :

$J^1X = \{m, X(m), \delta_\alpha X(m), \alpha = 0...3\}$ with : $\delta_\alpha X(m) = \sum_{i=1}^n \delta_\alpha X_i e_i^\alpha$

$\Leftrightarrow J^1X = \{m, X(m), \delta X(m)\}$ with : $\delta X(m) = \sum_{\alpha=0}^3 \sum_{i=1}^n \delta_\alpha X_i e_i^\alpha$

or equivalently $\delta_\alpha X(m) = \sum_{\alpha=0}^3 \sum_{i=1}^n \delta_\alpha X_i e_i$. A section of J^1E is then equivalent to 5 independent sections of E.

The jet prolongation of maps : $X : [0,T] \to E :: X(t)$ is a map : $[0,T] \to J^1E :: J^1X = \left(q(t), X(t), \frac{dX}{dt}\right)$ and to a section of J^1E corresponds a map : $[0,T] \to J^1E :: J^1X = (q(t), X(t), \delta X)$ where δX is independent from X.

The great advantage of the r-jet formalism in Physics, and specially in GR, is that it provides a tool to deal with the derivatives of the *components* and *not* :

$\frac{d}{dt}X(t) = \frac{d}{dt}\left(\sum_{i=1}^n u_i(t) e_i(q(t))\right)$

valued in the tangent bundle of E, which involves the derivative $\frac{d}{dt}e_i(q(t))$.

If the vector bundle is associated to a principal bundle, the link is done through maps :

$e_i(m) = \varphi_E(\mathbf{p}(m), e_i) = (\mathbf{p}(m), e_i)$

For instance with $P_G\left[\mathbb{R}^4, \mathbf{Ad}\right]$ the link is done through the tetrad :
$\varepsilon_i(m) = \varphi_E(\mathbf{p}(m), \varepsilon_i) = (\mathbf{p}(m), e_i) = \sum_{\alpha=0}^{3} P_i^\alpha(m)\,\partial\xi_\alpha$
Then the 1st jet extension can be similarly defined as :
$J^1X = \{m, X(m), \delta_\alpha X(m), \alpha = 0..3\}$ with $\delta_\alpha X(m) = \sum_{\alpha=0}^{3}\sum_{i=1}^{3} \delta_\alpha X_i(m)\,\varepsilon_i(m)$
$\delta_\alpha X(m) = \sum_{\alpha=0}^{3}\sum_{i=1}^{3} \delta_\alpha X_i(m) \sum_{\beta=0}^{3} P_i^\beta(m)\,\partial\xi_\beta = \sum_{\alpha,\beta,i=0}^{3} \delta_\alpha X_i(m)\,P_i^\beta(m)\,\partial\xi_\beta$

Differential operators

The principal application of the r-jet formalism is in Differential Equations and Differential Operators.

A r differential operator is a base preserving morphism $D : \mathfrak{X}(J^rE_1) \to \mathfrak{X}(E_2)$ between two vector bundles (Maths.32). It maps fiberwise $Z(m)$ in J^rE_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^rE_1). Combined with the map : $J^r : \mathfrak{X}(E_1) \to \mathfrak{X}(J^rE_1)$, $D \circ J^r$ maps sections on E_1, to sections on E_2.

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 $Z \in \mathfrak{X}(J^rE_1)$ 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...

3.5.6 Jet representation of the motion

Jet prolongation of P_G

A section σ of a principal bundle $P(M, G, \pi)$ is given by a map : $g : M \to G$:
$\sigma(m) = \varphi_P(m, g(m))$

The tangent space T_gG is isomorphic to the Lie algebra T_1G so it is more natural to define the 1st jet prolongation by the left or the right derivatives, $L'_{g^{-1}}g(X)$ or $R'_{g^{-1}}g(X)$ which belong to the Lie algebra. For the principal bundle the Lie algebra $T_1Spin(3,1)$ belongs to the Clifford Algebra.

The Spin Group, is not a vector space, but a 6 dimensional manifold embedded in $Cl(3,1)$.

At each point we have a copy of the Clifford algebra with the Clifford bundle $Cl(TM)$, which is a vector bundle.

An element of the 1st jet extension $J^1Cl(TM)$ can be represented by $(m, X, Y_\alpha, \alpha = 0..3)$ where X, Y_α are vectors of the Clifford algebra.

The 1st jet prolongation of a section $\sigma \in \mathfrak{X}(P_G)$ can be represented by :
$J^1\sigma(m) = \left(m, \sigma, \partial_\alpha\sigma \cdot \sigma^{-1}, \alpha = 0..3\right) \in J^1Cl(TM)$
and a section of $J^1Cl(TM)$ such as :
$j^1\sigma(m) = (m, \sigma, \upsilon(X_{r\alpha}, X_{w\alpha}), \alpha = 0..3) \in J^1Cl(TM)$
can represent a section of J^1P_G.

$Cl(TM)$ is a vectorial bundle associated to P_G, so the components $\sigma, \upsilon(X_{r\alpha}, X_{w\alpha})$ belong to the fixed vector space $Cl(3,1)$ and change according to the usual rules in a change of gauge.

Motion of a deformable solid

A deformable solid can be represented by a section $\sigma \in \mathfrak{X}(P_G)$. By definition the motion is continuous, and the section $\sigma \in \mathfrak{X}(P_G)$ defines the section $J^1\sigma \in \mathfrak{X}\left(J^1P_G\right)$ thus : $\upsilon(X_{r\alpha}(m), X_{w\alpha}(m)) =$

$\partial_\alpha \sigma \cdot \sigma^{-1}$ and in the jet formalism we have

$$\sigma \in \mathfrak{X}(P_G) \to J^1\sigma = \left(m, \sigma(m), \partial_\alpha \sigma \cdot \sigma^{-1}, \alpha = 0...3\right) \in J^1 Cl(TM)$$

and a section of $J^1 Cl(TM)$ corresponding to the motion of a deformable solid reads :

$$j^1\sigma : M \to J^1 Cl(TM) :: j^1\sigma(m) = (m, \sigma(m), \upsilon(X_{r\alpha}(m), X_{w\alpha}(m)), \alpha = 0...3) \tag{3.67}$$

The trajectories and the arrangement are defined by $\sigma(m)$ as above,
$U = -\frac{c}{\langle \mathbf{Ad}_\sigma \varepsilon_0, \varepsilon_0 \rangle_{Cl}} \mathbf{Ad}_\sigma \varepsilon_0$
$V^\alpha = \sum_{i=0}^3 [P']_i^\alpha U^i$
$V = \frac{dq}{dt} = c\varepsilon_0 + \overrightarrow{v}$
which does not involve the derivatives.
The motion itself is such that :
$\forall i, \alpha = 0..3$:
$\delta_\alpha e_i = [\upsilon(X_{r\alpha}, X_{w\alpha}), e_i]$
$\delta_\alpha U = \frac{U}{c} \langle [\upsilon(X_{r\alpha}, X_{w\alpha}), U], \varepsilon_0 \rangle_{Cl} + [\upsilon(X_{r\alpha}, X_{w\alpha}), U]$
$\delta_\alpha V^\beta = \sum_{i=0}^3 [P']_i^\beta \delta_\alpha U^i$
so the motion is not necessarily continuous : the trajectory is differentiable, in the meaning that the tangent V is always defined through σ, but not necessarily continuously differentiable : physically it addresses the case where the particle takes, at a point, another direction.

This representation is useful to model discontinuous processes, but also necessary, at least from a mathematical point of view, when the trajectories are themselves a variable in a model (such as with a lagrangian).

Particles

The representation depends on the problem.

i) If one considers single, isolated particles, then the motion is represented by a map :
$\sigma : \mathbb{R} \to P_1 G : \sigma(t) = \varphi_G(q(t), \sigma(t))$
And by extension we will write :
$J^1\sigma = \left(q(t), \sigma(t), \frac{d\sigma}{dt} \cdot \sigma^{-1}\right) \in J^1 Cl(TM)$
$j^1\sigma : \mathbb{R} \to J^1 Cl(TM) :: j^1\sigma(t) = (q(t), \sigma(t), \upsilon(X_r(t), X_w(t)))$
If the motion is continuous then $\upsilon(X_r(t), X_w(t)) = \frac{d\sigma}{dt} \cdot \sigma^{-1}$.

This assumes that the trajectory is known, which is the case for bonded particles (they are spatially immobile, then $w = 0$ and the motion is limited to a rotation).

ii) Usually the trajectory is itself part of the problem. Then the convenient representation is actually through a section of P_G. It provides a vector field, to which the trajectory must belong, that is a "general solution" which is adjusted to the problem at hand through the initial conditions (which covers location, arrangement and tangent to the trajectory) . So it can address all the cases, but the section is not necessarily unique. However we will see that actually, in a given physical environment (that is for a given value of the fields), the motion of particles which have the same physical characteristics can be represented by the same section of P_G. This is the foundation for the idea of "matter field". Of course this representation holds for a collection of particles which have trajectories belonging to the same vector field.

3.6 SOME ISSUES ABOUT RELATIVITY

It is useful to review here some issues which arise frequently about Relativity.

3.6.1 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 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 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.

3.6.2 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 respect to another observer located at the same point, travel at the speed of light. Afterwards its 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.

3.6.3 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.

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 : $\ell_{AB} = c(\tau_B - \tau_A)$. 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 for the total duration of their travel to 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 : $OA : p_A(\tau_A) = c\tau_A\varepsilon_0$

The twin B moves in the direction of the first axis. His coordinates are then : $OB : p_B(\tau_B) = c\tau_B\varepsilon_0 + x_B(\tau_B)\varepsilon_1$

The spatial speed of B with respect to A is : $\frac{dOB}{d\tau_A} = V(\tau_B)\varepsilon_1$

The velocity of B is : $u_B = \frac{dOB}{d\tau_B} = \frac{1}{\sqrt{1-\frac{V^2}{c^2}}}(V\varepsilon_1 + c\varepsilon_0)$

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 $\gamma = \frac{1}{\sqrt{1-\frac{V^2}{c^2}}}$

$p_B(\tau_B) = \int_0^{\tau_B} \frac{c}{\sqrt{1-(\gamma t)^2}}(\gamma t\varepsilon_1 + \varepsilon_0)\, dt = \frac{c}{\gamma}\left[\sqrt{1-y^2}\varepsilon_1 + \varepsilon_0 \arcsin y\right]_0^{\gamma\tau_B}$

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.

3.6.4 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 even if 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.

3.6.5 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 of the universe, and indeed the existence of a universal time, which provides a foliation in hypersurfaces analogous to $\Omega_3(t)$, is one of their key component.

3.6.6 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 : 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 topics 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. 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 foliation of M in hypersurfaces similar to $\Omega_3(t)$, which represent the present for the observers who are located on them (see Chapter 4 and 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. And of course material points do not swell, only the vacuum, which separates material bodies, dilates.

3.6.7 For a full understanding of motion

We have built a comprehensive and consistent Geometry of General Relativity starting from the way one proceeds to measures, some general principles of Physics, and the concepts of space, time, material bodies and their motion, with their characteristic properties. We have not started from scratch, but from the usual, well known and proven formalism of Galilean Geometry. Relativity extends the framework, it does not negate it. And it leads to uncover some troubling facts which were actually already present in Galilean Geometry.

Exploring the concept of motion, we have seen that the idea of an orthonormal frame is actually present in our perception and understanding of the motion of a material body. We are so well used to deal with rotation that we forget two significant features : it is a property of material bodies, and it adds 3 parameters to characterize, geometrically, a material body, even in Galilean Geometry. Observers use a tetrad, but actually a tetrad is attached to any material body, and it must be seen as a property of matter, whatever the scale. The tetrad is orthonormal, and thus defined with respect to the metric, which is of physical nature. As well as particles travel with constant velocity, the tetrad attached to a material body must adjust (in a chart), to adapt to a changing metric. This is where the use of the tetrad formalism finds all its worth, compared to the usual computations with banal charts : it has a physical meaning, and is closer to the way measures are done.

The right way to deal with a metric is by principal bundles. But the representation of the concept of motion leads to see the Clifford bundle as the natural, and physical, framework to represent any change in the geometric state of a material body, be it its location or its arrangement. The Clifford bundle replaces the tangent bundle TM as the true physical domain where any change in the geometric characteristics of material bodies occurs.

Moreover the motion is essentially characterized by two vectors $r, w \in \mathbb{R}^3$ which have a clear physical meaning, and related to the 6 parameters used in Galilean Geometry. With all the tools of Clifford Algebra, it is then easy to work on and compute all the geometric problems in RG, even problems involving rotation which would have been intractable in the usual framework. Actually in the most part of the computations one can forget the chart, and the $\partial\xi_\alpha, d\xi^\beta$ which have been the nightmare of Physicists.

Chapter 4

KINEMATICS

Fields acts on particles by forces which change the motion of particles, according to kinematic characteristics of these particles. They are expressed as mass and inertial tensors, from which are defined translational and rotational momenta. Newtonian Mechanics has developed a comprehensive and sophisticated theory of Kinematics, and Analytic Mechanics has provided much of the initial framework for QM. Relativity introduces a totally new concept of motion, which is now absolute in a quadridimensional universe, and the usual concept of rigid solid does not hold any longer. If the usual concepts of Kinematics can more or less be fitted to Special Relativity, General Relativity requires a totally new approach, with spinors, which have been introduced, by a very different way, in the Quantum Theory of Fields.

As we have done for the Geometric concepts, it is useful to rediscover the main concepts of Kinematics in Newtonian Mechanics.

4.1 USUAL REPRESENTATIONS IN KINEMATICS

4.1.1 In Newtonian Mechanics

Motion and momentum are two different, but related, physical quantities. They are measured by different protocols. Momenta can be computed but actually this is the change in the value of the momenta which is measured, through inertial forces which express the resistance of a material body to change its motion.

As for motion, there is a translational momentum and a rotational momentum, to which are associated linear forces (or "forces") and torques.

The balance of energy exchanged by a material body with the forces exercised on it is then expressed by the kinetic energy, and there are a translational and a rotational kinetic energy.

The picture is clear for rigid solids, but can be extended to deformable solids, which are of a greater interest because they can be defined in the relativist context.

Translational Momentum

To a material point with mass m and speed $\overrightarrow{v} = \frac{dq}{dt}$ is associated the translational momentum $\overrightarrow{p} = m\overrightarrow{v}$. And the Fundamental Law of Mechanics states the relation $\overrightarrow{F} = \frac{d\overrightarrow{p}}{dt}$ between a force exercised on the material point and the change of its momentum. The assumption that m is a scalar constant leads then to a direct relation between the force and the motion. So a *change* of motion can be measured (by accelerometers as in smartphones) without any measure of the motion, even by an observer attached to the material body. And if $\overrightarrow{F} = 0$ then the momentum is constant.

For a system of material points the picture is more complicated, because actually the forces are localized quantities : they should be represented, not by a single vector $\overrightarrow{F}$, but by a couple $\left(q, \overrightarrow{F}\right)$. However Galilean Geometry has the special feature that one can define a center of mass G for any system of material points : $\left(\sum_a m_a\right)\overrightarrow{OG} = \sum_a m_a \overrightarrow{OM}_a$. Then the system is equivalent to a particle of mass $\sum_a m_a$ located at G and the sum $\overrightarrow{F}_G = \sum_a \overrightarrow{F}_a$, exercised at G, has a physical meaning. And the Law of Mechanics can be written :

$$\sum_a \frac{d\overrightarrow{p}_a}{dt} = \frac{d\overrightarrow{p}_G}{dt} = \overrightarrow{F}_G$$

Torque

Another consequence of the localization of the forces is the existence of torques, similar to forces, but which are distinct physical quantities.

For a force $\left(M_a, \overrightarrow{F_a}\right)$ the torque is defined with respect to any fixed point O by $\tau_a(O) = \overrightarrow{OM_a} \times \overrightarrow{F_a}$ with the cross product. $\tau_a(O)$ reads : $\tau_a(O) = j\left(\overrightarrow{OM_a}\right)\overrightarrow{F_a} = j\left(\overrightarrow{F_a}\right)\overrightarrow{M_aO}$ so this is actually an operator, acting on O, with an antisymmetric matrix, which can then be represented by a vector of $\mathbb{R}^3$ with the usual convention. As with the translational momentum, the rotational momentum is then defined by :

$$\Gamma_a(O) = j\left(\overrightarrow{OM_a}\right)\overrightarrow{p_a} = m_a j\left(\overrightarrow{OM_a}\right)\left(\frac{d}{dt}\overrightarrow{OM}_a\right)$$

$$\frac{d}{dt}\Gamma_a(O) = j\left(\frac{d}{dt}\overrightarrow{OM_a}\right)\overrightarrow{p_a} + j\left(\overrightarrow{OM_a}\right)\frac{d}{dt}\overrightarrow{p_a} = \frac{d}{dt}\tau_a(O)$$

Because in Galilean Geometry one can define a center of mass :

$$\Gamma_a(O) = j\left(\overrightarrow{OG}\right)\overrightarrow{p_a} + j\left(\overrightarrow{GM_a}\right)\overrightarrow{p_a}$$

$$\sum_a \Gamma_a(O) = j\left(\overrightarrow{OG}\right)\sum_a \overrightarrow{p_a} + \sum_a j\left(\overrightarrow{GM_a}\right)\overrightarrow{p_a} = \sum_a j\left(\overrightarrow{GM_a}\right)\overrightarrow{p_a}$$

$$= \sum_a j\left(\overrightarrow{GM_a}\right)m_a\left(\frac{d}{dt}\overrightarrow{OG} + \frac{d}{dt}\overrightarrow{GM}_a\right) = \sum_a j\left(\overrightarrow{GM_a}\right)m_a\frac{d}{dt}\overrightarrow{GM}_a = \sum_a \Gamma_a(G)$$

and one can define a total torque :

$\tau = \sum_a \tau_a(O) = \sum_a \tau_a(G)$

For a rigid solid :

$\overrightarrow{G(t)\, M_a(t)} = R(t)\, \overrightarrow{X}_a$ with $\overrightarrow{X}_a = Ct$

$\sum_a \Gamma_a(G) = \sum_a m_a j\left(\overrightarrow{GM_a}\right) \frac{d}{dt}\overrightarrow{GM_a} = \sum_a m_a j\left(R(t)\, \overrightarrow{X}_a\right) \frac{dR}{dt}\overrightarrow{X}_a$

$= \sum_a m_a j\left(R(t)\, \overrightarrow{X}_a\right) R(t)\, R(t)^{-1} \frac{dR}{dt}\overrightarrow{X}_a$

$= R(t) \sum_a m_a j\left(\overrightarrow{X}_a\right) R(t)^{-1} \frac{dR}{dt}\overrightarrow{X}_a$

$= -R(t) \sum_a m_a j\left(\overrightarrow{X}_a\right) j\left(\overrightarrow{X}_a\right) r(t)$

$[J] = -\sum_a m_a j\left(\overrightarrow{X}_a\right) j\left(\overrightarrow{X}_a\right)$ is a fixed symmetric matrix, the inertial tensor, and $[J]\, r(t)$ is the rotational momentum.[1]

$\sum_a \Gamma_a(G) = R(t)\, [J]\, r(t)$

and $\sum_a \tau_a(G) = \frac{dR}{dt}[J]\, r(t) + R[J]\frac{dr}{dt} = R\left(j(r)\, [J]\, r + [J]\frac{dr}{dt}\right)$

Kinetic energy

Mechanical Energy is defined as the work done by a force along a path : $W = \int_{q_1}^{q_2} \left\langle \overrightarrow{F}, \overrightarrow{dq} \right\rangle$ thus with $\overrightarrow{F} = \frac{d\overrightarrow{p}}{dt}$:

$W = \int_{t_1}^{t_2} \frac{1}{m}\left\langle \overrightarrow{p}, \frac{d\overrightarrow{p}}{dt}\right\rangle dt = \frac{1}{2}\int_{t_1}^{t_2} \frac{1}{m}\frac{d}{dt}\left\langle \overrightarrow{p}, \overrightarrow{p}\right\rangle dt$ which leads to the definition of the variation of kinetic energy : $\delta K = \frac{1}{m}\left\langle \overrightarrow{p}, \overrightarrow{\delta p}\right\rangle$, that is the energy that the body exchanges with the exterior in a change $\overrightarrow{\delta p}$ of momentum, and the kinetic energy $K = \frac{1}{2m}\left\langle \overrightarrow{p}, \overrightarrow{p}\right\rangle$ when, in a continuous motion, $\overrightarrow{\delta p} = \frac{d\overrightarrow{p}}{dt}$. It is defined with respect to an observer, as well as $\overrightarrow{p}$.

Kinetic energy being a scalar, one can sum the kinetic energy related to the translational momentum of a set of material points :

$K = \sum_a \frac{1}{2m_a}\left\langle \overrightarrow{p_a}, \overrightarrow{p_a}\right\rangle$

$= \sum_a \frac{1}{2} m_a \left\langle \frac{d}{dt}\overrightarrow{OG} + \frac{d}{dt}\overrightarrow{GM}_a, \frac{d}{dt}\overrightarrow{OG} + \frac{d}{dt}\overrightarrow{GM}\right\rangle$

$= \frac{1}{2} M \left\|\overrightarrow{v_G}\right\|^2 + \sum_a \frac{1}{2} m_a \left\langle \frac{d}{dt}\overrightarrow{GM}_a, \frac{d}{dt}\overrightarrow{GM_a}\right\rangle$

For a solid :

$\overrightarrow{G(t)\, M_a(t)} = R(t)\, \overrightarrow{X}_a$

$\left\langle \frac{d}{dt}\overrightarrow{GM}_a, \frac{d}{dt}\overrightarrow{GM_a}\right\rangle = \left\langle \frac{dR}{dt}\overrightarrow{X}_a, \frac{dR}{dt}\overrightarrow{X_a}\right\rangle = \left\langle R^{-1}\frac{dR}{dt}\overrightarrow{X}_a, R^{-1}\frac{dR}{dt}\overrightarrow{X_a}\right\rangle = \left\langle j(r)\, \overrightarrow{X}_a, j(r)\, \overrightarrow{X_a}\right\rangle$

$= \left\langle j\left(\overrightarrow{X}_a\right) r, j\left(\overrightarrow{X}_a\right) r\right\rangle = -[r]^t \left[j\left(\overrightarrow{X}_a\right)\right]\left[j\left(\overrightarrow{X}_a\right)\right][r]$

$\sum_a \frac{1}{2} m_a \left\langle \frac{d}{dt}\overrightarrow{GM}_a, \frac{d}{dt}\overrightarrow{GM_a}\right\rangle = \frac{1}{2}[r]^t [J]\, [r]$

$K = \frac{1}{2} M \left\|\overrightarrow{v_G}\right\|^2 + \frac{1}{2}[r]^t [J]\, [r]$

And the variation of rotational kinetic energy is :

$\frac{d}{dt}\left(\frac{1}{2}[r]^t [J]\, [r]\right) = \frac{1}{2}\left[\frac{dr}{dt}\right]^t [J]\, [r] + \frac{1}{2}[r]^t [J]\left[\frac{dr}{dt}\right]$

The torque on the solid :

$\tau(G) = \sum_a \tau_a(G) = \frac{d}{dt}\left(R(t)\, [J]\, r(t)\right) = R\left(j(r)\, [J]\, r + [J]\frac{dr}{dt}\right)$

$[J]\frac{dr}{dt} = R^t [\tau(G)] - j(r)\, [J]\, r$

$\frac{1}{2}\left[\frac{dr}{dt}\right]^t [J]\, [r] = \frac{1}{2}\left([\tau(G)]^t R + [r]^t [J]\, j(r)\right)[r]$

$\frac{1}{2}[r]^t [J]\left[\frac{dr}{dt}\right] = \frac{1}{2}[r]^t \left([R]^t [\tau(G)] - j(r)\, [J]\, [r]\right)$

[1] Matrices like $j(X)\, j(X)$ have negative eigen values, so the minus sign induces positive momenta along the eigen vectors.

$\frac{d}{dt}\left(\frac{1}{2}[r]^t[J][r]\right)=\frac{1}{2}[\tau(G)]^t[R][r]+\frac{1}{2}[r]^t[J]j(r)[r]+\frac{1}{2}[r]^t[R]^t[\tau(G)]-\frac{1}{2}[r]^tj(r)[J][r]$

$=[r]^t[R]^t[\tau(G)]$ is the work done by the torque $\tau(G)$

$\delta K=\frac{1}{m}\left\langle\overrightarrow{p},\overrightarrow{\delta p_G}\right\rangle+[r]^t[R]^t[\delta\Gamma(G)]$

The representation and computations above rely heavily on the existence of a center of mass, the fact that $SO(3)$ has the same dimension as the space $\mathbb{R}^3$, and on the properties of solids. However the definitions can be extended to deformable solids.

Units

The formula : $K=\frac{1}{2}M\|\overrightarrow{v_G}\|^2+\frac{1}{2}[r]^t[J][r]$ is consistent if the rotational motion is measured by $[r]$ with the same units as $\overrightarrow{v_G}$,, as noticed before, then $[J]$ has the dimension of a mass. The rotational momentum $[J]r(t)$ is measured in the same units as the translational momentum $M\overrightarrow{v_G}$. However the torque $\tau_a(O)=\overrightarrow{OM_a}\times\overrightarrow{F_a}$ is not measured with the same units as the force $\overrightarrow{F_a}$.

Density

Material bodies are comprised of material points, so it is natural to introduce a density μ, seen as the number of identical material points at the same location x at the time t : $\mu(x,t)$. The density defines, with a volume form $\varpi_3=\varepsilon_1\wedge\varepsilon_2\wedge\varepsilon_3$ a measure $\mu\varpi_3$ such that the mass of the material body in an area Ω at t is $M(t)=\int_\Omega\mu(x,t)\varpi_3(x)$.

In the model of deformable solid introduced previously a material point is labeled by its position q at $t=0$ and its position at t is given by a differentiable map : $X(q,t)=\phi(q,t)$. A basis, e_i attached at q, orthonormal at $t=0$ is transported by $\phi'_q(q,t)$: $e_i(q,t)=\phi'_q(q,t)e_i(q,0)$. It is no longer orthonormal at t and defines a metric $g_{ij}(q,t)=\langle e_i(q,t),e_j(q,t)\rangle$ and a volume form $\varpi(q,t)=\sqrt{\det g}\varepsilon_1\wedge\varepsilon_2\wedge\varepsilon_3=\det\phi'_q(q,t)\varepsilon_1\wedge\varepsilon_2\wedge\varepsilon_3$.

$\varpi(q,t)$ is just the push forward of ϖ_3 by ϕ. The material points which occupy a volume $\varepsilon_1\wedge\varepsilon_2\wedge\varepsilon_3$ at $t=0$ occupy a volume $\det\phi'_q(q,t)\varepsilon_1\wedge\varepsilon_2\wedge\varepsilon_3$ at t. Then the conservation of mass, which is equivalent to the conservation of the number of particles, leads to :

$\frac{\partial}{\partial t}\left(\mu(q,t)\det\phi'_q(q,t)\right)=0$

$\frac{\partial\mu}{\partial t}\det\phi'_q+\mu\left(\det\phi'_q\right)Tr\left(\left[\frac{\partial}{\partial t}\phi'_q\right]\left[\phi'_q\right]^{-1}\right)=0$

The trajectories of the particles are : $\frac{\partial}{\partial t}X(q,t)=\frac{\partial}{\partial t}\phi(q,t)$

Let us define : $V:\mathbb{R}^3\times\mathbb{R}\to\mathbb{R}^3::\overrightarrow{V}(x,t)$ such that : $\overrightarrow{V}(X(q,t),t)=\frac{\partial}{\partial t}X(q,t)$. It is called the “flow velocity".

$\Rightarrow\frac{\partial^2\phi_i}{\partial t\partial q_k}(q,t)=\frac{\partial V_i}{\partial q_k}=\sum_{j=1}^3\frac{\partial V_i}{\partial x_j}\frac{\partial x_j}{\partial q_k}=\sum_{j=1}^3\frac{\partial V_i}{\partial x_j}\frac{\partial\phi_j}{\partial q_k}$

$\Rightarrow\left[\frac{\partial}{\partial t}\phi'_q\right]=\left[\frac{\partial V}{\partial x}\right]\left[\frac{\partial\phi}{\partial q}\right]$

$Tr\left(\left[\frac{\partial}{\partial t}\phi'_q\right]\left[\phi'_q\right]^{-1}\right)=Tr\left[\frac{\partial V}{\partial x}\right]=div\overrightarrow{V}$

and we get the continuity equation : $\frac{\partial\mu}{\partial t}+\mu div\overrightarrow{V}=0$.

The reasoning is done usually for fluids but it holds for a deformable solid.

Stress tensor, Energy-momentum tensor

The motion of each material point of a deformable solid can be represented by :

- a translation, given by $\frac{dq}{dt}=\frac{\partial}{\partial t}\phi(q,t)=\overrightarrow{V}(X(q,t),t)$

- a deformation of its orthonormal basis $(e_i(q,t))_{i=1}^3$ given by : $\frac{\partial}{\partial t}e_i(q,t)=[\gamma(q,t)]e_i(q,t)$ with the deformation tensor $[\gamma]=\left[\phi"_{qt}(q,t)\right]\left[\phi'_q(q,t)\right]^{-1}=[\partial_xV]$ which can be decomposed in a symmetric part $[s]=\frac{1}{2}\left([\gamma]+[\gamma]^t\right)$ and an antisymmetric part $[j(\rho)]=\frac{1}{2}\left([\gamma]-[\gamma]^t\right)$

The usual momentum of the material point is : $\overrightarrow{p}(q,t)=\mu(q,t)\frac{dq}{dt}=\mu(q,t)\overrightarrow{V}(X(q,t),t)$.

The deformation of the solid is the effect of forces, or conversely the solid opposes forces to its deformation. From :

$\frac{d}{dt}\left(\mu\left(q,t\right)V\left(X\left(q,t\right),t\right)\right)=\frac{d\mu}{dt}V+\mu\left(\left[\partial_{x}V\right]\left[\phi_{t}'\right]+\left[\partial_{t}V\right]\right)=\frac{d\mu}{dt}V+\mu\left(\left(\left[s\right]+\left[j\left(\rho\right)\right]\right)V+\left[\partial_{t}V\right]\right)$

one can identify :

- a force corresponding to a variation of the translational momentum : $\mu\left[\partial_{t}\overrightarrow{V}\right]$

- the forces, similar to a pressure (they act symmetrically), opposed to the variation of the volume :

$\frac{d\mu}{dt}V+\mu\left[s\right]V=\mu\left(\left[s\right]-divV\right)V$

- a torque $\mu\left[j\left(\rho\right)\right]V$

The variation of the kinetic energy can be computed as above.

$\frac{dK}{dt}=\frac{1}{2}\frac{1}{\mu}\left\langle \overrightarrow{p}\left(q,t\right),\frac{d}{dt}\overrightarrow{p}\left(q,t\right)\right\rangle =\frac{1}{2}\left[V\right]^{t}\left\{ \mu\left[\partial_{t}V\right]+\mu\left(\left[s\right]-divV\right)V+\mu\left[j\left(\rho\right)\right]V\right\}$

$=\frac{1}{2}\mu\left[V\right]^{t}\left[\partial_{t}V\right]+\frac{1}{2}\mu\left[V\right]^{t}\left[s\right]\left[V\right]-\frac{1}{2}\mu\left(divV\right)\left[V\right]^{t}\left[V\right]+\mu\frac{1}{2}\left[V\right]^{t}\left[j\left(\rho\right)\right]V$

and we have a kinetic energy corresponding to the rotational momentum $\mu\frac{1}{2}\left[j\left(V\right)\right]\rho$.

This is usually written with a "stress tensor" $T=T_{j}^{i}\varepsilon^{j}\otimes\varepsilon_{i}$ such that the forces, on the surface $d\sigma$ with normal $\overrightarrow{n}$, opposing the deformation (the "stress"), are $\overrightarrow{dF}=T\left(\overrightarrow{n}\right)=\sum_{i,j=1}^{3}\left[T\right]_{j}^{i}\left[n\right]^{j}\varepsilon_{i}d\sigma$.

By considering a small volume Ω with border $\partial\Omega$:

- the sum of the stress on Ω is :

$\int_{\partial\Omega}\sum_{i,j=1}^{3}\left[T\right]_{j}^{i}\left[n\right]^{j}\varepsilon_{i}d\sigma=\int_{\partial\Omega}\left\langle \sum_{i=1}^{3}\left[T\right]^{i}\varepsilon_{i},\overrightarrow{n}\right\rangle d\sigma=\int_{\Omega}\overrightarrow{dF_{v}}\varpi_{3}$ that is a force by unit of volume $\overrightarrow{dF}_{v}=\sum_{j=1}^{3}div\left(\sum_{i=1}^{3}\left[T\right]_{j}^{i}\varepsilon_{i}\right)\varepsilon_{j}$

- the torque with respect to the origin :

$\tau\left(O\right)=\int_{\partial\Omega}\overrightarrow{X}\times\sum_{i,j=1}^{3}\left[T\right]_{j}^{i}\left[n\right]^{j}\varepsilon_{i}d\sigma=\int_{\partial\Omega}\left\langle j\left(X\right)\sum_{i=1}^{3}\left[T\right]^{i}\varepsilon_{i},\overrightarrow{n}\right\rangle d\sigma=\int_{\Omega}d\tau\varpi_{3}$

with the fixed orthonormal basis ε_{j}

$d\tau=\sum_{i,j=1}^{3}div\left(j\left(X\right)\sum_{i=1}^{3}\left[T\right]_{j}^{i}\varepsilon_{i}\right)\varepsilon_{j}=\sum_{i,j=1}^{3}\frac{\partial}{\partial x_{i}}\left(\left[j\left(X\right)\left[T\right]_{j}\right]\right)^{i}\varepsilon_{j}=\sum_{i,j=1}^{3}j\left(\varepsilon_{i}\right)\left[T\right]_{j}\varepsilon_{j}+j\left(X\right)\sum_{i,j=1}^{3}\left[\left[\frac{\partial}{\partial x_{i}}T\right]_{j}^{i}\right]\varepsilon_{j}$

$=\sum_{i,j=1}^{3}j\left(\varepsilon_{i}\right)\left[T\right]_{j}^{i}\varepsilon_{j}+\overrightarrow{X}\times\overrightarrow{dF}_{v}$

Thus there is an elementary torque located at X equal to $\sum_{i,j=1}^{3}j\left(\varepsilon_{i}\right)\left[T\right]_{j}^{i}\varepsilon_{j}$

Symmetries

In Kinematics symmetries of a solid can be understood as symmetries of the density $\mu:\mathbb{R}^{3}\rightarrow\mathbb{R}$, and as symmetries in the momentum. For instance the instantaneous rotation of a solid, represented by $\left[R\right]^{-1}\left[\frac{dR}{dt}\right]=j\left(r\right)$ with a matrix $\left[R\right]\in SO\left(3\right)$, does not change in a change of observer given by a global rotation $g\in SO\left(3\right)$:

$R\rightarrow\widetilde{R}=gR,$

$R^{-1}\frac{dR}{dt}=j\left(r\right)\rightarrow\widetilde{R}^{-1}\frac{d\widetilde{R}}{dt}=j\left(r\right)$

but the rotational momentum : $\sum_{a}\Gamma_{a}\left(G\right)=R\left(t\right)\left[J\right]r\left(t\right)$ changes,as :

$\sum_{a}\Gamma_{a}\left(G\right)\rightarrow\widetilde{R\left(t\right)}\left[J\right]r\left(t\right)=g\sum_{a}\Gamma_{a}\left(G\right)$

So that all rotational motions are not equivalent, as all engineers know. There is a symmetry if :

$gR\left(t\right)\left[J\right]r\left(t\right)=R\left(t\right)\left[J\right]r\left(t\right)$

that is if $R\left(t\right)\left[J\right]r\left(t\right)$ is an eigen vector of g. The only real eigen vector of g is given by the axis with the eigen value 1. The matrix $\left[J\right]$ is symmetric, and has 3 orthogonal eigen vectors r_{a}, with real eigen values λ_{a}. If the motion is a constant rotation with axis one of this eigen vectors r_{a}, then $R\left(t\right)\left[J\right]r\left(t\right)=\lambda_{a}\left(\exp tr_{a}\right)r_{a}=\lambda_{a}r_{a}$ and there is a symmetry for $g=\exp r_{a}$.

Energy momentum tensor

A more general way to deal with these issues is with the "Energy-Momentum" tensor, which comes from the Principle of Least Action. A system represented by variables $z^i(m), i = 1...N$ defined over a manifold with coordinates $(\xi_\alpha)_{\alpha=1}^3$, and their first partial derivatives $z^i_\alpha(m)$ is endowed with a scalar lagrangian such that the equilibrium is reached when the functional $\int_\Omega L\left(z^i(m), z^i_\alpha(m)\right)\varpi(m)$ is stationary. Then the quantity :

$T = \sum_{i\alpha\beta} \frac{\partial L}{\partial z^i_\alpha} z^\beta_i \partial\xi_\alpha \otimes d\xi^\beta$

is a tensor, called the Energy-Momentum tensor. The Lagrangian has the meaning of the energy of the system, and a change $\delta z^i = \sum_\beta z^i_\beta \delta v^\beta$ of the variables $z^i(m)$ along $\delta v = \sum_{\beta=0}^3 v^\beta \delta\xi_\beta$ changes the energy by $\delta\ell = \int_\Omega div\left(T(\delta v)\right)\varpi$ (this is seen in more details in the Chapter 6) so that $T(\delta v)$ can be seen as a reaction of the system to a change by δv, that is as a force. Then the quantities $\Pi_i = \sum_{\alpha\beta} \frac{\partial L}{\partial z^i_\alpha} z^\beta_i \partial\xi_\alpha \otimes d\xi^\beta$ are the momenta associated to the scalar variable z^i. They are the generalized definition of the translational and rotational momenta, as they apply to any motion. If $z = \left(z^i\right)_{i=1}^n$ are the components of a vector in some vector space E then there is a momentum Π_z expressed as a tensor valued in the dual space E^*. For a system with Lagrangian $L\left(z^i(t), \dot{z}^i\right)$ and variables depending on t only, the infinitesimal variations are $\delta \dot{z}^i = \frac{dz^i}{dt}\delta t$ and $\Pi_i = \frac{\partial L}{\partial \dot{z}^i}\frac{dz^i}{dt}$.

To sum up, in Newtonian Mechanics :

i) The kinematic of a material body is represented by a translational momentum and a rotational momentum, which are distinct and read : $\overrightarrow{p} = m\overrightarrow{v}; \Gamma = R(t)[J]r$

ii) Each momentum is related to the motion, and overall the kinematic characteristics of a solid are represented by 7 independent scalars (the mass and 6 parameters for $[J]$).

iii) The momenta can be computed, but this is the change in the momenta which is measured, through the inertial forces.

iv) The representation of the momenta by vectors of $\mathbb{R}^3$ is conventional. If it is natural for $\overrightarrow{p}$, the vector $R(t)[J]r$ has no direct relation with a physical basis $\overrightarrow{\varepsilon}_i$.

v) For deformable solids and systems the definition of momenta is less straightforward and comes from the identification of the forces, inertial and external, acting on the body. The representation of momenta and forces is given through the lagrangian.

vi) The conservation of the momentum in the transformation of a system is only a special case of the laws of the transformation, meanwhile the conservation of energy is just the balance of the energy exchanged by its different components.

4.1.2 Usual representations in the relativist framework

Translational momentum

The translational momentum is defined as the 4 dimensional vector : $P = mV$. It depends on the observer, and here it means the choice of the time t in the derivative $V = \frac{dq}{dt}$.

In the relativist context location and motion are absolute. So there is an intrinsic definition of the momentum, *for an observer who is attached to the particle* with the proper time and velocity $u = \frac{dq}{d\tau}$: $p = mu$. Then, if we take the same definition for the kinetic energy, with respect to this observer, it is constant : $K = \frac{1}{m}\langle p, p\rangle = -mc^2$.

For any other observer :

$P = mV = p\frac{d\tau}{dt} = p\sqrt{1 - \frac{\|\overrightarrow{v}\|^2}{c^2}}$

$-\frac{1}{m}\langle P, P\rangle = -\left(1 - \frac{\|\overrightarrow{v}\|^2}{c^2}\right)mc^2 = \left(1 - \frac{\|\overrightarrow{v}\|^2}{c^2}\right)K \leq K$

$p = m\frac{1}{\sqrt{1-\frac{\|\overrightarrow{v}\|^2}{c^2}}}\left(c\varepsilon_0\left(q\left(t\right)\right)+\overrightarrow{v}\right)$ is a 4 dimensional vector. However the common practice is to distinguish its spatial and time components. The spatial part : $\overrightarrow{p_r} = m\frac{\overrightarrow{v}}{\sqrt{1-\frac{\|\overrightarrow{v}\|^2}{c^2}}}$ which is similar to the usual translational momentum, and $m\frac{c}{\sqrt{1-\frac{\|\overrightarrow{v}\|^2}{c^2}}}$ is then related to the energy E, defined by :

$E = c^2 m\frac{1}{\sqrt{1-\frac{\|\overrightarrow{v}\|^2}{c^2}}} = \langle Pc, \varepsilon_0\rangle = mc\left\langle \frac{dq}{d\tau}, \varepsilon_0\right\rangle$

$\Rightarrow E^2 = c^2\left\|\overrightarrow{p}_r\right\|^2 + m^2c^4$ which is just $\langle pc, pc\rangle = -m^2c^4 = c^2\left\|\overrightarrow{p}_r\right\|^2 - E^2$

The advantage of this expression is that for small speed it gives :

$E = c^2 m\frac{1}{\sqrt{1-\frac{\|\overrightarrow{v}\|^2}{c^2}}} \simeq c^2 m\left(1+\frac{1}{2}\frac{\|\overrightarrow{v}\|^2}{c^2}\right) = \frac{1}{2}m\left\|\overrightarrow{v}\right\|^2 + mc^2$

The total energy of the particle E has one part corresponding to a kinetic energy and another one to an "energy at rest". So keep the principle of conservation of energy leads to accept that mass itself can be transformed into energy, according to the famous relation $E = mc^2$.

However it mixes two concepts - momentum and energy - which are usually seen as distinct and are measured by different protocols.

The definition of rigid solid of Newtonian Mechanics does not extend to the relativist Geometry, and there is no satisfying definition for the rotational momentum.

In GR what is considered is the energy-momentum tensor T, which is a key part of the Einstein equation. There is no general formula to specify T, only phenomenological laws. The most usual are based on the behavior of dust clouds, including sometimes thermodynamic components.

The Dirac's equation

In writing $pc = (c\overrightarrow{p_r}, E)$ the energy E and p_r are two separate quantities which can be measured [2]. In the usual interpretation of QM to E and p_r are associated operators acting on scalar wave functions ψ.

In common QM, "quantization" is just an operation where mathematical symbols are substituted to other symbols. Starting from : $E^2 = c^2\left\|\overrightarrow{p}_r\right\|^2 + m^2c^4$ the "minimal substitution rule" : $E \rightarrow i\hbar\frac{\partial}{\partial t}; p_{r\alpha} \rightarrow -i\hbar\partial_\alpha$ gives the Klein-Gordon equation : $\left(\Box + m^2\right)\psi = 0$ which, checked for the spectrum of Hydrogen, provides wrong results.

In order to have first order derivatives Dirac proposed another equation, starting from $E = \sqrt{c^2\left\|\overrightarrow{p}_r\right\|^2 + m^2c^4}$ assuming that :

$E = A.p_r + Bm$ the substitution gives : $i\hbar\frac{\partial\psi}{\partial t} = \left(Ai\hbar\nabla + B\mu\right)\psi$

But 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 $\gamma = \left(\gamma_j\right)_{j=0..3}$ with the relation : $\gamma_i\gamma_j + \gamma_j\gamma_i = 2\eta_{ij}I_4$. The wave functions $\psi\left(t, \xi_1, \xi_2, \xi_3\right)$ are then vectors, called spinors, belonging to a 4 dimensional complex vector space F, and (F, γ) is the representation of the Lorentz group with action given by the matrices γ.

The Dirac's equation then reads :

$i\frac{\partial\psi}{\partial t} = -i\sum_{\alpha=1}^{3}\gamma_\alpha\frac{\partial\psi}{\partial\xi_\alpha} + m\gamma_0\psi$

and can be seen as a propagation equation for ψ or, in the usual QM, as a substitute for the Schrödinger equation. Its eigen values correspond to the energy. Its eigen vectors, which provide a basis for observables quantities, correspond to "plane waves" :

[2] Actually only the change of momentum and energy can be physically measured.

with positive energies : $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \exp(-imt), \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \exp(-imt)$

with negative energies : $\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \exp(imt), \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \exp(imt)$

The existence of the last two solutions leads to antiparticles. The proof of their existence has not closed the issue of the interpretation of these solutions, the most common being that antiparticles are "holes" in a sea of virtual particles, and that they moved backwards in time.

$\psi(t,x)$ is such that $\rho = \psi(t,x)^* \psi(t,x)$ gives the probability to find the particle at (t,x). Then the Dirac's currents $j^a = \overline{\psi}^t \gamma_a \psi$ gives the probability to find the particle in $\xi^a, a = 1,2,3$ and the solutions of the Dirac's equation meet the continuity equation :

$\frac{\partial \rho}{\partial t} + \sum_{\alpha=1}^{3} \frac{\partial j_\alpha}{\partial \xi_\alpha} = 0$

The scheme has been extended to account for the action of the fields, and leads to the standard model. But its construction is totally abstract, and justified only by the results that it provides, through complicated computations.

So in the Relativist context we have two representations which proceed form totally different principles. And this is at the core of the belief that QM and GR are not reconciliable.

4.2 MOMENTA REVISITED

Our purpose is to find an efficient way to represent the kinematic characteristics of particles and material bodies in the framework of General Relativity. We focus on the properties assigned to momenta of material bodies :

- momenta are physical quantities, related to the motion but distinct, and a change in the value of the momenta can be physically measured through the inertial forces, by specific protocols;

- they are computed from their properties. A particle is defined not only by its location and transversal motion, but also by an orthonormal basis attached to it, with its rotational motion. So we must consider translational and rotational momenta.

- they must be expressed in a format which is equivariant with respect to the group $Spin(3,1)$.

- momenta are localized quantities : a momentum is defined at each location $q(t)$ of the particle.

- momenta are expressed by vectorial quantities : the linear combination of momenta at the same point has a physical meaning (such as in a collision).

- in a continuous motion the momenta are related to the motion by some fixed relation.

- for a free particle, which is not submitted to any force, the momenta are constant along its world line.

We will naturally look for a fiber bundle representation. It should be a vector bundle associated to a principal bundle based on $Spin(3,1)$, and the natural choice is $P_G[E,\gamma]$, with some vector space E and action γ. E is the vector space in which are represented forces and torques. In Newtonian Mechanics they are represented by 2 distinct vectors in $\mathbb{R}^3$, but, at least for the torque, this is just a convention. So we are quite free in the choice of the vector space E. It is legitimate to look for two vectors in the Minkovski space, or a vector in a complex 4 dimensional space.

The motion $\frac{d\sigma}{dt}\cdot\sigma^{-1}$ is represented in the Lie Algebra. The derivative $\gamma'(1)$ provides a representation of the Lie algebra $T_1Spin(3,1)$ but with the bracket as internal operation, which has little interest here, so we look for a representation (E,γ) of the Clifford algebra itself. This is consistent with the assessment that the Clifford bundle $Cl(TM)$ is the right framework to represent arrangement and motion of material bodies.

The motion is represented by $(q,\sigma,\upsilon(X_r,X_w))$ in $J^1Cl(TM)$.

The momentum is represented in the associated vector bundle $P_G[E,\gamma]$

The motion comes, in a continuous motion, from the derivation of the arrangement σ. It is then natural to consider a quantity $S=\gamma(\sigma)S_0\in E$ representing the state of the particle, with a fixed vector S_0 representing the kinematic characteristics of the body, from which the momentum is computed by derivation.

4.2.1 Representation of the Clifford Algebra

Principles

A geometric representation (E,γ) of a Clifford algebra is an isomorphism $\gamma : Cl \to L(E;E) :: [\gamma(X)]$ where $[\gamma(X)]$ is the matrix of an endomorphism of E, represented in some basis. 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, the generators, $(\gamma_i)_{i=0}^3$, representing each vector $(\varepsilon_i)_{i=0}^3$ of an orthonormal basis. The choice of these matrices is not 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 $\widetilde{\gamma}_j=M\gamma_jM^{-1}$ with a fixed matrix M gives an equivalent representation. An element of the Clifford algebra is then represented by a linear combination of generators :

$\gamma(X)=\gamma\left(\sum_{\{i_1...i_r\}}X^{i_1...i_r}\varepsilon_{i_1}\cdot...\cdot\varepsilon_{i_r}\right)=\sum_{\{i_1...i_r\}}X^{i_1...i_r}\gamma_{i_1}...\gamma_{i_r}$

A Clifford algebra has, up to isomorphism, a unique faithful algebraic irreducible representation in an algebra of matrices. 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), acting on a 4 dimensional vector space;

for $Cl(1,3)$ this is $H(2)$ the 2×2 matrices with quaternionic elements.

So the choice of a representation raises the issue of the signature. However the vector space E upon which are represented the momenta can be a 4 dimensional complex vector space. The representation of complex Clifford algebras are on complex vector spaces. 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 ω, which exists in any complex algebra, but not in $Cl(1,3), Cl(3,1)$. As chirality is a defining feature of particles and of the rotational motion, this is an additional argument for using a complex Clifford algebra.

The first step is to expand $Cl(1,3), Cl(3,1)$ into $Cl(\mathbb{C},4)$.

Complexification of real Clifford algebras

We have seen how to introduce a complex structure on the Clifford algebra. There is another method, more usual, by extending the set such that the operations hold with complex numbers (Maths.6.5.2). One starts by he complexification of the vector space F : it 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 $(++++)^3$, 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 $Cl(F_{\mathbb{C}},\langle\rangle)$ which is the complexified of $Cl(F,\langle\rangle)$. In $Cl(F_{\mathbb{C}},\langle\rangle)$ the product of vectors is :

$\forall u,v\in F_{\mathbb{C}}: u\odot v+v\odot u=2\langle u,v\rangle_{\mathbb{C}}$

with the bilinear *symmetric* form $\langle u,v\rangle_{\mathbb{C}}$ of signature $(++++)$. $Cl(3,1)$ and $Cl(1,3)$ have the same complexified algebraic structure $Cl(\mathbb{C},4)$. Any orthonormal basis of $Cl(3,1)$ or $Cl(1,3)$ is an orthonormal basis of $Cl(\mathbb{C},4)$ and : $\varepsilon_i\odot\varepsilon_j+\varepsilon_j\odot\varepsilon_i=2\delta_{ij}$ and $\varepsilon_0\odot\varepsilon_0=+1$

$Cl(3,1)$ and $Cl(1,3)$ are real vector subspaces of $Cl(\mathbb{C},4)$.

There are real algebras morphisms (injective but not surjective) from the real Clifford algebras to $Cl(\mathbb{C},4)$.

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 :

$\widetilde{C}:(F,\langle\rangle)\to Cl(\mathbb{C},4)::\widetilde{C}(u)=(u+\langle\varepsilon_0,u\rangle_F\varepsilon_0)-i\langle\varepsilon_0,u\rangle_F\varepsilon_0=u+\langle\varepsilon_0,u\rangle_F(\varepsilon_0-i\varepsilon_0)$

(this is just the map : $\widetilde{C}(\varepsilon_j)=\varepsilon_j, j=1,2,3;\widetilde{C}(\varepsilon_0)=i\varepsilon_0$)

$\widetilde{C}(u)\odot\widetilde{C}(v)+\widetilde{C}(v)\odot\widetilde{C}(u)$

$=(u+\langle\varepsilon_0,u\rangle_F(\varepsilon_0-i\varepsilon_0))\odot(v+\langle\varepsilon_0,v\rangle_F(\varepsilon_0-i\varepsilon_0))$

$+(v+\langle\varepsilon_0,v\rangle_F(\varepsilon_0-i\varepsilon_0))\odot(u+\langle\varepsilon_0,u\rangle_F(\varepsilon_0-i\varepsilon_0))$

$=u\odot v+\langle\varepsilon_0,v\rangle_F u\odot(\varepsilon_0-i\varepsilon_0)+\langle\varepsilon_0,u\rangle_F(\varepsilon_0-i\varepsilon_0)\odot v$

$+\langle\varepsilon_0,u\rangle_F\langle\varepsilon_0,v\rangle_F(\varepsilon_0-i\varepsilon_0)\odot(\varepsilon_0-i\varepsilon_0)$

$+v\odot u+\langle\varepsilon_0,u\rangle_F v\odot(\varepsilon_0-i\varepsilon_0)+\langle\varepsilon_0,v\rangle_F(\varepsilon_0-i\varepsilon_0)\odot u$

$+\langle\varepsilon_0,v\rangle_F\langle\varepsilon_0,u\rangle_F(\varepsilon_0-i\varepsilon_0)\odot(\varepsilon_0-i\varepsilon_0)$

$=2\langle u,v\rangle_{\mathbb{C}}+2\langle\varepsilon_0,v\rangle_F\langle u,\varepsilon_0-i\varepsilon_0\rangle_{\mathbb{C}}+2\langle\varepsilon_0,u\rangle_F\langle\varepsilon_0-i\varepsilon_0,v\rangle_{\mathbb{C}}$

$+2\langle\varepsilon_0,u\rangle_F\langle\varepsilon_0,v\rangle_F\langle\varepsilon_0-i\varepsilon_0,\varepsilon_0-i\varepsilon_0\rangle_{\mathbb{C}}$

$=2\langle u+\langle\varepsilon_0,u\rangle_F(\varepsilon_0-i\varepsilon_0),v+\langle\varepsilon_0,v\rangle_F(\varepsilon_0-i\varepsilon_0)\rangle_{\mathbb{C}}$

$=2\left\langle\widetilde{C}(u),\widetilde{C}(v)\right\rangle_{\mathbb{C}}$

As a consequence, by the universal property of Clifford algebras, there is a unique real algebra morphism $C:Cl(3,1)\to Cl(\mathbb{C},4)$ such that $\widetilde{C}=C\circ\jmath$ where $\jmath$ is the canonical injection $(F,\langle\rangle)\to Cl(3,1)$. We will denote for simplicity $\widetilde{C}=C$. The image $C(Cl(3,1))$ is a real subalgebra of

[3] 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.

$Cl(\mathbb{C},4)$, which can be identified with $Cl(3,1)$ so it does not depend on the choice of ε_0 (but the map C depends on ε_0).

Similarly with $\widetilde{C}'(\varepsilon_j) = i\varepsilon_j, j = 1,2,3; \widetilde{C}'(\varepsilon_0) = \varepsilon_0$ we have a real algebra morphism $C' : Cl(1,3) \to Cl(\mathbb{C},4)$ and $C'(Cl(1,3))$ is a real subalgebra of $Cl(\mathbb{C},4)$. Moreover $C'(\varepsilon_j) = -i\eta_{jj}C(\varepsilon_j)$.

Chirality

In $Cl(\mathbb{C},4)$ the special element is : $\omega = \pm\varepsilon_0 \odot \varepsilon_1 \odot \varepsilon_2 \odot \varepsilon_3 \in Spin(\mathbb{C},4)$. Thus there is a choice and we will use : $\omega = \varepsilon_0 \odot \varepsilon_1 \odot \varepsilon_2 \odot \varepsilon_3$.

$\omega \cdot \omega = 1$ and the map : $f : Cl(\mathbb{C},4) \to Cl(\mathbb{C},4) :: f(X) = \omega \cdot X$ is linear and has for eigen values ± 1. There are two eigen spaces, which are subalgebras :

$Cl(\mathbb{C},4) = Cl^R(\mathbb{C},4) \oplus Cl^L(\mathbb{C},4)$:

$Cl^R(\mathbb{C},4) = \{X \in Cl(\mathbb{C},4) : \omega \odot X = X\}$,

$Cl^L(\mathbb{C},4) = \{X \in Cl(\mathbb{C},4) : \omega \odot X = -X\}$

denoted : $Cl_\epsilon(\mathbb{C},4), \epsilon = \pm 1$

For the representation (E,γ) of $Cl(4,\mathbb{C})$:

$\gamma(\omega)\gamma(\omega) = \gamma(1) = I$ and we have similarly : $E = E^R \oplus E^L$ with

$E^R = \{S \in E : \gamma(\omega) S = S\}, E^L = \{S \in E : \gamma(\omega) S = -S\}$

and the projections : $\gamma_\epsilon(S) = \frac{1}{2}(S + \epsilon(\omega) S)$.

For any homogeneous element $X = v_1 \odot v_2 ... \odot v_k \in Cl(\mathbb{C},4) : \omega \odot X = (-1)^k X \odot \omega$

$\Rightarrow \gamma(\omega)\gamma(X) = (-1)^k \gamma(X)\gamma(\omega)$

$\gamma(\omega)\gamma(X) S = (-1)^k \gamma(X)\gamma(\omega) S$

If $\gamma(\omega) S = \epsilon S : \gamma(\omega)\gamma(X) S = \epsilon(-1)^k \gamma(X) S$. Thus for k even $\gamma(X)$ preserves both E^R, E^L, for k odd $\gamma(X)$ exchanges E^R, E^L.

The choice of the representation γ

A 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 practical choice. Let e_1, e_2 be a basis of E^R and e_3, e_4 a basis of E^L. Then :

$\gamma(\omega) = \gamma_R - \gamma_L = \begin{bmatrix} I_2 & 0 \\ 0 & -I_2 \end{bmatrix}$

Denote : $\gamma_j = \begin{bmatrix} A_j & B_j \\ C_j & D_j \end{bmatrix}$ with four 2×2 complex matrices $j = 0...3$.

$\gamma(\omega)\gamma(\varepsilon_j) = -\gamma(\varepsilon_j)\gamma(\omega)$ 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 :

$\begin{bmatrix} B_j C_k + B_k C_j & 0 \\ 0 & C_j B_k + C_k B_j \end{bmatrix} = 2\delta_{jk} I_4$

$j \neq k : B_j C_k + B_k C_j = C_j B_k + C_k B_j = 0$

$j = k : B_j C_j = C_j B_j = I_2 \Leftrightarrow C_j = B_j^{-1}$

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 : $k=1,2,3: M_k=-iB_kB_0^{-1}$
The matrices $(M_k)_{k=1}^3$ are such that :
$M_k^2=-\left(B_jB_0^{-1}\right)^2=-I_2$
$M_jM_k+M_kM_j=-B_jB_0^{-1}B_kB_0^{-1}-B_kB_0^{-1}B_jB_0^{-1}$
$=-\left(-B_jB_k^{-1}B_0-B_kB_j^{-1}B_0\right)B_0^{-1}$
$=B_jB_k^{-1}+B_kB_j^{-1}=0$
that is $k=1,2,3: M_jM_k+M_kM_j=2\delta_{jk}I_2$
Moreover : $\gamma(\omega)=\gamma_0\gamma_1\gamma_2\gamma_3 \Rightarrow$
$B_0B_1^{-1}B_2B_3^{-1}=I_2$
$B_0^{-1}B_1B_2^{-1}B_3=-I_2$
with $B_k=iM_kB_0, B_k^{-1}=-iB_0^{-1}M_k^{-1}$
$B_0\left(-iB_0^{-1}M_1^{-1}\right)(iM_2B_0)\left(-iB_0^{-1}M_3^{-1}\right)=I_2=-iM_1^{-1}M_2M_3^{-1}$
$B_0^{-1}(iM_1B_0)\left(-iB_0^{-1}M_2^{-1}\right)(iM_3B_0)=-I_2=iB_0^{-1}M_1M_2^{-1}M_3B_0$
which reads :
$iM_2=-M_1M_3=M_3M_1$
$-M_1^{-1}M_3^{-1}=iM_2^{-1} \Leftrightarrow iM_2=M_3M_1$
$M_2M_3+M_3M_2=0=iM_1M_3M_3+M_3M_2 \Leftrightarrow iM_1=-M_3M_2=M_2M_3$
$M_1M_2+M_2M_1=0=iM_3M_2M_2+M_2M_1 \Rightarrow iM_3=-M_2M_1=M_1M_2$
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 :

$$\sigma_1=\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix};\sigma_2=\begin{bmatrix}0 & -i\\i & 0\end{bmatrix};\sigma_3=\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix};\sigma_0=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \tag{4.1}$$

$$\sigma_i^2=\sigma_0;\ \text{For } j\neq k: \sigma_j\sigma_k=\epsilon(j,k,l)\,i\sigma_l \tag{4.2}$$

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 $\gamma(z)=zI_4$

Thus :

$$\gamma_0=\begin{bmatrix}0 & -i\sigma_0\\ i\sigma_0 & 0\end{bmatrix};\gamma_1=\begin{bmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{bmatrix};\gamma_2=\begin{bmatrix}0 & \sigma_2\\ \sigma_2 & 0\end{bmatrix};\gamma_3=\begin{bmatrix}0 & \sigma_3\\ \sigma_3 & 0\end{bmatrix}; \tag{4.3}$$

The matrices γ_j are then unitary and Hermitian :

$$\gamma_j=\gamma_j^*=\gamma_j^{-1} \tag{4.4}$$

which is extremely convenient.

We will use the following (see the annex for more formulas) :

Notation 67 $j=1,2,3: \widetilde{\gamma}_j=\begin{bmatrix}\sigma_j & 0\\ 0 & \sigma_j\end{bmatrix}$

$j\neq k, l=1,2,3: \gamma_j\gamma_k=-\gamma_k\gamma_j=i\epsilon(j,k,l)\,\widetilde{\gamma}_l$
$j=1,2,3: \gamma_j\gamma_0=-\gamma_0\gamma_j=i\begin{bmatrix}\sigma_j & 0\\ 0 & -\sigma_j\end{bmatrix}=i\gamma_5\widetilde{\gamma}_j$

Representation of the real Clifford Algebras

Notice that the choice of the matrices is done in $Cl(\mathbb{C},4)$, so it is independent of the signature. We get the representations of the real algebras by the matrices $\gamma C(\varepsilon_j)$ and $\gamma C'(\varepsilon_j)$

$$\begin{array}{l} Cl(3,1): \gamma C(\varepsilon_j) = \gamma_j, j = 1,2,3; \gamma C(\varepsilon_0) = i\gamma_0; \gamma C(\varepsilon_5) = i\gamma_5 \\ Cl(1,3): \gamma C'(\varepsilon_j) = i\gamma_j, j = 1,2,3; \gamma C'(\varepsilon_0) = \gamma_0; \gamma C'(\varepsilon_5) = \gamma_5 \end{array} \tag{4.5}$$

However, because C is a real, and not a complex map : $\gamma C(\lambda X) \neq \lambda \gamma C(X)$ if $\lambda \in \mathbb{C}$.

The representation that we have chosen here is not unique and others, equivalent, would hold. However the defining relations are rather strong and the choices which give manageable matrices are limited. In the Standard Model the representation of $Cl(1,3)$ is by the matrices : $\widetilde{\gamma}_0 = i\gamma_0, \widetilde{\gamma}_j = \gamma_j, j = 1,2,3$ and $\widetilde{\gamma}_5 = -i\widetilde{\gamma}_0\widetilde{\gamma}_1\widetilde{\gamma}_2\widetilde{\gamma}_3$.

Invariant vector subspaces

(E,γ) is a faithful, and thus irreducible, representation of $Cl(4,\mathbb{C})$, and because $C(Cl(3,1)), C'(Cl(1,3))$ are real subalgebras of $Cl(4,\mathbb{C})$, the set of vectors of E which are invariant by γC is the set invariant by $\gamma C(\varepsilon_j), j = 0..3$ and similarly with $\gamma C'$.

Let be the vector subspaces :

$E_\epsilon = \left\{ \begin{bmatrix} S_R \\ S_L \end{bmatrix} \in E : S_L = \epsilon i S_R = \begin{bmatrix} u \\ v \end{bmatrix} \in \mathbb{C}^2 \right\}, \epsilon = \pm 1$

then :

with $Cl(3,1)$

$i\gamma_0 \begin{bmatrix} S_R \\ \epsilon i S_R \end{bmatrix} = \begin{bmatrix} 0 & \sigma_0 \\ -\sigma_0 & 0 \end{bmatrix} \begin{bmatrix} S_R \\ \epsilon i S_R \end{bmatrix} = \begin{bmatrix} i\epsilon S_R \\ -S_R \end{bmatrix} = \begin{bmatrix} i\epsilon S_R \\ \epsilon i (i\epsilon S_R) \end{bmatrix}$

$\gamma_j \begin{bmatrix} S_R \\ \epsilon i S_R \end{bmatrix} = \begin{bmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{bmatrix} \begin{bmatrix} S_R \\ \epsilon i S_R \end{bmatrix} = \begin{bmatrix} i\epsilon \sigma_j S_R \\ \sigma_j S_R \end{bmatrix} = \begin{bmatrix} i\epsilon S_R \sigma_j \\ -i\epsilon (i\epsilon \sigma_j S_R) \end{bmatrix}$

$S_0 \in E_\epsilon \Rightarrow \gamma_0 C(\varepsilon_0) S_0 \in E_\epsilon, j = 1,2,3 : \gamma C(\varepsilon_j) S_0 \in E_{-\epsilon}$

with $Cl(1,3)$

$\gamma_0 \begin{bmatrix} S_R \\ \epsilon i S_R \end{bmatrix} = \begin{bmatrix} 0 & -i\sigma_0 \\ i\sigma_0 & 0 \end{bmatrix} \begin{bmatrix} S_R \\ \epsilon i S_R \end{bmatrix} = \begin{bmatrix} \epsilon S_R \\ i S_R \end{bmatrix} = \begin{bmatrix} \epsilon S_R \\ i\epsilon (\epsilon S_R) \end{bmatrix}$

$i\gamma_j \begin{bmatrix} S_R \\ \epsilon i S_R \end{bmatrix} = \begin{bmatrix} 0 & i\sigma_j \\ i\sigma_j & 0 \end{bmatrix} \begin{bmatrix} S_R \\ \epsilon i S_R \end{bmatrix} = \begin{bmatrix} -\epsilon \sigma_j S_R \\ i\sigma_j S_R \end{bmatrix} = \begin{bmatrix} -\epsilon \sigma_j S_R \\ -\epsilon i (-\epsilon \sigma_j S_R) \end{bmatrix}$

$S_0 \in E_\epsilon \Rightarrow \gamma_0 C'(\varepsilon_0) S_0 \in E_\epsilon, j = 1,2,3 : \gamma C'(\varepsilon_j) S_0 \in E_{-\epsilon}$

So the set $E_0 = E_+ \cup E_-$ is globally invariant by both $Cl(3,1), Cl(1,3)$. It is not a vector space.

Expression of the γ matrices

Complex notation with the Dirac's matrices

With complex vector spaces the following notation is very convenient.

Define, for any $z \in \mathbb{C}^3$:

Notation 68 $\sum_{a=1}^3 z_a \sigma_a = \sigma(z)$ *with* $z \in \mathbb{C}^3$

$\sigma(z) = \begin{bmatrix} z_3 & z_1 - iz_2 \\ z_1 + iz_2 & -z_3 \end{bmatrix} \in sl(\mathbb{C},2)$

Then we have the identities (see Formulas in the Annex for more) :

$(\sigma(z))^* = \sigma(\overline{z})$

$\sigma(z)\sigma(z') = \sigma(j(z)z') + z^t z' \sigma_0$

$\sigma(z)\sigma(z') - \sigma(z')\sigma(z) = \sigma(j(z)z') - \sigma(j(z')z) = 2\sigma(j(z)z')$

$\sigma(z')\sigma(z)\sigma(z') = \left((z')^t z'\right)\sigma(z)$

Representations of the elements of the Lie algebras

In Cl(3,1) :

$$\gamma C\left(\upsilon\left(r,w\right)\right)=-\frac{1}{2}i\begin{bmatrix}\sigma\left(r+iw\right) & 0\\ 0 & \sigma\left(r-iw\right)\end{bmatrix}=-\frac{1}{2}i\begin{bmatrix}\sigma\left(Z\right) & 0\\ 0 & \sigma\left(\overline{Z}\right)\end{bmatrix} \tag{4.6}$$

In Cl(1,3) :

$$\gamma C'\left(\upsilon\left(r,w\right)\right)=\frac{1}{2}i\begin{bmatrix}\sigma\left(r-iw\right) & 0\\ 0 & \sigma\left(r+iw\right)\end{bmatrix} \tag{4.7}$$

Representations of the elements of the Spin group

$\gamma C\left(a+\upsilon\left(r,w\right)+b\varepsilon_5\right)=aI_4+\gamma C\left(\upsilon\left(r,w\right)\right)+b\gamma_5$

In Cl(3,1) :

$\gamma C\left(a+\upsilon\left(r,w\right)+b\varepsilon_5\right)=\begin{bmatrix}a+ib-\frac{1}{2}i\sigma\left(r+iw\right) & 0\\ 0 & a-ib-\frac{1}{2}i\sigma\left(r-iw\right)\end{bmatrix}$

$=\begin{bmatrix}A-\frac{1}{2}i\sigma\left(Z\right) & 0\\ 0 & \overline{A}-\frac{1}{2}i\sigma\left(\overline{Z}\right)\end{bmatrix}$

In Cl(1,3) :

$\gamma C'\left(a+\upsilon\left(r,w\right)+b\varepsilon_5\right)=\begin{bmatrix}a-ib+\frac{1}{2}i\sigma\left(r-iw\right) & 0\\ 0 & a+ib+\frac{1}{2}i\sigma\left(r+iw\right)\end{bmatrix}$

$=\begin{bmatrix}\overline{A}-\frac{1}{2}i\sigma\left(\overline{Z}\right) & 0\\ 0 & A-\frac{1}{2}i\sigma\left(Z\right)\end{bmatrix}$

Some properties of the $\gamma C\left(\sigma\right)$ matrices

Because γC is a representation of the Clifford algebra, the operations on $Spin\left(3,1\right)$ extend to the matrices with the exception of the multiplication by a complex number.

$\gamma C\left(A+Z\right)^{-1}=\gamma C\left(A-Z\right)$

$\gamma C\left(A_1+Z_1\right)\gamma C\left(A_2+Z_2\right)=\gamma C\left(\left(A_1+Z_1\right)\cdot\left(A_2+Z_2\right)\right)$

$\gamma C\left(\sigma^{-1}\cdot\partial_\alpha\sigma\right)=\gamma C\left(D\left(-Z\right)\frac{\partial Z}{\partial x}\right)=\gamma C\left(D\left(-Z\right)\right)\gamma C\left(\frac{\partial Z}{\partial x}\right)$

$\gamma C\left(\frac{\partial\sigma}{\partial x}\cdot\sigma^{-1}\right)=\gamma C\left(D\left(Z\right)\frac{\partial Z}{\partial x}\right)=\gamma C\left(D\left(Z\right)\right)\gamma C\left(\frac{\partial Z}{\partial x}\right)$

$\gamma C\left(\mathbf{Ad}_gX\right)=\gamma C\left(\left(1+Aj\left(Z\right)+\frac{1}{2}j\left(Z\right)j\left(Z\right)\right)\left[X\right]\right)$

4.2.2 Scalar product of Spinors

We need a scalar product on E, preserved by a gauge transformation, that is by $Spin(3,1), Spin(1,3)$.

Theorem 69 *The only scalar products on E, preserved by $\{\gamma C\left(\sigma\right),\sigma\in Spin(3,1)\}$ are $G=\begin{bmatrix}0 & k\sigma_0\\ \overline{k}\sigma_0 & 0\end{bmatrix}$ with $k\in\mathbb{C}$*

Proof. The scalar product is represented in the basis of E by a 4×4 Hermitian matrix G such that : $G=G^*$

$\forall s\in Spin(3,1):\left[\gamma C\left(s\right)\right]^*G\left[\gamma C\left(s\right)\right]=G$

or $\forall s\in Spin(1,3):\left[\gamma C'\left(s\right)\right]^*G\left[\gamma C'\left(s\right)\right]=G$

$\left[\gamma C\left(s\right)\right]^*\left[G\right]=\left[G\right]\left[\gamma C\left(s\right)\right]^{-1}=\left[G\right]\left[\gamma C\left(s^{-1}\right)\right]$

$[\gamma C(s)] = \gamma C(A+Z) = \begin{bmatrix} A\sigma_0 - \frac{1}{2} i\sigma(Z) & 0 \\ 0 & \overline{A}\sigma_0 - \frac{1}{2} i\sigma(\overline{Z}) \end{bmatrix}$

$[\gamma C(s)]^* = \begin{bmatrix} \overline{A}\sigma_0 + \frac{1}{2} i\sigma(\overline{Z}) & 0 \\ 0 & A\sigma_0 + \frac{1}{2} i\sigma(Z) \end{bmatrix}$

$[\gamma C(s^{-1})] = \gamma C(A-Z) = \begin{bmatrix} A\sigma_0 + \frac{1}{2} i\sigma(Z) & 0 \\ 0 & \overline{A}\sigma_0 + \frac{1}{2} i\sigma(\overline{Z}) \end{bmatrix}$

$G = \begin{bmatrix} M & P \\ P^* & N \end{bmatrix}$, with $M = M^*, N = N^*$

$[G][\gamma C(s)]^{-1} = [G][\gamma C(s^{-1})] \Leftrightarrow$

$\begin{bmatrix} \left(\overline{A}\sigma_0 + \frac{1}{2} i\sigma(\overline{Z})\right) M & \left(\overline{A}\sigma_0 + \frac{1}{2} i\sigma(\overline{Z})\right) P \\ \left(A\sigma_0 + \frac{1}{2} i\sigma(Z)\right) P^* & \left(A\sigma_0 + \frac{1}{2} i\sigma(Z)\right) N \end{bmatrix}$

$= \begin{bmatrix} M\left(A\sigma_0 + \frac{1}{2} i\sigma(Z)\right) & P\left(\overline{A}\sigma_0 + \frac{1}{2} i\sigma(\overline{Z})\right) \\ P^*\left(A\sigma_0 + \frac{1}{2} i\sigma(Z)\right) & N\left(\overline{A}\sigma_0 + \frac{1}{2} i\sigma(\overline{Z})\right) \end{bmatrix}$

We must have the identities, $\forall Z$:

$\left(\overline{A}\sigma_0 + \frac{1}{2} i\sigma(\overline{Z})\right) M = M\left(A\sigma_0 + \frac{1}{2} i\sigma(Z)\right)$

$\left(\overline{A}\sigma_0 + \frac{1}{2} i\sigma(\overline{Z})\right) P = P\left(\overline{A}\sigma_0 + \frac{1}{2} i\sigma(\overline{Z})\right)$

$\left(A\sigma_0 + \frac{1}{2} i\sigma(Z)\right) P^* = P^*\left(A\sigma_0 + \frac{1}{2} i\sigma(Z)\right)$

$\left(A\sigma_0 + \frac{1}{2} i\sigma(Z)\right) N = N\left(\overline{A}\sigma_0 + \frac{1}{2} i\sigma(\overline{Z})\right)$

Let us consider first $s \in Spin(3) \Leftrightarrow A, Z \in \mathbb{R}$

The conditions read

$\sigma(Z) M = M\sigma(Z)$

$\sigma(Z) P = P\sigma(Z)$

$\sigma(Z) P^* = P^*\sigma(Z)$

$\sigma(Z) N = N\sigma(Z)$

The only matrices which commute with all Dirac matrices are scalar, thus :

$M = m\sigma_0, N = n\sigma_0, P = p\sigma_0$

$G = \begin{bmatrix} m\sigma_0 & p\sigma_0 \\ \overline{p}\sigma_0 & n\sigma_0 \end{bmatrix}$, with $m, n \in \mathbb{R}$

Then for $s \in Spin(3,1)$ the conditions become :

$\left(\overline{A}\sigma_0 + \frac{1}{2} i\sigma(\overline{Z})\right) m = m\left(A\sigma_0 + \frac{1}{2} i\sigma(Z)\right)$

$\left(\overline{A}\sigma_0 + \frac{1}{2} i\sigma(\overline{Z})\right) p = p\left(\overline{A}\sigma_0 + \frac{1}{2} i\sigma(\overline{Z})\right)$

$\left(A\sigma_0 + \frac{1}{2} i\sigma(Z)\right) \overline{p} = \overline{p}\left(A\sigma_0 + \frac{1}{2} i\sigma(Z)\right)$

$\left(A\sigma_0 + \frac{1}{2} i\sigma(Z)\right) n = n\left(\overline{A}\sigma_0 + \frac{1}{2} i\sigma(\overline{Z})\right)$

$\overline{A}m = mA \Rightarrow m = 0$

$An = n\overline{A} \Rightarrow n = 0$

The only solution is :

$G = \begin{bmatrix} 0 & k\sigma_0 \\ \overline{k}\sigma_0 & 0 \end{bmatrix}$ ■

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 :

$S = \sum_{i=1}^4 S^i e_i = \begin{bmatrix} S_R \\ S_L \end{bmatrix}$ with 2 vectors $S_R, S_L \in \mathbb{C}^2$

The scalar product of two vectors S, S' of E is then:

$$\langle S, S' \rangle_E = [S]^* [\gamma_0] [S'] = i\left([S_L]^* [S'_R] - [S_R]^* [S'_L]\right) \tag{4.8}$$

It is not definite positive but :

$[S_L]^* [S_R] = \left([S_L]^* [S_R]\right)^t = [S_R]^t \overline{[S_L]} = \overline{\left([S_R]^* [S_L]\right)}$
$\Rightarrow$

$$\left[\langle S, S\rangle_E = -2\operatorname{Im}\left([S_L]^* [S_R]\right)\right] \tag{4.9}$$

And if $S \in E_\epsilon : S_L = \epsilon i S_R : \langle S, S\rangle_E = -2\operatorname{Im}\left(-\epsilon i [S_R]^* [S_R]\right) = 2\epsilon [S_R]^* [S_R]$ thus the scalar product is definite positive on E_+ and definite negative on E_-. These two vector spaces are Hilbert spaces.

An orthonormal basis is :

$\frac{1}{\sqrt{2}}(e_1 + ie_3), \frac{1}{\sqrt{2}}(e_2 + ie_4), \frac{1}{\sqrt{2}}(-e_1 + ie_3), \frac{1}{\sqrt{2}}(-e_2 + ie_4) : \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ i & 0 & i & 0 \\ 0 & i & 0 & i \end{bmatrix}$

Norm on the space of spinors

E_ϵ are Hilbert spaces, so normed vector spaces. More generally there is a norm on E : $\|S\| = \sqrt{[S]^* [S]}$

4.3 THE SPINOR REPRESENTATION OF MOMENTA

4.3.1 The Spinor bundle

Because M is endowed with the structure of the principal bundle P_G, there is a structure of spin bundle (Maths.2110), an associated vector bundle $P_G[E,\gamma C]$ such that, at each point of M, any element of $Cl(3,1)$ acts on the vectors of $P_G[E,\gamma C]$ through γC.

Definition 70 *The* ***Spinor bundle*** *is the associated vector bundle* $P_G[E,\gamma C]$

Its elements S are **spinors**. They are measured by observers in the standard gauge defined through the holonomic basis : $\mathbf{e}_i(m) = (\mathbf{p}(m), e_i)$

In a change of gauge the holonomic basis becomes :

$$\begin{gathered}\mathbf{p}(m) = \varphi_G(m,1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m) \cdot \chi(m)^{-1} : \\ \mathbf{e}_i(m) = (\mathbf{p}(m), e_i) \to \widetilde{\mathbf{e}}_i(m) = \gamma C\left(\chi(m)^{-1}\right)\mathbf{e}_i(m) \\ (\mathbf{p}(m), S) \sim (\widetilde{\mathbf{p}}(m), \gamma C(\chi(m)) S)\end{gathered} \tag{4.10}$$

$(\mathbf{e}_i(m))_{i=1}^4$, are defined through the standard gauge $\mathbf{p}(m)$ chosen by the observer.

A jet in $J^1P_G[E,\gamma C]$ is represented by : $j^1S = (m, S, \delta_\alpha S, \alpha = 0..3)$ where $S, \delta_\alpha S \in E$ and change as in $P_G[E,\gamma C]$.

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 :

$\left\langle \mathbf{S}, \mathbf{S}' \right\rangle = \int_\Omega \left\langle \mathbf{S}(m), \mathbf{S}'(m) \right\rangle_E \varpi_4(m)$

4.3.2 Definition of the Momenta

Definition

Proposition 71 *The kinematic characteristics of a particle are represented in the first jet extension* $J^1P_G[E,\gamma C]$.

Along any trajectory by a map $j^1S : \mathbb{R} \to J^1P_G[E,\gamma C] :: j^1S(t) = (q(t), S(t), \delta S(t))$

$S(t), \delta S(t) \in P_G[E,\gamma C]$ *are located at* $q(t)$.

In a continuous motion j^1S *is a the first jet prolongation of a map :*

$S : \mathbb{R} \to J^1P_G[E,\gamma C] :: \left(q(t), S(t), \frac{dS}{dt}(t)\right)$

Momenta and motion are two distinct concepts. The maps :

$j^1\sigma : \mathbb{R} \to J^1P_G :: (q(t), \sigma(t), \upsilon(X_r, X_w))$

$j^1S : \mathbb{R} \to J^1P_G[E,\gamma C] :: (q(t), S(t), \delta S(t))$

are a priori distinct. The main physical assumption is that there is a relation between the motion and the momentum. In the usual representations the relation is given, for the translational momentum by a scalar, the mass, and for the rotational momentum by a matrix, the inertial tensor. Because we assume that to any particle is associated an orthonormal basis, the momentum requires more than a scalar.

We compute momenta, and measure change of momenta. For any observer $\mathbf{p}(q(t)) = \varphi_G(q(t), 1)$ the motion of the body is along the trajectory : $(q(t), \sigma(t), \upsilon(X_r, X_w))$. The position of the particle is $\sigma(t)$, its motion is $\upsilon(X_r, X_w)$, which, in continuous motion, is $\frac{d\sigma}{dt} \cdot \sigma^{-1}$.

The changes of momenta are related to the instantaneous motion with respect to the previous state of the particle and they are measured by forces and torques. The momenta are then related to the position of the particle, that is to $\sigma(t)$.

We assume that $\exists S_0 \in E : S(t) = \gamma C(\sigma(t)) S_0$

In a continuous motion, the observer measures the change, through inertial forces :

$\frac{d}{dt}S(t) = \gamma C\left(\frac{d}{dt}\sigma(t)\right)S_0 = \gamma C\left(\frac{d}{dt}\sigma(t)\cdot\sigma(t)^{-1}\right)\gamma C(\sigma(t))S_0 = \gamma C\left(\frac{d}{dt}\sigma(t)\cdot\sigma(t)^{-1}\right)S(t)$

And we generalize as : for a, not necessarily continuous, motion $(q(t), \sigma(t), \upsilon(X_r, X_w))$ the momenta follow :

$(q(t), S(t) = \gamma C(\sigma(t))S_0, \delta S(t) = \gamma C(\upsilon(X_r, X_w))S(t))$

Proposition 72 *For any particle there is a fixed differential operator $\mathcal{M}$ such as, for the motion $j^1\sigma = (q(t), \sigma(t), \upsilon(X_r, X_w))$:*

$$\begin{gathered}\mathcal{M} : J^1 Cl(TM) \to J^1 P_G[E, \gamma C] :: \\ \mathcal{M}(q(t), \sigma(t), \upsilon(X_r, X_w)) = (q, S = \gamma C(\sigma)S_0, \delta S = \gamma C(\upsilon(X_r, X_w))S)\end{gathered} \tag{4.11}$$

*where $S_0 \in E$ is a fixed vector called the **inertial spinor**.*

In a change of gauge :

$\mathbf{p}(m) = \varphi_G(m, 1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m)\cdot\chi(m)^{-1}$:

$\sigma \to \widetilde{\sigma} = \chi\cdot\sigma$

$(\mathbf{p}(m), \upsilon(X_r, X_w)) \sim \left(\widetilde{\mathbf{p}}(m), \mathbf{Ad}_{\chi(m)}\upsilon(X_r, X_w)\right) \Leftrightarrow \upsilon\widetilde{(X_r, X_w)} = \mathbf{Ad}_{\chi(m)}\upsilon(X_r, X_w)$

$S \to \widetilde{S} = \gamma C(\chi(m))S$

$\delta S \to \widetilde{\delta S} = \gamma C(\chi(m))\delta S$

$(q(t), S(t), \delta S(t)) \to \left(q(t), \widetilde{S}(t), \widetilde{\delta S}(t)\right)$

and :

$\widetilde{S} = \gamma C(\chi)S = \gamma C(\chi)\gamma C(\sigma)S_0 = \gamma C(\chi\cdot\sigma)S_0 = \gamma C(\widetilde{\sigma})S_0$

$\widetilde{\delta S} = \gamma C(\chi)\delta S = \gamma C(\chi)\gamma C(\upsilon(X_r, X_w))S = \gamma C(\chi)\gamma C(\upsilon(X_r, X_w))\gamma C(\chi^{-1})\widetilde{S}$

$= \gamma C(\mathbf{Ad}_\chi\upsilon(X_r, X_w))\widetilde{S} = \gamma C\left(\upsilon\widetilde{(X_r, X_w)}\right)\widetilde{S}$

So S_0 *does not change* : this is an intrinsic property of the particle, which is measured by an observer through $S = \gamma C(\sigma)S_0$.And $\sigma = 1$ for an observer attached to the particle.

The spinor, which characterizes the momentum (corresponding to $m\overrightarrow{v}$), is $S = \gamma C(\sigma)S_0$.

The change of momentum, equal to the inertial forces (corresponding to $m\overrightarrow{\gamma}$), is $\delta S = \gamma C(\upsilon(X_r, X_w))S$

$\delta S_R = \sum_{\alpha=0}^3 \gamma C(\upsilon(X_r, 0))S$ is the equivalent of a change of rotational momentum or an inertial torque.

$\delta S_T = \sum_{\alpha=0}^3 \gamma C(\upsilon(0, X_w))S$ is the equivalent of a change of translational momentum or a translational inertial force.

Forces, torques and Spinors

i) The motion is represented in the real Clifford algebra. It is legitimate to assume that S_0 belongs to a subset which is invariant by $Cl(3,1)$ (or similarly by $Cl(1,3)$). So we can state :

Proposition 73 *For particles the inertial spinor S_0 belongs to the set of vectors :*

$E_0 = \left\{\begin{bmatrix} S_R \\ S_L \end{bmatrix} \in E : S_L = \pm i S_R\right\}$

Then $\forall s \in Cl(3,1) : \gamma C(s)S_0 \in E_0$ and idem for $Cl(1,3)$ because the set is globally invariant.

ii) (E, γ) is a faithful representation of $Cl(4, \mathbb{C})$ and $(E, \gamma C)$ is a faithful representation of $Cl(3,1)$:

$\forall X, X' \in Cl(3,1), S \in E : \gamma C(X)S = \gamma C(X')S \Leftrightarrow X = X'$

iii) A vector of E, with 4 complex components, can represent :

either a combination of a translational and rotational momentum (S)

or a combination of force and torque (δS).

Forces and torques are measured through the change of motion of known particles.

The action of the fields is represented by a differential operator acting on $j^1 S$:

$D_F : J^1 P_G [E, \gamma C] \to J^1 P_G [E, \gamma C]$

The relation $\sigma \to S$ through S_0 is the mathematical expression of the continuity of the particle. The condition : $S(t) = \gamma C(\sigma(t)) S_0$ provides differential equations with respect to σ which give the motion. Their solutions depend on the value of S_0, which enables to estimate S_0.

The vectors e_i of the basis of E have no universal physical meaning : it depends on the system (as it can be seen in the measure of the spin of an atom by an analyzer). Actually forces and torques are identified by the change of motion with which they are associated, that is by $\upsilon(X_r, X_w)$ in the basis of $T_1 Spin(3,1)$ and not to vectors of the basis ε_i as in Newtonian Mechanics : forces correspond to $\upsilon(0, X_w)$ and torques to $\upsilon(X_r, 0)$. And the identification of the axes e_i can be done, for a rigid solid, through the inertial vector as we will see.

4.3.3 Mass and Kinetic Energy

Mass

The scalar product is invariant by the action of γ, thus :

$\langle S(t), S(t)\rangle = \langle \gamma C(\sigma(t)) S_0, \gamma C(\sigma(t)) S_0\rangle = \langle S_0, S_0\rangle = -2\,\mathrm{Im}\left([S_L]^* [S_R]\right)$

By similarity with $\langle P, P\rangle = -M_p^2 c^2$ it is then natural to state that $\langle S_0, S_0\rangle$ represents the square of the mass of the particle, up to a constant depending on the units.

With the proposition above : $[S_L] = \epsilon i [S_R] \Rightarrow \langle S_0, S_0\rangle = 2\epsilon [S_R]^* [S_R]$

This quantity can be positive or negative. We will come back on this issue later and define the mass “at rest” of the particle by :

$$M_p = \sqrt{|\langle S_0, S_0\rangle|} = \sqrt{2\left|\mathrm{Im}\left([S_L]^* [S_R]\right)\right|} = \sqrt{2 [S_R]^* [S_R]} \tag{4.12}$$

Then S_R reads :

$S_R = \frac{M_p}{\sqrt{2}} \begin{bmatrix} a \\ b \end{bmatrix}$ and $1 = |a|^2 + |b|^2$

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.

And we can write similarly :

$$S_R = \frac{M_p}{\sqrt{2}} \begin{bmatrix} e^{i\alpha_1} \cos \alpha_0 \\ e^{i\alpha_2} \sin \alpha_0 \end{bmatrix} \tag{4.13}$$

Kinetic Energy

$\frac{d}{dt}\langle S(t), S(t)\rangle = 0 = \left\langle \frac{d}{dt} S(t), S(t)\right\rangle + \left\langle S(t), \frac{d}{dt} S(t)\right\rangle$ thus $\left\langle S(t), \frac{d}{dt} S(t)\right\rangle$ is pure imaginary.

The variation of the kinetic energy is defined in Newtonian Mechanics as :

$\delta K = \frac{1}{m}\left\langle \overrightarrow{p}, \overrightarrow{\delta p_G}\right\rangle + [r]^t [R]^t [\delta\Gamma(G)]$

It involves both the present state of momentum and its evolution. The natural generalization is

$\delta K = \frac{1}{M_p}\frac{1}{i}\langle S, \delta S\rangle = \frac{1}{M_p}\frac{1}{i}\langle \gamma C(\sigma) S_0, \gamma C(\upsilon(X_r, X_w)) \gamma C(\sigma) S_0\rangle$

$= \frac{1}{M_p}\frac{1}{i}\langle S_0, \gamma C(\mathbf{Ad}_{\sigma^{-1}} \upsilon(X_r, X_w)) S_0\rangle$

In a continuous motion along the trajectory :

$\upsilon(X_r, X_w) = \frac{d\sigma}{dt} \cdot \sigma^{-1}$

$\frac{dK}{dt} = \frac{1}{M_p}\frac{1}{i}\left\langle \gamma C(\sigma) S_0, \gamma C\left(\frac{d\sigma}{dt} \cdot \sigma^{-1}\right) \gamma C(\sigma) S_0\right\rangle = \frac{1}{M_p}\frac{1}{i}\left\langle S_0, \gamma C\left(\sigma^{-1} \cdot \frac{d\sigma}{dt}\right) S_0\right\rangle$

$$\begin{aligned} \delta K &= \frac{1}{M_p}\frac{1}{i}\langle S, \delta S\rangle = \frac{1}{M_p}\frac{1}{i}\langle S_0, \gamma C(\mathbf{Ad}_{\sigma^{-1}} \upsilon(X_r, X_w)) S_0\rangle \\ \frac{dK}{dt} &= \frac{1}{M_p}\frac{1}{i}\left\langle S_0, \gamma C\left(\sigma^{-1} \cdot \frac{d\sigma}{dt}\right) S_0\right\rangle \end{aligned} \tag{4.14}$$

The scalar product does not depend on the observer, however in a continuous motion the observer is involved in the definition of t.

$\mathbf{Ad}_{\sigma^{-1}}\upsilon\left(X_r,X_w\right)=\sum_{a=1}^{6}\delta X^a\overrightarrow{\kappa}_a$ is the instantaneous change of motion

$\delta K_a=\frac{1}{M_p}\frac{1}{i}\left\langle S_0,\gamma C\left(\delta X^a\overrightarrow{\kappa}_a\right)S_0\right\rangle$ is the variation of kinetic energy due to a change of motion δX^a in the direction $\overrightarrow{\kappa}_a$.

Inertial vector

Let us denote $[S_0]=\begin{bmatrix}S_R\\S_L\end{bmatrix}$, $Z\in T_1Spin\left(3,1\right)$ in the complex formalism.

$\gamma C\left(Z\right)[S_0]=-\frac{i}{2}\begin{bmatrix}\sigma\left(Z\right)&0\\0&\sigma\left(\overline{Z}\right)\end{bmatrix}\begin{bmatrix}S_R\\S_L\end{bmatrix}$

$\left\langle S_0,\gamma C\left(Z\right)S_0\right\rangle=-\frac{i}{2}\begin{bmatrix}S_R^*&S_L^*\end{bmatrix}\begin{bmatrix}0&-i\sigma_0\\i\sigma_0&0\end{bmatrix}\begin{bmatrix}\sigma\left(Z\right)S_R\\\sigma\left(\overline{Z}\right)S_L\end{bmatrix}$

$=\frac{1}{2}\left(-S_R^*\sigma\left(\overline{Z}\right)S_L+S_L^*\sigma\left(Z\right)S_R\right)$

$S_L^*\sigma\left(Z\right)S_R=\left(S_L^*\sigma\left(Z\right)S_R\right)^t=S_R^t\left[\sigma\left(Z\right)\right]^t\overline{S_L}=\overline{\overline{S}_R^t\overline{\left[\sigma\left(Z\right)\right]}^tS_L}=\overline{S_R^*\sigma\left(\overline{Z}\right)S_L}$

$\left\langle S_0,\gamma C\left(Z\right)S_0\right\rangle=\frac{1}{2}\left(-\overline{S_L^*\sigma\left(Z\right)S_R}+S_L^*\sigma\left(Z\right)S_R\right)=i\operatorname{Im}S_L^*\sigma\left(Z\right)S_R$

Denote the vector : $k\in\mathbb{C}^3:k^a=S_L^*\sigma_aS_R$ then $S_L^*\sigma\left(Z\right)S_R=\sum_{a=1}^{3}Z^aS_L^*\sigma_aS_R=k^tZ$. And one can check that : $k^tk=\left(\left[S_L\right]^*\left[S_R\right]\right)^2$

$\left\langle S_0,\gamma C\left(Z\right)S_0\right\rangle=i\operatorname{Im}k^tZ$

$a=1,2,3$

Take $\upsilon\left(X_r,X_w\right)=\overrightarrow{\kappa_a}$

$\left\langle S_0,\gamma C\left(\overrightarrow{\kappa_a}\right)S_0\right\rangle_E=i\operatorname{Im}k^a=-\frac{1}{2}i\left\langle S_0,\widetilde{\gamma}_aS_0\right\rangle_E$

$\operatorname{Im}k^a=-\frac{1}{2}\left\langle S_0,\widetilde{\gamma}_aS_0\right\rangle_E$

$\delta K_a=\frac{1}{M_p}\frac{1}{i}\left\langle S_0,\gamma C\left(\overrightarrow{\kappa}_a\right)S_0\right\rangle=\frac{1}{M_p}\operatorname{Im}k^a$

Take $\upsilon\left(X_r,X_w\right)=\overrightarrow{\kappa_{a+3}}$

$\left\langle S_0,\gamma C\left(\overrightarrow{\kappa_{a+3}}\right)S_0\right\rangle_E=\left\langle S_0,\gamma C\left(i\overrightarrow{\kappa_a}\right)S_0\right\rangle_E=i\operatorname{Im}ik^a=i\operatorname{Re}k^a=\frac{1}{2}i\left\langle S_0,\gamma_0\gamma_aS_0\right\rangle_E$

$\operatorname{Re}k^a=\frac{1}{2}\left\langle S_0,\gamma_0\gamma_aS_0\right\rangle_E$

$\delta K_{a+3}=\frac{1}{M_p}\frac{1}{i}\left\langle S_0,\gamma C\left(\overrightarrow{\kappa}_a\right)S_0\right\rangle=\frac{1}{M_p}\operatorname{Re}k^a$

$k^a=\frac{1}{2}\left\langle S_0,\gamma_0\gamma_aS_0\right\rangle_E+i\left(-\frac{1}{2}\left\langle S_0,\widetilde{\gamma}_aS_0\right\rangle_E\right)=\frac{1}{2}\left\langle S_0,\left(\gamma_0\gamma_a-i\widetilde{\gamma}_a\right)S_0\right\rangle_E$

$k^a=\frac{1}{2}\left\langle S_0,\left(\gamma_0\gamma_a-i\widetilde{\gamma}_a\right)S_0\right\rangle_E$ corresponds to the Dirac's current. It does not depend on the trajectory or the motion.

So far $k\in\mathbb{C}^3$, however when $S_0\in E_0:S_R=\frac{M_p}{\sqrt{2}}\begin{bmatrix}e^{i\alpha_1}\cos\alpha_0\\e^{i\alpha_2}\sin\alpha_0\end{bmatrix}$, $S_L=i\epsilon S_R$:

$k^a=S_L^*\sigma_aS_R=-i\epsilon S_R^*\sigma_aS_R$

$k=-i\epsilon\frac{M_p^2}{2}\begin{bmatrix}\left(\sin2\alpha_0\right)\cos\left(\alpha_1-\alpha_2\right)\\-\left(\sin2\alpha_0\right)\sin\left(\alpha_1-\alpha_2\right)\\\cos2\alpha_0\end{bmatrix}=-i\epsilon\frac{M_p^2}{2}k_0$

with $k_0^tk_0=1$

Then

$\delta K=\frac{1}{M_p}\frac{1}{i}\left\langle S_0,\gamma C\left(\mathbf{Ad}_{\sigma^{-1}}\upsilon\left(X_r,X_w\right)\right)S_0\right\rangle=\frac{1}{M_p}\operatorname{Im}k^t\mathbf{Ad}_{\sigma^{-1}}\upsilon\left(X_r,X_w\right)\delta K$

$=-\epsilon\frac{M_p}{2}k_0^t\operatorname{Re}\mathbf{Ad}_{\sigma^{-1}}\upsilon\left(X_r,X_w\right)$

In a continuous motion :

$\upsilon\left(X_r,X_w\right)=\frac{d\sigma}{dt}\cdot\sigma^{-1}\Leftrightarrow\sigma^{-1}\cdot\frac{d\sigma}{dt}=\mathbf{Ad}_{\sigma^{-1}}\upsilon\left(X_r,X_w\right)$

$\frac{dK}{dt}=\frac{1}{M_p}\operatorname{Im}k^t\left(\sigma^{-1}\cdot\frac{d\sigma}{dt}\right)=-\epsilon\frac{M_p}{2}k_0^t\operatorname{Re}\left(\sigma^{-1}\cdot\frac{d\sigma}{dt}\right)$

We sum up the results :

$$\begin{gathered} a=1,2,3: k^a=S_L^* \sigma_a S_R=\frac{1}{2} i\left\langle S_0,\left(\gamma_0 \gamma_a-\widetilde{\gamma}_a\right) S_0\right\rangle_E \\ \delta K=\frac{1}{M_p} \frac{1}{i}\left\langle S_0, \gamma C\left(\mathbf{Ad}_{\sigma^{-1}} \upsilon\left(X_r, X_w\right)\right) S_0\right\rangle=\frac{1}{M_p} \operatorname{Im} k^t \mathbf{Ad}_{\sigma^{-1}} \upsilon\left(X_r, X_w\right) \\ S_0 \in E_0: k=-i \epsilon \frac{M_p^2}{2} k_0 \\ k_0=\left[\begin{array}{c} \left(\sin 2 \alpha_0\right) \cos \left(\alpha_1-\alpha_2\right) \\ -\left(\sin 2 \alpha_0\right) \sin \left(\alpha_1-\alpha_2\right) \\ \cos 2 \alpha_0 \end{array}\right] ; k_0^t k_0=1 \\ \delta K=-\epsilon \frac{M_p}{2} k_0^t \operatorname{Re} \mathbf{Ad}_{\sigma^{-1}} \upsilon\left(X_r, X_w\right) \end{gathered} \tag{4.15}$$

The vector k, that we will call the **inertial vector**, does not depend on the state or the motion of the particle. In a change of gauge S_0 does not change, so $k^a = S_L^* \sigma_a S_R$ does not change. k and $\langle S_0, S_0 \rangle$ characterize the kinematic features of the material body. They are defined by 7 independent parameters, as we have in Newtonian Mechanics, and 4 when $S_0 \in E_0$. Two material bodies such that $S_0' = e^{i\alpha} S_0$ with $\alpha \in \mathbb{R}$ have the same kinematic characteristics.

In all practical applications this is the vector k_0 which is involved, the basis $(e_i)_{i=1}^4$ and the inertial spinor S_0 are only used to identify the forces and torques, and this is done in conventional bases depending on the problem, as required (that is in relation with physical measures).

Units

We have seen that the motion, expressed as $\upsilon(X_r, X_w)$, is measured in units $[T]^{-1}$. From $M_p = \sqrt{|\langle S_0, S_0 \rangle|}$, $S = \gamma C(\sigma) S_0$ the spinor is measured in units $[M]^{1/2}$, $k^a = S_L^* \sigma_a S_R$ in units of mass (as well as the inertial tensor $[J]$ in Newtonian Mechanics), and $k_0 = 2i\epsilon M_p^{-2} k$ in units $[M]^{-1}$.

The Dirac's current $k^a = -i\epsilon \frac{M_p^2}{2} k_0^a$ represents the flow of matter through a plane orthogonal to ε_a

The variation of kinetic energy :
$\frac{dK}{dt} = \frac{1}{M_p} \frac{1}{i} \langle S_0, \gamma C(\mathbf{Ad}_{\sigma^{-1}} \upsilon(X_r, X_w)) S_0 \rangle = \frac{1}{M_p} \operatorname{Im} k^t \mathbf{Ad}_{\sigma^{-1}} \upsilon(X_r, X_w)$
is measured in units of $[M][L]^2[T]^{-3}$ so we need a universal constant C_I of units of energy $[E] = [M][L]^2[T]^{-2}$:
$\frac{dK}{dt} = -C_I \epsilon \frac{M_p}{2} k_0^t \operatorname{Re} \mathbf{Ad}_{\sigma^{-1}} \upsilon(X_r, X_w)$
The variation of kinetic energy due to the motion $\upsilon(X_r, X_w)$ can be represented as a flow and $-\epsilon C_I \frac{M_p}{2} \operatorname{Re} \mathbf{Ad}_{\sigma^{-1}} \upsilon(X_r, X_w) = \sum_{a=1}^3 \phi^a \varepsilon_a$
and $k_0^a \phi^a$ is the flow of variation of kinetic energy measured through a plane orthogonal to ε_a.

4.3.4 Momenta of Deformable Solids

Spinor Fields

A section of P_G can represent the motion of particles whose trajectories do not cross and have similar behavior. And a section of $P_G[E, \gamma C]$ can represent the kinematic characteristics of identical particles.

Definition 74 *A* ***Spinor field*** *is a section* $\mathbf{S} \in \mathfrak{X}\left(J^1 P_G[E, \gamma C]\right)$ *which represents the kinematics characteristics of a particle.* $S = (m, S(m), \delta_\beta S(m), \beta = 0..3)$

From a Mathematical point of view the condition is that there is a section $J^1\sigma \in \mathfrak{X}\left(J^1 P_G\right)$ and an inertial spinor S_0 such that :

$S(m) = \gamma C(\sigma(m)) S_0, \delta_\alpha S(m) = \gamma C(\upsilon(X_{r\alpha}(m), X_{w\alpha}(m))) S(m)$. A necessary condition is that : $\langle S(m), S(m) \rangle_E = Ct$.

From a Physical point of view such a section represents particles which have the same kinematics characteristics and whose trajectories do not cross. As a consequence the motion is continuous and $\upsilon\left(X_{r\alpha}\left(m\right),X_{w\alpha}\left(m\right)\right)=\partial_{\alpha}\sigma\cdot\sigma^{-1}$.

Conversely a vector $S_0\in E$ and a section $J^1\sigma\in\mathfrak{X}\left(J^1P_G\right)$ defines a spinor field.

Density

With a population of similar particles represented by a spinor field it is natural to consider a density of particles, that is a function $\mu:M\rightarrow\mathbb{R}$ such that $\mu\left(m\right)$ represents the number of identical particles located at the same point. Then for any observer the conservation of the number of particles implies that :

$\mathcal{N}\left(t\right)=\int_{\Omega(t)}\mu_3\left(t,x\right)\varpi_3=Ct$

which can be written :

$\int_{\Omega(t)}i_V\left(\mu\varpi_4\right)=Ct$

where V is the vector field representing the trajectories, as it is deduced from $\sigma:V=-\frac{c}{\left\langle\mathbf{Ad}_{\sigma}\varepsilon_0,\varepsilon_0\right\rangle_{Cl}}\mathbf{Ad}_{\sigma}\varepsilon_0$.

Consider the manifold $\Omega\left(\left[t_1,t_2\right]\right)$ with borders $\Omega\left(t_1\right),\Omega\left(t_2\right)$:

$\mathcal{N}\left(t_2\right)-\mathcal{N}\left(t_1\right)=\int_{\partial\Omega([t_1,t_2])}i_V\left(\mu\varpi_4\right)=\int_{\Omega([t_1,t_2])}d\left(i_V\mu\varpi_4\right)$

$d\left(i_V\mu\varpi_4\right)=\pounds_V\left(\mu\varpi_4\right)-i_Vd\left(\mu\varpi_4\right)=\pounds_V\left(\mu\varpi_4\right)-i_V\left(d\mu\wedge\varpi_4\right)-i_V\mu d\varpi_4=\pounds_V\left(\mu\varpi_4\right)$

$\mathcal{N}\left(t_2\right)-\mathcal{N}\left(t_1\right)=\int_{\omega([t_1,t_2])}\pounds_V\left(\mu\varpi_4\right)$

with the Lie derivative $\pounds$. The conservation of the number of particles is equivalent to the condition $\pounds_V\left(\mu\varpi_4\right)=0$.

$\pounds_V\mu\varpi_4$

$=\frac{d\mu}{dt}\varpi_4+\mu\pounds_V\varpi_4$

$=\frac{d\mu}{dt}\varpi_4+\mu\left(divV\right)\varpi_4$

$=\left(\frac{d\mu}{dt}+\mu\left(divV\right)\right)\varpi_4$

and we retrieve the usual **continuity equation** :

$$\frac{d\mu}{dt}+\mu divV=0 \tag{4.16}$$

With :

$divV=\sum_{\alpha=0}^{3}\frac{\partial V^{\alpha}}{\partial\xi^{\alpha}}+\frac{1}{2}\frac{1}{\det g}\sum_{\alpha=0}^{3}V^{\alpha}\partial_{\alpha}\left(\det g\right)$

$V=c\varepsilon_0+\sum_{\alpha,j=1}^{3}Q_j^{\alpha}U^j\partial\xi_{\alpha}$ because the motion is defined in the standard chart of the observer

$\alpha>0:\partial_{\beta}V^{\alpha}=\sum_{j=0}^{3}\left(P_j^{\alpha}-\frac{1}{c}P_j^0V^{\alpha}\right)\left[\partial_{\beta}\sigma\cdot\sigma^{-1},U\right]^j$

$[g]=\begin{bmatrix}-1&0\\0&[g_3]\end{bmatrix}$ because the observer uses his standard gauge

$divV=\sum_{\alpha,j=1}^{3}Q_j^{\alpha}\left\{\left[\partial_{\alpha}\sigma\cdot\sigma^{-1},U\right]^j+\frac{1}{2}\frac{1}{\det g}U^j\partial_{\alpha}\left(\det g\right)\right\}$

Let us define :

$T:TM\otimes J^1P_G\left[E,\gamma C\right]\rightarrow\mathbb{R}::$

$T\left(\sum_{\alpha}U^{\alpha}\partial\xi_{\alpha}\right)=\frac{1}{i}\sum_{\alpha,\beta=0}^{3}\mu\left\langle S_0,U^{\alpha}\upsilon\left(X_{r\beta},X_{w\beta}\right)\right\rangle=-\frac{1}{2}\mu\epsilon k_0^t\operatorname{Im}\left(\sum_{\alpha,\beta=0}^{3}\frac{U^{\alpha}}{c}\left(\sigma^{-1}\cdot\partial_{\beta}\sigma\right)\right)$

T is a tensor : its action is linear, and the result does not depend on the chart or the gauge. It gives the resistance of the particle to change its motion by $\sigma^{-1}\cdot\partial_{\beta}\sigma$ in the direction U^{α}.This is the energy-momentum tensor of the Spinor field.

The trace $Tr\left(T\right)$ of the tensor T is the tensor :

$Tr\left(T\right)\left(\sum_{\alpha}U^{\alpha}\partial\xi_{\alpha}\right)=\frac{1}{i}\sum_{\alpha=0}^{3}\mu\left\langle S_0,U^{\alpha}\upsilon\left(X_{r\alpha},X_{w\alpha}\right)\right\rangle$

that is the kinetic energy (up to a constant).

Take $\upsilon\left(X_{r\alpha},X_{w\alpha}\right)=\upsilon\left(0,\delta_{\alpha}w\right)$

$Tr\left(T\right)\left(\sum_{\alpha}U^{\alpha}\partial\xi_{\alpha}\right)=-\frac{1}{2}\mu\epsilon k_0^t\sum_{\alpha=0}^{3}\frac{U^{\alpha}}{c}\delta_{\alpha}w$

can be seen as the pressure of the flow of matter in the spatial direction $\delta_{\alpha}w$.

Spinor field for a deformable solid

One can define, for any observer, a deformable solid by a section $\sigma \in P_G$. The particles travel on trajectories V defined by σ_w with parameter the time of the observer. Adding a density μ, and an inertial spinor S_0, then, because S is valued in the vector space E, the integral : $\int_{\omega(t)} \mu(m) S(m) \varpi_3(t,m)$ where $\omega(t) = \Phi_V(\omega(0), t)$ and $\omega(0)$ is a compact subset of $\Omega_3(0)$ makes sense.

$S(t) = \gamma C\left(\int_{\omega(t)} \sigma(m) \mu(m) \varpi_3(m)\right) S_0$

$\Gamma = \int_{\omega(t)} \sigma(m) \mu(m) \varpi_3(m) \in Cl(3,1)$

We have several cases of interest.

If the solid is rigid : $\sigma(\Phi_V(t,x)) = s(t) \cdot g(\Phi_V(0,x))$ with $s(t) \in Spin(3,1)$. Then $\int_{\omega(t)} \sigma(m) \mu(m) \varpi_3(m) = s(t) \int_{x\in\omega(0)} g(x) \mu((\Phi_V(t,x))) \varpi_3(\Phi_V(t,x))$

and $S(t) = \gamma C(s(t)) S_B(t)$ with $S_B(t) = \gamma C\left(\int_{x\in\omega(0)} g(x) \mu((\Phi_V(t,x))) \varpi_3(\Phi_V(t,x))\right) S_0$.

The variation of $S_B(t)$ can be computed as above :

$S_B(t_2) - S_B(t_1) = \int_{\omega([t_1,t_2])} \gamma C(\pounds_V(g\mu\varpi_4)) S_0 = \int_{\omega([t_1,t_2])} \gamma C\left(\frac{dg\mu}{dt} + g\mu(divV)\right) \varpi_4 S_0$

$= \int_{\omega([t_1,t_2])} \gamma C\left(g\left(\frac{d\mu}{dt} + \mu(divV)\right)\right) \varpi_4 S_0$

With the continuity equation : $S_B(t) = Ct$ and $S(t) = \gamma C(s(t)) S_B$.

Theorem 75 *From a kinematic point of view, a rigid solid can be replaced by a particle moving along one integral curve of the vector field V with spinor $S(t) = \gamma C(s(t)) S_B$ where S_B is the constant inertial spinor.*

$$S_B = \gamma C\left(\int_{x\in\omega(0)} g(x) \mu(x) \varpi_3(x)\right) S_0 \tag{4.17}$$

This is the generalization of the rule of Newtonian Mechanics.

The computation of the integral $\Gamma = \int_{x\in\omega(0)} g(x) \mu(x) \varpi_3(x) \in Cl(3,1)$ can be done in any chart, adjusted for the symmetries of the solid. And if $S_0 \in E_0$ then $S_B \in E_0$. However Γ does not necessarily belong to $Spin(3,1)$.

A rigid solid has an inertial vector defined by S_B :

$k_{0B}^a = \frac{1}{i} \epsilon \frac{1}{M_B^2} \langle S_B, (\gamma_0\gamma_a - i\widetilde{\gamma}_a) S_B \rangle_E$ with $k_{0B}^t k_{0B} = 1$

The Dirac's current $(\gamma_0\gamma_a - i\widetilde{\gamma}_a) S_B$ can be identified with the flow of matter in the 3 spatial directions corresponding to $\gamma_a = \gamma(\varepsilon_a)$.For any rigid solid in Newtonian Mechanics there is an inertial tensor, represented by a symmetric matrix $[J]$ with 3 orthogonal eigen vectors and real eigen values λ_a. So we can identify these vectors to ε_a and $k_{0B}^a = \frac{1}{\sqrt{\sum_{a=1}^3 \lambda_a^2}} \lambda_a$.

Example : For an ellipsoid with mass m and axes of spatial lengths a, b, c

$k_{0B} = \frac{1}{\frac{\sqrt{2}}{5}\sqrt{(a^4+b^4+c^4+a^2b^2+a^2c^2+b^2c^2)}} \left(\frac{b^2+c^2}{5}, \frac{a^2+c^2}{5}, \frac{b^2+a^2}{5}\right)$

In the general case the deformation tensor is $\partial_\alpha \sigma \cdot \sigma^{-1}$. This is a 1 form on M valued in $T_1 Spin(3,1)$.

The stress tensor is then : $\gamma C(\partial_\alpha \sigma) S_0 \otimes d\xi^\alpha = \gamma C\left(\partial_\alpha \sigma \cdot \sigma^{-1}\right) \gamma C(\sigma) S_0 \otimes d\xi^\alpha$. This is a 1 form on M valued in E. On a trajectory $\delta U = \sum_{\alpha=0}^3 \delta U^\alpha \partial\xi_\alpha$ the inertial forces, similar to stress forces, which preserve the integrity of the solid are :

$\delta F = \sum_{\alpha=0}^3 \delta U^\alpha \delta_\alpha S \in E$.

We still have $S(t) = \gamma C(\Gamma(t)) S_0 \in E_0$ if $S_B \in E_0$.

Symmetries

For particles symmetries are symmetries in the motion, and they are essentially periodic symmetries. For periodic instantaneous motions the momentum and the kinetic energy are periodic.

Sections of $P_G[E, \gamma C]$ are defined by a section of P_G and a constant spinor, so they have the same symmetries than sections of P_G.

For deformable solids, defined through $S(\varphi_o(t,\xi)) = \mu(\varphi_o(t,\xi))\gamma C(\sigma(\varphi_o(t,\xi)))S_0$, the continuity equation involves the metric g. If both σ and μ have the same symmetry, the symmetry is preserved over time if $\det g$ is constant. And in a periodic motion it implies that the metric itself is periodic.

Classic Mechanics provides efficient and simpler tools, and the use of spinors would be just pedantic in common problems. However 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 can account for the rotational momentum of the bodies, which is significant and contributes to the total kinetic energy of the system.

To go further in the study of Spinors for elementary particles we need to remind some results about the representations of the groups $Spin(3,1), Spin(3)$.

4.3.5 Representation of the Spin group

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 : $\lambda : G \times E \to E :: \lambda(g,x)$

Define the left action Λ of G on H : $\Lambda : G \times H \to H :: \Lambda(g,\varphi)(x) = \varphi\left(\lambda\left(g^{-1},x\right)\right)$

Thus G acts *on the argument* of φ. Then (H,Λ) is a representation of G.

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) 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_1G where $X \in T_1G$ 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, if the set of polynomials is algebraically closed under the action of the group.

Isomorphisms of groups

Most of the groups encountered in Physics are related to the group $SL(\mathbb{C},2)$ of 2×2 complex matrices with determinant 1 (on these topics Maths.V.24). Its Lie algebra $sl(\mathbb{C},2)$ is the set of 2×2 complex matrices with null trace. They can be written :

$\sigma(Z) = \begin{bmatrix} z_3 & z_1 - iz_2 \\ z_1 + iz_2 & -z_3 \end{bmatrix}$ with $Z = (z_1,z_2,z_3) \in \mathbb{C}^3$

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 :

$\exp\sigma(Z) = I\cosh D + i\frac{\sinh D}{D}\sigma(Z)$ with $D^2 = \det\sigma(Z) = Z^tZ$

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 the set of matrices $\sigma(ir)$ with $r\in\mathbb{R}^3$. The exponential is surjective on $SU(2)$:

$\exp\sigma(ir) = I\cos\sqrt{r^tr} - \frac{\sin\sqrt{r^tr}}{\sqrt{r^tr}}\sigma(ir)$

$T_1Spin(3,1)$ is isomorphic to $sl(\mathbb{C},2)$: $\upsilon(r,w)\to\sigma(r+iw)$ so with the complex structure $T_1Spin(3,1)_C \sim sl(\mathbb{C},2)$ and $T_1Spin(3,1), T_1Spin(1,3)$ are isomorphic.

$Spin(3,1)$ is isomorphic to $SL(\mathbb{C},2)$: $A+Z\to\exp\sigma(Z) = I\cosh D + i\frac{\sinh D}{D}\sigma(Z)$ with $D^2 = Z^tZ = 4\left(1-A^2\right)$ and $Spin(3,1)$ is isomorphic to $Spin(1,3)$

$T_1Spin(3)$ is isomorphic to $su(2)$: $\upsilon(r,0)\to\sigma(r)$ and $so(3) : \upsilon(r,0)\to j(r)$

$Spin(3)$ is isomorphic to $SU(2)$:

$a_r + \upsilon(r,0)\to\exp\sigma(r) = I\cosh\sqrt{r^tr} + i\frac{\sinh\sqrt{r^tr}}{r^tr}\sigma(r)$

Representations of Spin(3,1)

$SL(\mathbb{C},2), Spin(1,3)$ and $Spin(3,1)$ have the same representations.

There is a unique (up to equivalence) *non unitary*, irreducible representation of dimension n, 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).

$(Cl(3,1),\mathbf{Ad}), (Cl(1,3),\mathbf{Ad})$ are 2 non equivalent non unitary representations of real dimension 16, they are reducible : $(T_1Spin(3,1),\mathbf{Ad})$ is a 6 dimensional irreducible representation, isomorphic to $\left(P^1\otimes P^1, D_1\times D_1\right)$.

The only known unitary representations are over spaces of complex functions : they are infinite dimensional and each *irreducible* representation is parametrized by 2 scalars $z\in\mathbb{Z}, k\in\mathbb{R}$.

Representations of the group $Spin(3)$

$SU(2)$ as $Spin(3)$ are compact groups, so their unitary representations are reducible in a sum of orthogonal, finite dimensional, unitary representations. The usual 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 $-j\le m\le +j$. 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 irreducible, unitary, representations of $SO(3)$ are then given by $\left(P^j, D^j\right)$ with j integer.

We have seen that $(T_1Spin(3)_C,\mathbf{Ad})$ is a unitary representation of $Spin(3)$. It is reducible : the real and imaginary part, or equivalently the vector subspaces L_0, P_0 are invariant. $(L_0,\mathbf{Ad}), (P_0,\mathbf{Ad})$

are 3 dimensional unitary representations of $Spin(3)$, parametrized by the choice of vector ε_0. They are orthogonal and equivalent. So they are isomorphic to $\left(P^1, D^1\right)$, and representations of $SO(3)$.

4.3.6 Quantization of the Spinor

Quantization

The vector space E is normed, and E_ϵ are Hilbert spaces. In a model involving a particle the spinor is represented by a map : $J^1S : [0,T] \subset \mathbb{R} \to J^1P_G[E, \gamma C]$ for some $S_0 \in E_0$. The map is assumed to be such that $\int_0^T \max\left(\|S(t)\|, \|\delta S(t)\|\right) dt < \infty$ then it belongs to a separable Fréchet space F and the theorems of the 2nd Chapter apply.

With the action :

$\lambda : Spin(3,1) \times F \to F :: \lambda(g,S)(t) = \gamma C(g) S(t)$

(F, λ) is a representation of $Spin(3,1)$. An observable of S is an irreducible representation, characterized by 2 scalars $a \in \mathbb{R}, z \in \mathbb{Z}$.

Each map depends on the kinematic characteristics of the particle. We can assume that a is the mass. Then there is a countable number of possible values for the inertial vector k, which can be labeled by z.

The spin is represented by $\upsilon(X_r(t), 0) \in T_1Spin(3)$ which is globally invariant by $Spin(3)$. Then an observable of the spatial spinor $S_r(t) = \gamma C(\upsilon(X_r(t), 0)) S(t)$ corresponding to the rotational momentum belongs to an irreducible representation of $Spin(3)$, and is characterized by some $j \in \frac{1}{2}\mathbb{N}$. The change $S_r(t) \to -S_r(t)$ is a discontinuous operation.

And we can state :

Theorem 76 *An observable of the momentum of a particle is characterized by the mass and a scalar* $z \in \mathbb{Z}$.

An observable of the rotational momentum of a particle is characterized by a scalar $j \in \frac{1}{2}\mathbb{N}$..

Periodic states

We have seen that a periodic motion can be represented by a map :

$\sigma : \mathbb{R} \to Spin(3,1) :: \sigma(t) = A(t) + Z(t)$ where $Z(t+T) = Z(t)$ for some fixed period

with :

$Z(t) = \sum_{n\in\mathbb{Z}} \widehat{Z}(n) \exp in\omega t$ with $\widehat{Z}(n) = \frac{1}{T}\int_0^T Z(t) \exp(-in\omega t)\, dt$ and $\omega = \frac{2\pi}{T}$

$Z(0) = \sum_{n\in\mathbb{Z}} \widehat{S}(n)$

$A(t) = \sum_{n\in\mathbb{Z}} \widehat{A}(n) \exp in\omega t$ with $\widehat{A}(n) = \frac{1}{T}\int_0^T A(t) \exp(-in\omega t)\, dt$

$A(t)^2 = 1 - \frac{1}{4} Z(t)^t Z(t)$

The spinor is then :

$S(t) = \gamma C(\sigma(t)) S_0 = \sum_{n\in\mathbb{Z}} \widehat{S}(n) \exp in\omega t$ with $\widehat{S}(n) = \frac{1}{T}\int_0^T S(t) \exp(-in\omega t)\, dt$

$\widehat{S}(n) = \gamma C\left(\widehat{A}(n) + \widehat{Z}(n)\right) S_0$

By derivation :

$\frac{dS}{dt} = \sum_{n\in\mathbb{Z}} in\omega \widehat{S}(n) \exp in\omega t$

we have necessarily the relation :

$\widehat{\frac{dS}{dt}}(n) = in\omega \widehat{S}(n)$

and $\frac{dS}{dt}|_{t=0} = \sum_{n\in\mathbb{Z}} in\omega \widehat{S}(n)$

The average energy on the trajectory is : $\frac{1}{M_p}\frac{1}{T}\int_0^T \frac{1}{i}\left\langle S(t), \frac{d}{dt}S(t)\right\rangle dt$

The variables belong to a Hilbert space H with scalar product :

$\langle Y_1, Y_2\rangle_H = \frac{1}{T}\int_0^T \langle Y_1(t), Y_2(t)\rangle_E\, dt = \sum_{n\in Z} \left\langle \widehat{Y_1}(n), \widehat{Y_2}(n)\right\rangle_E$

Thus :

$\frac{1}{T}\int_0^T \frac{1}{i}\left\langle S(t), \frac{d}{dt}S(t)\right\rangle dt = \sum_{n\in Z}\left\langle \widehat{S}(n), \widehat{\frac{dS}{dt}}(n)\right\rangle = \sum_{n\in Z} in\omega \left\langle \widehat{S}(n), \widehat{S}(n)\right\rangle$

$\left\langle \widehat{S}(n), \widehat{S}(n)\right\rangle$ can be computed with : $S_0 = \begin{bmatrix} S_R \\ \epsilon i S_R \end{bmatrix}$ and one gets :

$\left\langle \widehat{S}(n), \widehat{S}(n)\right\rangle = M_p^2\left(\left(\operatorname{Re}\widehat{A}(n)\right)^2 - \left(\operatorname{Im}\widehat{A}(n)\right)^2 + \frac{1}{4}\left(\left(\operatorname{Re}\widehat{Z}(n)^t\widehat{Z}(n)\right)^2 - \left(\operatorname{Im}\widehat{Z}(n)^t\widehat{Z}(n)\right)^2\right)\right)$

$\frac{dS}{dt}|_{t=0} = \sum_{n\in\mathbb{Z}} in\omega\widehat{S}(n) \Rightarrow \sum_{n\in Z}\left\langle in\omega\widehat{S}(n), in\omega\widehat{S}(n)\right\rangle = \omega^2\sum_{n\in Z} n^2\left\langle \widehat{S}(n), \widehat{S}(n)\right\rangle < \infty$

$\Rightarrow \sum_{n\in Z} n\left\langle \widehat{S}(n), \widehat{S}(n)\right\rangle S_R^* S_R < \infty$

$\frac{1}{M_p}\frac{1}{T}\int_0^T \frac{1}{i}\left\langle S(t), \frac{d}{dt}S(t)\right\rangle dt$

$= \omega M_p \sum_{n\in Z} n\left(\left(\operatorname{Re}\widehat{A}(n)\right)^2 - \left(\operatorname{Im}\widehat{A}(n)\right)^2 + \frac{1}{4}\left(\left(\operatorname{Re}\widehat{Z}(n)^t\widehat{Z}(n)\right)^2 - \left(\operatorname{Im}\widehat{Z}(n)^t\widehat{Z}(n)\right)^2\right)\right)$

Theorem 77 *In a periodic motion of a particle the average kinetic energy is proportional to the frequency.*

From the previous theorem the trajectories are characterized by the mass and an integer $z \in Z$. Moreover, from the theorem 24, if we add the energy as variable each irreducible representation belongs to a class of solutions which gives the same value to the average energy. From there we can conclude that the frequencies are quantized : there is only a countable number of observable frequencies in the periodic state of a particle, and each one corresponds to a level of energy.

4.3.7 Spinors for elementary particles

Particles and Anti-particles

The inertial spinor is a starting point in the identification of "elementary particles", that is the ultimate constituent of matter.

The first natural requisite is that $S_0 \in E_0$. The value of ϵ is related to a choice of a basis of E_ϵ. In the usual cases ϵ is purely conventional. However for elementary particles it is an issue because, for a given value of the mass, there are a countable set of possible values for k_0. The relation between S_0 and k_0 is not linear and we cannot expect to find vector subspaces of elementary particles, but the basis of E matters and one cannot discard ϵ.

The logical explanation is that the value of ϵ distinguishes particles and antiparticles. The mass is $M_p^2 = \epsilon 2\left[S_R\right]^*\left[S_R\right]$. 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 $\langle S_0, S_0\rangle = \epsilon 2 M_p^2$ so we can choose any sign for M_p, and it seems more appropriate to take $M_p > 0$ both for particles and antiparticles.

The inertial spinor of particles is then :

$$S_0 = \frac{M_p}{\sqrt{2}}\begin{bmatrix} e^{i\alpha_1}\cos\alpha_0 \\ e^{i\alpha_2}\sin\alpha_0 \\ ie^{i\alpha_1}\cos\alpha_0 \\ ie^{i\alpha_2}\sin\alpha_0 \end{bmatrix}$$

and of antiparticles :

$$S_0 = \frac{M_p}{\sqrt{2}}\begin{bmatrix} e^{i\alpha_1}\cos\alpha_0 \\ e^{i\alpha_2}\sin\alpha_0 \\ -ie^{i\alpha_1}\cos\alpha_0 \\ -ie^{i\alpha_2}\sin\alpha_0 \end{bmatrix}$$

It is characterized by 4 parameters : $M_p, \alpha_0, \alpha_1, \alpha_2$.

Chirality

In the Spinor representation 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 elementary particles : chirality. The representation (E, γ) has been chosen because of this property. If the real Clifford algebras leave invariant E_0, some of their elements exchange E_ϵ and $E_{-\epsilon}$.

$S_0 \in E_\epsilon \Rightarrow \gamma_0 C(\varepsilon_0) S_0 \in E_\epsilon, j = 1, 2, 3 : \gamma C(\varepsilon_j) S_0 \in E_{-\epsilon}$ with the same property in $Cl(1, 3)$.

So E_ϵ is preserved by $X \in T_1 Spin(3), \sigma \in Spin(3)$.

Space reversal is the operation :

$u = u^0\varepsilon_0 + u^1\varepsilon_1 + u^2\varepsilon_2 + u^3\varepsilon_3 \rightarrow u^0\varepsilon_0 - u^1\varepsilon_1 - u^2\varepsilon_2 - u^3\varepsilon_3$

corresponding to $s = \varepsilon_0, s^{-1} = -\varepsilon_0$ in $Cl(3, 1)$, $s^{-1} = \varepsilon_0$ in $Cl(1, 3)$ so it preserves E_ϵ.

Time reversal is the operation :

$u = u^0\varepsilon_0 + u^1\varepsilon_1 + u^2\varepsilon_2 + u^3\varepsilon_3 \rightarrow -u^0\varepsilon_0 + u^1\varepsilon_1 + u^2\varepsilon_2 + u^3\varepsilon_3$

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 $Cl(1, 3)$ so it exchanges E_ϵ and $E_{-\epsilon}$.

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.

Inertial vector

The inertial vector is : $k = -i\epsilon \frac{M_p}{2} k_0 = -i\epsilon \frac{M_p}{2} \begin{bmatrix} (\sin 2\alpha_0)\cos(\alpha_1 - \alpha_2) \\ -(\sin 2\alpha_0)\sin(\alpha_1 - \alpha_2) \\ \cos 2\alpha_0 \end{bmatrix}$. Particles and antiparticles with the same parameters $M_p, \alpha_0, \alpha_1, \alpha_2$ have opposite inertial vectors, and so opposite momenta and kinematic behaviors.

Particles whose inertial vectors differ by a complex scalar of module 1 have the same kinematic behavior. This is the starting point for the idea of rays in QM.

Spin

$Spin(3)$ preserves E_ϵ, then $(E_\epsilon, \gamma C), (E_\epsilon, \gamma C')$ are representations of $Spin(3)$. Moreover the scalar product is definite positive or negative and preserved by $Spin(3)$ so we have unitary representations, which are isomorphic to one of the representations (P^j, D^j) with $j \in \frac{1}{2}\mathbb{N}$. Actually, for elementary particles $j = \frac{1}{2}$ and this is the origin of the name “particles of spin $\frac{1}{2}$”.

Because the spatial spin is quantized, the rotational motion is itself quantized. The natural representation is then by a periodic motion : the particle spins at a constant rotational speed. The average rotational kinetic energy is proportional to the frequency. The speed does not change, but the axis of rotation can change (by the action of $Spin(3)$). Moreover the spin can take the opposite value, corresponding to $\upsilon(X_r, 0) \rightarrow \upsilon(-X_r, 0)$. This is a discontinuous process (because the spin is quantized, it cannot take intermediate values) which requires an external action and entails a change of kinetic energy.

The variation of kinetic energy is $\frac{dK}{dt} = C_I \frac{1}{M_p} \operatorname{Im} k^t \mathbf{Ad}_{\sigma^{-1}} \upsilon(X_r, X_w)$ with a universal constant C_I which has the dimension of energy. X_r, X_w have the unit $[T]^{-1}$ and we have seen the conventions for the measure of rotations of solids. However this formula provides another point of view : $C_I X_r$ represents the variation of rotational energy, expressed in units $[E][T]^{-1}$ or J/s. In the relation $E = h\nu$ of Quantum Physics the Plank’s constant h is expressed in Js and the frequency ν in Hz, or cycles / s. So, for atomic or subatomic particles $C_I X_r$ can be measured as a multiple of hX_r. Their rotational motion is then measured in rad/s or in cycle / s that in Hz, in concordance with our representation.

To each particle corresponds an antiparticle with the same mass. And particles show polarization characteristics similar to waves. The picture is similar to the Dirac's spinors, with different definitions of the γ matrices.

Charge

Assume that we study a system comprising 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.

What the quantization theorem 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 k, their inertial vector, but they belong also to classes of maps characterized by $z \in \mathbb{Z}$. One can assume that the signed integer z is related to a charge. But we see that any particle which has the same inertial vector k belongs to a definite class characterized by the same z : these particles have the same behavior in a field. This is the starting point for the representation of charged particles and we can guess that the inertial vector is more than a kinematic feature.

4.3.8 Composite particles and Atoms

Representation by tensorial products

Stable combinations of elementary particles are represented by the tensorial product of the spinors, as composite system, following the theorem 29 of QM. Then the motion is represented in the universal enveloping algebra U of $T_1Spin(3,1)$. This is a vector space, built from *tensorial powers* of the Lie algebra $T_1Spin(3,1)$, such that the elements of the form : $X \otimes Y - Y \otimes X - [X, Y] \sim 0$. A basis of U consists of the ordered tensorial products of vectors of a basis of the Lie algebra. That is for $T_1Spin(3,1)$: 1 and the tensorial products $\overrightarrow{\kappa}_{\alpha_1} \otimes \overrightarrow{\kappa}_{\alpha_2} \otimes ... \otimes \overrightarrow{\kappa}_{\alpha_n}$ with $\alpha_1 \leq \alpha_2 \leq ... \leq \alpha_n$.

Any representation (E, f) of the Lie algebra can be extended to a representation (E, F) of its universal enveloping algebra where the action is :

$F\left(\kappa_{i_1}^{n_1}...\kappa_{i_p}^{n_p}\right) = f(\kappa_{i_1})^{n_1} \circ ... \circ f(\kappa_{i_p})^{n_p}$

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 order $n_1 + n_2 + ... + n_p$.

In the representation of $T_1Spin(3,1)$ by matrices of $so(3,1)$ the universal enveloping algebra is actually an algebra of matrices.

Casimir element

The **Casimir element** is a special element Ω of U, defined through the Killing form. 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 : $F(\Omega)\varphi(x) = D_2\varphi(x) = k\varphi(x)$: φ is an eigen vector of D_2. As a consequence, if there is a scalar product on E : $\langle F(\Omega)u, F(\Omega)u\rangle = \langle ku, ku\rangle = k^2\langle u, u\rangle$. And k has the same value in all equivalent representations.

The Killing form on $T_1Spin(3,1)$ is :

$B(v(r,w), v(r',w')) = 4(w^tw' - r^tr')$

We have an orthonormal basis for B with the elements

$\kappa_1 = -\frac{1}{8}\varepsilon_3 \cdot \varepsilon_2, \kappa_2 = -\frac{1}{8}\varepsilon_1 \cdot \varepsilon_3, \kappa_3 = -\frac{1}{8}\varepsilon_2 \cdot \varepsilon_1,$

$\kappa_4 = \frac{1}{8}\varepsilon_0 \cdot \varepsilon_1, \kappa_5 = \frac{1}{8}\varepsilon_0 \cdot \varepsilon_2, \kappa_6 = \frac{1}{8}\varepsilon_0 \cdot \varepsilon_3$

and the Casimir element of $U(T_1Spin(3,1))$ is :

$\Omega = \sum_{i=4}^{6} \kappa_i \otimes \kappa_i - \sum_{i=1}^{3} \kappa_i \otimes \kappa_i$

The action of the Casimir element in the representation $(E, \gamma C)$ of $Spin(3,1)$ is :

$\Gamma(\Omega) u = \left(\sum_{i=4}^{6} (\gamma C(\kappa_i))^2 - \sum_{i=1}^{3} (\gamma C(\kappa_i))^2 \right) u = \frac{3}{2} u$

In the representation (P^j, d^j) of $T_1 Spin(3)$, if we denote $L_x = f(\kappa_1), L_y = f(\kappa_2), L_z = f(\kappa_3)$ with 3 arbitrary orthogonal axes :

$F(\Omega) |j,m\rangle = L^2 |j,m\rangle = (L_x^2 + L_y^2 + L_z^2) |j,m\rangle = j(j+1) |j,m\rangle$

$d^j(\kappa_i) \left(\sum_{m=-j}^{m=+j} X^m |j,m> \right) = \sum_{m=-j}^{m=+j} X^m d^j(\kappa_i) |j,m>$

Measure of the spatial spin of a particle

An observable of the spatial spinor $S_r(t)$ belongs to a vector space of maps isomorphic to some (P^j, D^j) : $S_r(t) = \sum_{p=-j}^{+j} S_r^p |j,p>$ where S_r^p are fixed scalars and $|j,p>$ are, for a given system, fixed maps $|j,p>: \Omega \rightarrow E_0$, 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. Under the action of $Spin(3)$ the vectors $S_r(m)$ transform according to the same matrices as in D^j.

There is one important difference in the behavior of the spin, according to the value of j. $Spin(3)$ 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 actions of $+s$ and $-s$ give opposite results. The representations (P^j, D^j) 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.

The measure 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 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.

The action of the device can be modelled as an operator $L(\rho)$ acting on the space of vectors $|j,p>$, 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 :

$L(\rho)(S_r) = \sum_{p=-j}^{+j} S_r^p [d_j(\rho)] |j,p>$

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, and 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 $L^2(S_r) = (L_x^2 + L_y^2 + L_z^2)(S_r) = j(j+1)(S_r)$

Atoms

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.

In our framework we can build a model of an Atom, comprising a nucleus, and N electrons. We will not look at the physical laws which allow such a system to be stable, but focus on the representation of the spinors of the electrons. In an atom they stay together, close to the nucleus (at the atomic scale). Their trajectories $q_p(t)$ do not cross, so the natural model is to assume that each electron $p = 1...N$ is represented by a map : $S_p : \mathbb{R} \rightarrow E_+$ with the same inertial spinor $S_0 : S_p(t) = \gamma C(\sigma_p(t)) S_0$, and that $\sigma_p(t) = \sigma(q_p(t))$ for a common section $\sigma \in P_G$. It is reasonable to assume that their motion is periodic. We have seen (Periodic Motion, chap.3) how to build σ from :

- a periodic map $w : \mathbb{R} \rightarrow \mathbb{R}^3 :: w(t+T) = w(t)$, the trajectories are then $q : \mathbb{R} \rightarrow M :: q(t) = \varphi_o(ct, x(t))$ with : $\left[\frac{dx}{dt}\right] = \epsilon \frac{\sqrt{1+\frac{1}{4}w^t w}}{1+\frac{1}{2}w^t w} w(t)$, they are integral curves of the section $s(\varphi_o(ct, x(t))) = a_w + \upsilon(0, w(t))$ passing through the point $x(0)$

- a periodic map : $r : \mathbb{R}^4 \rightarrow \mathbb{R}^3 :: r(t + T_e(\xi), x) = r(t, \xi)$ which defines the section $\sigma_r(\varphi_o(ct, \xi)) = a_r + \upsilon(r(t, \xi), 0)$. The period $T_e(\xi)$ is itself given by a map : $T_e : \Omega_3(0) \rightarrow \mathbb{R}$. Thus it depends on the initial location $\varphi_\Omega(\xi) = x \in \Omega_3(0)$. For instance we could assume that it depends on the distance from the nucleus.

Then the section $\sigma = s \cdot \sigma_r$ has the path $q(t) = \varphi_o(ct, \xi(t)) = \Phi_V(t, x(0))$ as integral curves, and for each $\varphi_o(ct, \xi(0))$ the instantaneous motion is periodic with period $T(\xi(0))$ depending on $x(0)$.

What is of a greater interest is the set of electrons, the electronic cloud. It constitutes a system in itself, and most of the properties of the atom (notably in Chemistry) comes from its behavior. It is represented by the maps w, r, which do not depend on the electron, and the initial location $a_p = x_p(0)$ of each electron. The state of an electron is an occurrence of the same map $S_p(t) = \gamma C(\sigma(q_p(t))) S_0$ identified by $x_p(0)$. We can apply the Theorems 32, 34. The states of the electronic cloud are then represented by the tensorial product $S_1 \otimes S_2 ... \otimes S_N = S(a_1) \otimes S(a_2) ... \otimes S(a_N)$ Electrons have a Spin 1/2, so the representation is by antisymmetric tensor, to account for the action of $Spin(3,1)$ on the spin. The stable states are then characterized by a class of conjugacy λ represented by :

- a decomposition λ of the integer N in p parts : $\lambda(N) = \{0 \leq n_1 \leq n_2 ... \leq n_p; n_1 + n_2 + ..n_p = N\}$.

- a set of p distinct vectors $\psi_1, \psi_2, ...\psi_p$ of a hermitian basis of the space of maps $S(t)$, which together define a vector subspace H_J.

- the states of the system are then represented by antisymmetric tensors belonging to $\wedge_{n_1} H_J \wedge ... \wedge_{n_p} H_J$

Each subset of n_k elements is an "electron shell", which is then characterized by n_k different states of the electrons. Two electrons which differ only by their spin are deemed to be in different states. Of course the integral curves to which belong the electrons is one of their features, and each shell corresponds to an integral curve, with the same period. We have seen that, for periodic motion, the average kinetic energy is proportional to the frequency, so the shells correspond to different levels of energy.

Additional observables must be added to differentiate the shells. The spin of the electrons can be represented in a finite dimensional vector space isomorphic to $P^{1/2}$. Tensorial products of representations (P^j, D^j) can be combined using Clebsch-Jordan coefficients. 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.V.24). It is then assumed that the arguments of the function $f(x)$ are related to the coordinates (in an euclidean frame) of the electron.

By adding a section to represent the nucleus we have a spinor representing the atom like a rigid solid. One can compute a total spinor for the atom. The map w is then the proper rotation of the atom, which can itself be represented by a spin.

Chapter 5

FORCE FIELDS

The concept of fields has appeared in the XIX° century, with the electromagnetism theory, to replace the picture of action at a distance between particles. Gravitation, whose laws were known since Kepler and Newton, can be easily fitted in the same framework. However it appeared that Maxwell's equations were not compatible with Galilean Geometry, and this was at the origin of Einstein's Special Relativity. All together they provide a consistent, well checked and efficient theory. New phenomena, essentially occurring in discontinuous processes, lead to come back to a corpuscular interpretation of the EM field with the photon, and it appeared that the nuclear forces could themselves be represented as new force fields with corresponding charges. This extension was made possible with a new mathematical tool : connection and gauge theories. From a totally different point of view Einstein proposed a new Theory of Gravitation, based on the Geometry of General Relativity. It if can be expressed in the framework of gauge theory, as we will show, it relies on original physical assumptions. The facts that the strength of the gravitational field is 10^{-49} weaker than the EM field (a fact which by itself needs to be explained), but has additive effects which are felt at very long range, make that, for any experimental and practical purpose, a Theory of Gravitation will always be special. But the same could be said about the nuclear forces. The concept of fields, and the mathematical apparatus of gauge theories, can be seen as somewhat artificial, and the goal of a Great Unification Theory a futile dream. But, if we want, one day, master the different forces that exist, as well as we do for the EM field, it does not seem to be another path.

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 field with itself.

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 with, for each one, 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 :

$Y = 2(Q - T_3)$

The total sum of weak isospin is conserved in interactions.

Each fermion (as it seems also true for the neutrinos) has a mass and 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 group together to form a meson (a quark and an 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 transform itself into another one, it can also disintegrate in other particles, and conversely particles can be created in discontinuous

[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.

processes, 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 cases 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 $SU(3) \times SU(2) \times U(1)$ 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 combinations 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 basis of the Lie algebra 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 $\frac{1}{r}\exp(-kmr)$ 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 $SU(3)$.

All together they are part of $Cl(\mathbb{R},10)$.

In the next sections we will see how the states of particles, force fields and their interactions can be represented, in the geometrical context of GR. They involve essentially the potentials, which are a component of the connection. The other feature of fields is their propagation, assumed to come from their interaction with themselves. This is a phenomenon which deserves a study by itself. Its model is based on a derivative of the connection, the strength of the field $\mathcal{F}$, and has special characteristics which are the topic of separate sections.

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.

We will include gravitation in this representation. This is natural : since the introduction of the concept of force field, gravitation was acknowledged as one of them. The gauge field theory provides a mathematical framework, which can easily address gravitation in the Geometry of General Relativity and is compatible with the Classical Theory. But the Eintein's General Relativity encompasses both a Geometrical Theory - that we have seen in the previous chapters - and a Theory of Gravitation which is based on a different point of view - gravitation is not a force field per se, but its effects result from the metric (the "curved space-time"). However, as it is necessary to formalize these effects in the usual framework of forces and action, Einstein's Theory uses mathematical tools such as connection and covariant derivative, which can be seen as special cases of a general gauge theory of gravitation. We will come back to these fundamental issues later.

5.2 STATES OF PARTICLES

Spinor fields can be characterized, beyond the inertial spinor, by a signed 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 elementary 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_ϵ : 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.

In the previous chapters we assumed that :

the arrangement and motion of a particle is represented in $J^1P_G : (m, \sigma, \upsilon(X_r, X_w))$

the momenta are represented in $J^1P_G[E, \gamma C]$

there is a differential operator :

$\mathcal{M} : J^1P_G \to J^1P_G[E, \gamma C] :: \mathcal{M}(m, \sigma, \delta\sigma) = (m, \gamma C(S_0), \gamma C(\upsilon(X_r, X_w)) S,)$

The relation $\sigma \to \upsilon(X_r, X_w)$ is the motion

The relation $S \to \gamma C(\upsilon(X_r, X_w)) S$ represents the action of the inertial forces

The principle of equivalence tells that the action of the gravitational forces can be represented as the action of the inertial forces : through the fiber bundle $(E, \gamma C)$ so the inertial spinor S_0 is the gravitational charge.

Proposition 78 *The Gravitational charge of a particle is represented by its inertial spinor 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 $\psi_{0p} = S_{0p}$.

Another way to see the principle of equivalence is to take an observer who is attached to a material body (say the Earth). His chart, by definition, is fixed (say the direction to distant stars), as well as the holonomic basis $(\partial\xi_\alpha)_{\alpha=0}^{3}$. However he can choose a tetrad attached to the material body (say a fixed orthonormal frame), and he can measure the change of the tetrad with respect to the holonomic basis (the rotation of Earth). The standard gauge is defined through the tetrad, thus the spinor is fixed with respect to this observer, however with respect to a tetrad which has fixed components in the holonomic basis its value changes and inertial forces appear and can be measured (Foucault's pendulum).

The gravitational charge is then represented by 5 scalars (with the inertial vector) and not just the mass : this is the consequence of the attachment of a tetrad to the particle, and it entails that the action of the gravitational field is more complicated than what is commonly seen.

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 are 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 79 *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 elementary particles are vectors of the tensorial product $E \otimes F$

The intrinsic characteristics of each type of elementary particles are represented by a fixed tensor $\psi_0 \in E \otimes F$,that we call a ***fundamental state****, and all particles sharing the same fundamental state behave identically under the actions of all the fields.*

Notation 80 *$(f_i)_{i=1}^n$ 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 :

$$\psi=\sum_{i=1}^{4}\sum_{j=1}^{n}[\psi]_j^i\, e_i\otimes f_j \tag{5.2}$$

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 :
$\vartheta\,(\sigma,\varkappa)\,(\psi)=\sum_{i,k=1}^{4}\sum_{j,l=1}^{n}[\gamma C\,(\sigma)]_k^i\,[\varrho\,(\varkappa)]_j^l\,[\psi]_l^k\,e_i\otimes f_j$
$=\sum_{i,k=1}^{4}\sum_{j,l=1}^{n}[\gamma C\,(\sigma)]_k^i\,[\psi]_l^k\,[\varrho\,(\varkappa)]_j^l\,e_i\otimes f_j$
that we will denote in matrices :

Notation 81

$$[\vartheta:Spin\,(3,1)\times U\to\mathcal{L}\,(E\otimes F;E\otimes F)::\vartheta\,(\sigma,\varkappa)\,[\psi]=[\gamma C\,(\sigma)]\,[\psi]\,[\varrho\,(\varkappa)]] \tag{5.3}$$

One can extend the action of the Spin group to the action of the Clifford algebra. We define the action ϑ of $Cl\,(3,1)\times U$ on $E\otimes F$ by the unique linear extension of :
$\vartheta:Cl\,(3,1)\times U\to\mathcal{L}\,(E\otimes F;E\otimes F)::$
$\vartheta\,(s,g)\,(S\otimes\phi)=\gamma C\,(s)\,(S)\otimes\varrho\,(g)\,(\phi)$
to all tensors on $E\otimes F$

This is a morphism from $Cl\,(3,1)$ 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 $Cl\,(3,1)\times U$ 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.

$$\langle\psi,\psi'\rangle=Tr\left([\psi]^*\,[\gamma_0]\,[\psi']\right) \tag{5.4}$$

Theorem 82 *ϑ preserves the scalar product on $E\otimes F$: $\langle\vartheta\,(\sigma,\varkappa)\,\psi,\vartheta\,(\sigma,\varkappa)\,\psi'\rangle=\langle\psi,\psi'\rangle$*

Proof. $\widetilde{\psi}^{ij}=\sum_{k=1}^{4}\sum_{l=1}^{n}[\gamma C(\sigma)]_{k}^{i}[\varrho(\varkappa)]_{l}^{j}\psi^{kl}$

$\left\langle\widetilde{\psi},\widetilde{\psi}'\right\rangle=\sum[\gamma_0]_k^i\overline{[\gamma C(\sigma)]}_p^i\overline{[\varrho(\varkappa)]}_q^j\overline{\psi}^{pq}[\gamma C(\sigma)]_r^k[\varrho(\varkappa)]_s^j\psi^{rs}$

$=\sum\left([\gamma C(\sigma)]^*[\gamma_0][\gamma C(\sigma)]\right)_r^p\left([\varrho(\varkappa)]^*[\varrho(\varkappa)]\right)_s^q\overline{\psi}^{pq}\psi'^{rs}$

$=\sum[\gamma_0]_r^p\overline{\psi}^{pq}\psi'^{rq}$ ■

The scalar product is definite, positive on $E_+\otimes F$, negative on $E_-\otimes F$, but not 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 $\|e_i\otimes f_j\|=\|e_i\|_E\|f_j\|_F$. So that $E\otimes F$ is a Banach vector space.

Particles and antiparticles

We will 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.

We discriminate particles and antiparticles by looking 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 83 *The only vector subspaces of $E\otimes F$ which meet these conditions are such that $\psi_L=\epsilon i\psi_R$ with $\epsilon=\pm1$*

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 : $[\psi]=[S][F]^t=\begin{bmatrix}S_RF^t\\ S_LF^t\end{bmatrix}$

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 : $S_L=\epsilon iS_R$ 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 :

$T=\begin{bmatrix}0&i\sigma_0\\ i\sigma_0&0\end{bmatrix}$ 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 84 *The fundamental states ψ_0 of elementary particles (fermions) are such that :*
$\psi_L = i\psi_R$ for particles, their mass M_p is such that
$\langle\psi_0,\psi_0\rangle = 2Tr\left(\psi_R^*\psi_R\right) = M_p^2$
$\psi_L = -i\psi_R$ for antiparticles, their mass is
$\langle\psi_0,\psi_0\rangle = -2Tr\left(\psi_R^*\psi_R\right) = -M_p^2$
To each fermion is associated an antiparticle which has the same mass

$$M_p = \sqrt{\epsilon\left\langle\psi_0,\psi_0\right\rangle} = \sqrt{\epsilon 2Tr\left(\psi_R^*\psi_R\right)} \tag{5.5}$$

With : $\vartheta(\sigma,\varkappa)[\psi] = [\gamma C(\sigma)][\psi][\varrho(\varkappa)]$

$[\psi_0]=\begin{bmatrix}\psi_R\\ \epsilon i\psi_R\end{bmatrix}\Rightarrow\vartheta(1,\varkappa)[\psi_0]=[\psi_0][\varrho(\varkappa)]=\begin{bmatrix}[\psi_R][\varrho(\varkappa)]\\ \epsilon i[\psi_R][\varrho(\varkappa)]\end{bmatrix}$

so the relation does not depend on $\varkappa$. This fact is of interest because it shows that the distinction particle / antiparticle is not related to the field forces, but only to the intrinsic kinematic characteristics of the particles.

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 :

$(E_\epsilon\otimes F)(\psi_0) = \{\vartheta(\sigma,\varkappa)\psi_0, \sigma\in Spin(3,1), \varkappa\in U\}$ for a fixed ψ_0 such that $\psi_L=\epsilon i\psi_R$

But these sets are not vector spaces. $\gamma C(\sigma_r)$ preserves E_ϵ,and similarly the chiral relation $\psi_L = \epsilon i\psi_R$.

Physical states of elementary particles

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 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 measure 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 85 *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 $\vartheta(\sigma,\varkappa)\psi=\vartheta(\sigma',\varkappa')\psi\Leftrightarrow(\sigma,\varkappa)=(\sigma',\varkappa')$

We had seen that this is the case for the spinor. This is extended to the states of particles.

The orbits are not vector subspaces :

Theorem 86 *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. ■

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.

The operators P, T act on the spinor part :

$P : \begin{bmatrix} 0 & \sigma_0 \\ -\sigma_0 & 0 \end{bmatrix} \begin{bmatrix} \psi_R \\ \epsilon i \psi_R \end{bmatrix}$

$T : i \begin{bmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{bmatrix} \begin{bmatrix} \psi_R \\ \epsilon i \psi_R \end{bmatrix}$

$PT : i \begin{bmatrix} -\sigma_0 & 0 \\ 0 & \sigma_0 \end{bmatrix}$

$PT\psi = \begin{bmatrix} -i\psi_R \\ \epsilon \psi_R \end{bmatrix}$

The operator $[C]$ acts on the charge part of the tensor. If we rank the vectors of the basis of F such that the $n/2$ first correspond to a "positive" charge and the last $n/2$ correspond to the opposite charge, for each vector, one can write :

$[\psi] = \begin{bmatrix} \psi_{R+} & \psi_{R-} \\ \epsilon i \psi_{R+} & \epsilon i \psi_{R-} \end{bmatrix}$

Then the action of $[C]$ is :

$$\begin{bmatrix} \psi_{R+} & \psi_{R-} \\ \epsilon i \psi_{R+} & \epsilon i \psi_{R-} \end{bmatrix}_{4\times m} \begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix}_{m\times m} = \begin{bmatrix} \psi_{R+}C_1 + \psi_{R-}C_3 & \psi_{R+}C_2 + \psi_{R-}C_4 \\ \epsilon i \psi_{R+}C_1 + \epsilon i \psi_{R-}C_3 & \epsilon i \psi_{R+}C_2 + \epsilon i \psi_{R-}C_4 \end{bmatrix}$$

and we have for $PT\psi$

$$\begin{bmatrix} -i\psi_{R+}C_1 - i\psi_{R-}C_3 & -i\psi_{R+}C_2 - i\psi_{R-}C_4 \\ \epsilon \psi_{R+}C_1 + \epsilon \psi_{R-}C_3 & \epsilon \psi_{R+}C_2 + \epsilon \psi_{R-}C_4 \end{bmatrix} = \begin{bmatrix} \psi_{R-} & \psi_{R+} \\ \epsilon i \psi_{R-} & \epsilon i \psi_{R+} \end{bmatrix}$$

and one deduces :

$[C] = \begin{bmatrix} 0 & iI \\ iI & 0 \end{bmatrix}$

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 fiber bundle representation

The action ϑ of the groups gives the value of ψ for any fundamental state ψ_0 :

$\psi : (E \otimes F) \times (Spin(3,1) \times U) \to \psi = \vartheta(\sigma, \varkappa)\psi_0$

Formally the action of $Spin(3,1)$ and of U are similar, but they have a different physical meaning.

σ represents the arrangement of the tetrad of the particle with respect to the tetrad of the observer. It changes if the observer changes (or if the observer changes his tetrad), but changes also with the motion. So σ represents a physical quantity.

The action of U is related to the choice of a gauge by the observer. The charge is measured by comparing the behavior of the particle to the behavior of known particles. The charges correspond to different vectors of the basis of F. They can be labelled differently, their physical properties do not change, but their representation changes. However it is assumed that the charge itself does not change, notably with the motion : indeed it would mean a change of particle. The observer is assumed to use a standard gauge : $\mathbf{q} \in \mathfrak{X}(Q) :: \mathbf{q}(m) = \varphi_Q(m, (1,1))$ but, according to the Principle of Relativity, he has freedom of gauge and we must consider any gauge.

We assume that there is a principal bundle $Q(M, Spin(3,1) \times U, \pi_U)$ with fiber $Spin(3,1) \times U$ which represents the gauges used by observers. 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. The fundamental state of the particle is ψ_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 87 *There is a principal bundle $Q(M, Spin(3,1) \times U, \pi_U)$ with trivialization $\varphi_Q(m,(\sigma,\varkappa))$.*
The state of the particles is represented as vectors of the associated bundle $Q[E \otimes F, \vartheta]$
The value of the state as measured by an observer is $\psi(m) = \vartheta(\sigma(m), \varkappa(m))\psi_0$.

$Q[E \otimes F, \vartheta]$ has for trivialization :
$(\varphi_Q(m,(1,1)), \psi) \sim (\varphi_Q(m,(s^{-1}, g^{-1})), \vartheta(s,g)\psi)$
and holonomic basis:
$(\mathbf{e}_i(m) \otimes \mathbf{f}_j(m))_{i=0..3}^{j=1...n} = (\varphi_Q(m,(1,1)), e_i \otimes f_j)$

$$\begin{array}{c} \psi(m) = \sum_{i=1}^{4} \sum_{j=1}^{n} [\gamma C(\sigma(m))]_k^i [\varrho(\varkappa(m))]_l^j \psi_0^{kl}(m)\, \mathbf{e}_i(m) \otimes \mathbf{f}_j(m) \\ [\psi]_{4\times n} = [\gamma C(\sigma)]_{4\times 4} [\psi]_{4\times n} [\rho(\varkappa)]_{n \times n} \end{array} \tag{5.6}$$

A change of trivialization with a section $\chi(m) \in \mathfrak{X}(Q)$ induces a change of gauge :

$$\begin{array}{c} \mathbf{q}(m) = \varphi_Q(m,(1,1)) \to \widetilde{\mathbf{q}}(m) = \widetilde{\varphi}_Q(m,(1,1)) = \mathbf{q}(m) \cdot \chi(m)^{-1} \\ (\sigma(m), \varkappa(m)) = \varphi_Q(m,(\sigma,\varkappa)) = \widetilde{\varphi}_Q(m,(\widetilde{\sigma},\widetilde{\varkappa})) : (\widetilde{\sigma},\widetilde{\varkappa}) = \chi(m) \cdot (\sigma,\varkappa) \\ \mathbf{e}_i(m) \otimes \mathbf{f}_j(m) = (\mathbf{p}(m), e_i \otimes f_j) \to \widetilde{\mathbf{e}}_i(m) \otimes \widetilde{\mathbf{f}}_j(m) = \vartheta\left(\chi(m)^{-1}\right)(\mathbf{e}_i(m) \otimes \mathbf{f}_j(m)) \\ [\psi(m)] \to \left[\widetilde{\psi}(m)\right] = \vartheta(\chi(m))[\psi(m)] = [\gamma C(s)][\psi][\varrho(g)] \end{array} \tag{5.7}$$

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 $P_U[F, \varrho]$:

$$\begin{array}{c} \mathbf{p}_U(m) = \varphi_U(m,1) \to \widetilde{\mathbf{p}}_U(m) = \widetilde{\varphi}_U(m,1) = \mathbf{p}_U(m) \cdot \chi(m)^{-1} \\ \varkappa(m) = \varphi_U(m, \varkappa(m)) = \widetilde{\varphi}_U(m, \chi(m) \cdot \varkappa(m)) \\ \mathbf{f}_j(m) = (\mathbf{p}(m), f_j) \to \widetilde{\mathbf{f}}_j(m) = \varrho\left(\chi(m)^{-1}\right)(\mathbf{f}_j(m)) \\ \phi(m) \to \widetilde{\phi}(m) = \varrho(\chi(m))\phi(m) \end{array} \tag{5.9}$$

5.2.2 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 real 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 Lie algebra is $T_1 U(1) = \mathbb{R}$ and $\varrho'(1) = +i$ for the standard representation, $-i$ for the contragredient, and 0 for the trivial representation. The action of the EM field is then :

$\delta\mathcal{M} = [\psi]\left[\varrho'(1)\left(\delta\grave{A}\right)\right] = i\left(\delta\grave{A}\right)[\psi]$, or $-i\left(\delta\grave{A}\right)[\psi]$, or 0

with the variation $\delta\grave{A}$ of the potential along the trajectory, $\delta\grave{A} \in \mathbb{R}$.

The EM field interacts similarly with the left and right part of a spinor, so, *when no other field is involved*, the space of states of the particles is the sum of decomposable tensors : $S \otimes f$. And $f = e^{i\phi} f_p$ with fixed vectors f_p.Rather than to deal with 3 different representations, it is more convenient to assign a charge to the particle : $q = +1, -1, 0$

F becomes : $F = \left\{e^{iq\phi} f, \phi \in \mathbb{R}, q = +1, -1, 0\right\}$, with the action : $\varrho'(1)\left(\delta\grave{A}\right) = iq\left(\delta\grave{A}\right)$.

E is a complex vector space. The quantization of spinors fields show that they can be differentiated by a scalar (the mass), the spin and a signed integer $z \in \mathbb{Z}$. It is then legitimate, when only the gravitational and EM fields are considered, to choose the vectors $f \in E$. Inertial Spinors S_0 which differ by a complex number of module 1 have the same mass and inertial vector k, are differentiated by their charge q. The space of states of elementary particles, when only the gravitational and EM field are considered, is then given by :

$\widehat{E}_\epsilon = \left\{e^{iq\phi}\gamma C(\sigma) S_0, \phi \in \mathbb{R}, q = +1, -1, 0, S_0 \in E_\epsilon\right\}$

The inertial spinors of elementary particles change, by the Charge operator as follows :

Positive particles :

$$\begin{bmatrix} S_R & 0 \\ \epsilon i S_R & 0 \end{bmatrix}\begin{bmatrix} 0 & i\sigma_0 \\ i\sigma_0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & iS_R \\ 0 & -\epsilon S_R \end{bmatrix} = \begin{bmatrix} 0 & iS_R \\ 0 & -i\epsilon(iS_R) \end{bmatrix}$$

Negative particles :

$$\begin{bmatrix} 0 & S_R \\ 0 & \epsilon i S_R \end{bmatrix}\begin{bmatrix} 0 & i\sigma_0 \\ i\sigma_0 & 0 \end{bmatrix} = \begin{bmatrix} iS_R & 0 \\ -\epsilon S_R & 0 \end{bmatrix} = \begin{bmatrix} (iS_R) & 0 \\ -i\epsilon(iS_R) & 0 \end{bmatrix}$$

The particle takes the opposite charge, and becomes an anti-particle, with the same mass and opposite inertial vector.

Because the phase factor ϕ has no impact on the kinematic behavior of particles, it can be neglected : two particles such that their states differ by a phase factor $e^{i\phi}$ behave the same way for the gravitational field, so they can be deemed representing the same state. *This is the origin of the introduction of rays in QM.*

And the charge needs to be introduced explicitly only when the action of the EM field is considered (then q acts through the derivative).

5.2.3 Momentum and energy

Momentum

The motion of a particle is still represented by an element of $J^1 Cl(M)$:

$j^1 p = (m, \sigma, \upsilon(X_r, X_w))$

The extension of the spinor representation leads to define the momentum of a particle as an element of $J^1 Q[E \otimes F, \vartheta]$:

$\mathcal{M} = (m, \psi, \delta\psi) \in J^1 Q[E \otimes F, \vartheta]$

and along the trajectory of a particle by a map :
$\mathbb{R} \to J^1Q[E\otimes F,\vartheta] :: (m(t), \psi(t), \delta\psi(t))$
The relation between the motion and the momentum is represented for a spinor by the inertial spinor S_0.
By derivation we have :
$\left[\frac{d}{dt}\psi(t)\right] = \left[\gamma C\left(\frac{d\sigma}{dt}\right)\right][\psi_0][\varrho(\varkappa)] + [\gamma C(\sigma)][\psi_0]\left[\varrho\left(\frac{d\varkappa}{dt}\right)\right]$
and we should consider quantities such as $\partial_\alpha \varkappa$. However even if $\sigma, \varkappa$ are formally similar, they do not have the same physical meaning as we have noticed. The charge is represented with respect to the behavior of known particles, its value is conventional and $\varkappa$ measures the impact of a change of gauge by the observer. The momentum is related to the motion, but *in a motion it is assumed that the charge does not change* : actually it would imply a change of the fundamental state.

Proposition 88 *The momentum of a particle with motion : $j^1p = (m, \sigma, \upsilon(X_r, X_w)) \in J^1Cl(M)$ is represented by :*

$$\mathcal{M} = (m, \psi = \vartheta(\sigma,\varkappa)\psi_0, \delta\psi = \vartheta(\upsilon(X_r,X_w)\cdot\sigma,\varkappa)\psi_0) \in J^1Q[E\otimes F,\vartheta] \tag{5.10}$$

$\vartheta(\upsilon(X_r,X_w)\cdot\sigma,\varkappa)\psi_0 = \vartheta(\upsilon(X_r,X_w),1)\vartheta(\sigma,\varkappa)\psi_0 = \vartheta(\upsilon(X_r,X_w),1)\psi$
The value of a force field depends on the location. Due to the motion of the particle on its world line the value of the field changes. The field acts on $\mathcal{M}$ by a differential operator as we will see in the next section.

Energy

For the kinetic energy we look for :
$\langle\psi,\delta\psi\rangle = \langle\vartheta(\sigma,\varkappa)\psi_0, \vartheta(\upsilon(X_r,X_w)\cdot\sigma,\varkappa)\psi_0\rangle$
$= \left\langle\psi_0, \vartheta\left(\sigma^{-1}\cdot\upsilon(X_r,X_w)\cdot\sigma,1\right)\psi_0\right\rangle = \langle\psi_0, \gamma C(\mathbf{Ad}_{\sigma^{-1}}\upsilon(X_r,X_w))\psi_0\rangle$
Let us denote $Z = \mathbf{Ad}_{\sigma^{-1}}\upsilon(X_r,X_w) \in T_1Spin(3,1)$ in complex notation.
There is a vector similar to the inertial vector.
$[\psi_0] = \begin{bmatrix}\psi_R\\ \psi_L\end{bmatrix}$
$\gamma C(Z) = -\frac{1}{2}i\begin{bmatrix}\sigma(Z) & 0\\ 0 & \sigma(\overline{Z})\end{bmatrix}$
$\vartheta(Z,1)\psi_0 = \gamma C(Z)[\psi_0][\rho(1)] = -\frac{1}{2}i\begin{bmatrix}\sigma(Z)[\psi_R]\\ \sigma(\overline{Z})[\psi_L]\end{bmatrix}$
$\langle\psi_0, \vartheta(Z,1)\psi_0\rangle$
$= -\frac{1}{2}iTr\left(\begin{bmatrix}\psi_R^* & \psi_L^*\end{bmatrix}\begin{bmatrix}0 & -i\sigma_0\\ i\sigma_0 & 0\end{bmatrix}\begin{bmatrix}\sigma(Z)[\psi_R]\\ \sigma(\overline{Z})[\psi_L]\end{bmatrix}\right)$
$= \frac{1}{2}Tr\left(-[\psi_R^*]\sigma(\overline{Z})[\psi_L] + [\psi_L^*]\sigma(Z)[\psi_R]\right)$
$Tr([\psi_L^*]\sigma(Z)[\psi_R]) = Tr([\psi_L^*]\sigma(Z)[\psi_R])^t = Tr\left([\psi_R]^t[\sigma(Z)]^t\overline{[\psi_L]}\right) = Tr\overline{\left([\psi_R]^*\left[\sigma\left(\overline{Z}\right)\right][\psi_L]\right)}$
$\langle\psi_0, \vartheta(Z,1)\psi_0\rangle = i\operatorname{Im}Tr[\psi_L^*]\sigma(Z)[\psi_R]$
$= i\operatorname{Im}Tr[\psi_L^*]\sum_{a=1}^3 Z^a\sigma_a[\psi_R]$
Let be : $k^a = Tr[\psi_L^*]\sigma_a[\psi_R]$ then $\langle\psi_0, \vartheta(Z,1)\psi_0\rangle = i\operatorname{Im}k^tZ$

$$\begin{gathered}\langle\psi_0, \vartheta(Z,1)\psi_0\rangle = i\operatorname{Im}k^tZ\\ k^a = Tr[\psi_L^*]\sigma_a[\psi_R]\end{gathered} \tag{5.11}$$

The vector k, as well as ψ_0, is invariant in a change of gauge.
$a = 1,2,3$
Take $Z = \overrightarrow{\kappa_a}$
$\langle\psi_0, \vartheta(\overrightarrow{\kappa_a},1)\psi_0\rangle_E = \langle\psi_0, \gamma C(\overrightarrow{\kappa_a})\psi_0\rangle_E = \left\langle\psi_0, -\frac{1}{2}i\widetilde{\gamma}_a\psi_0\right\rangle_E = i\operatorname{Im}k^a$
$\operatorname{Im}k^a = -\frac{1}{2}\langle\psi_0, \widetilde{\gamma}_a\psi_0\rangle_E = \frac{1}{i}\langle\psi_0, \vartheta(\overrightarrow{\kappa_a},1)\psi_0\rangle_E$

Take $Z = \overrightarrow{\kappa_{a+3}} = i\overrightarrow{\kappa_a}$
$\langle \psi_0, \vartheta\left(\overrightarrow{\kappa_{a+3}}, 1\right)\psi_0\rangle_E = \langle \psi_0, \gamma C\left(i\overrightarrow{\kappa_a}\right)\psi_0\rangle_E = \frac{1}{2}i\langle \psi_0, \gamma_0\gamma_a\psi_0\rangle_E = i\operatorname{Im} ik^a = i\operatorname{Re} k^a$
$\operatorname{Re} k^a = \frac{1}{2}\langle \psi_0, \gamma_0\gamma_a\psi_0\rangle_E = \frac{1}{i}\langle \psi_0, \vartheta\left(\overrightarrow{\kappa_{a+3}}, 1\right)\psi_0\rangle_E$
$k^a = \frac{1}{2}\langle \psi_0, \gamma_0\gamma_a\psi_0\rangle_E + i\left(-\frac{1}{2}\langle \psi_0, \widetilde{\gamma}_a\psi_0\rangle_E\right) = \frac{1}{2}\langle \psi_0, \left(\gamma_0\gamma_a - i\widetilde{\gamma}_a\right)\psi_0\rangle_E$
$k^a = \frac{1}{2}\langle \psi_0, \left(\gamma_0\gamma_a - i\widetilde{\gamma}_a\right)\psi_0\rangle_E$ corresponds to the Dirac's current.
If $[\psi_L] = \epsilon i[\psi_R] : k^a = -\epsilon i Tr\left([\psi_R]^*\sigma_a[\psi_R]\right)$
$\overline{Tr\left([\psi_R]^*\sigma_a[\psi_R]\right)} = Tr\left([\psi_R]^t[\sigma_a]^t\overline{[\psi_R]}\right) = Tr\left([\psi_R]^t[\sigma_a]^t\overline{[\psi_R]}\right)^t = Tr\left([\psi_R]^*[\sigma_a][\psi_R]\right)^t$
Thus $Tr\left([\psi_R]^*\sigma_a[\psi_R]\right) \in \mathbb{R}$
And we will denote as for spinors :
$k = -i\epsilon\frac{M_p^2}{2}k_0$
$M_p = \sqrt{\epsilon\langle \psi_0, \psi_0\rangle} = \sqrt{\epsilon 2Tr\left(\psi_R^*\psi_R\right)}$
$k_0^a = \frac{2}{M_p^2}Tr\left([\psi_R]^*\sigma_a[\psi_R]\right) = \frac{Tr\left([\psi_R]^*\sigma_a[\psi_R]\right)}{Tr\left(\psi_R^*\psi_R\right)} \in \mathbb{R}$
$\langle \psi_0, \vartheta(Z,1)\psi_0\rangle = i\operatorname{Im}\left(-i\epsilon\frac{M_p^2}{2}k_0^t\right)Z = -i\epsilon\frac{M_p^2}{2}k_0^t\operatorname{Re} Z$

$$\begin{gathered}
\langle \psi_0, \vartheta(Z,1)\psi_0\rangle = i\operatorname{Im} k^t Z \\
k^a = Tr\left[\psi_L^*\right]\sigma_a[\psi_R] \\
[\psi_L] = \epsilon i[\psi_R] \Rightarrow k = -i\epsilon\frac{M_p^2}{2}k_0 \in i\mathbb{R} \\
k_0^a = \frac{Tr\left([\psi_R]^*\sigma_a[\psi_R]\right)}{Tr\left(\psi_R^*\psi_R\right)} \\
\langle \psi_0, \vartheta(Z,1)\psi_0\rangle = -i\epsilon\frac{M_p^2}{2}k_0^t\operatorname{Re} Z
\end{gathered} \tag{5.12}$$

So we can define the variation of kinetic energy as :

$$\begin{gathered}
\delta K = \frac{1}{M_p}\frac{1}{i}\langle \psi, \delta\psi\rangle = \frac{1}{M_p}\frac{1}{i}\langle \psi_0, \vartheta\left(\mathbf{Ad}_{\sigma^{-1}}\upsilon(X_r, X_w), 1\right)\psi_0\rangle \\
\delta K = -\frac{1}{2}\epsilon M_p k_0^t\operatorname{Re}\left(\mathbf{Ad}_{\sigma^{-1}}\upsilon(X_r, X_w)\right)
\end{gathered} \tag{5.13}$$

In a continuous motion along the trajectory : $\upsilon(X_r, X_w) = \frac{d\sigma}{dt}\cdot\sigma^{-1}$

Quantization

The quantity ψ sums up everything (motion, kinematic, charge) about the particle. A particle is then represented as a map :
$j^1\psi : \mathbb{R} \to J^1Q[E\otimes F, \vartheta] :: j^1\psi(t) = (q(t), \psi(t), \delta\psi(t))$
and in a continuous motion :
$\psi(t) = \vartheta(\sigma(t), \varkappa(t))\psi_0$
$\delta\psi(t) = \vartheta(\upsilon(X_r, X_w)\cdot\sigma, \varkappa)\psi_0 = \vartheta\left(\frac{d}{dt}\sigma(t)\cdot\sigma(t)^{-1}\cdot\sigma, \varkappa\right)\psi_0 = \vartheta\left(\frac{d}{dt}\sigma(t), \varkappa\right)\psi_0$

ψ is then a map : $\psi : [0,T] \to Q[E\otimes F, \vartheta]$ belonging to a normed vector space V, and we can implement the theorems of QM. The vector space is invariant by the action of $Spin(3,1)\times U$: $\lambda(g\times\varkappa, \psi)(t) = \vartheta(g, \varkappa)\psi(t)$.

(V, ϑ) is a representation of $Spin(3,1)\times U$. The observables of ψ are irreducible representations, characterized by a scalar, the mass, and a signed integer, the charge. Moreover the representation is faithful : for given values of $\psi_0, \psi(t)$ there is a unique couple $(\sigma(t)\times\varkappa(t))$ and thus a unique $\sigma(t)$. *For a given observer* $\sigma(t)$ admits two decompositions $\sigma(t) = \epsilon\sigma_w(t)\cdot\sigma_r(t)$.

The spin is represented by $\upsilon(X_r(t), 0) \in T_1 Spin(3)$ which is globally invariant by $Spin(3)$. Then an observable of the spatial spinor $\vartheta(\upsilon(X_r(t), 0), \varkappa)\psi_0$ corresponding to the rotational momentum belongs to an irreducible representation of $Spin(3)$, and is characterized by some $j \in \frac{1}{2}\mathbb{N}$. For elementary particles $j = \frac{1}{2}$. The change $X_r(t) \to -X_r(t)$ is a discontinuous process.

5.2.4 Matter fields

Composite particles

Composite stable particles can be represented by tensorial product of the vectors of their constituents. And this is the only way when the weak or strong interactions are involved.

When only the EM and gravitational fields are involved the states of elementary particles can be represented in E. Mathematically the tensorial product of non equivalent representations is well defined, however particles with the same charge must behave similarly, and the action of $\varrho'(1)$ should be the same on all the components, which must then have a charge of the same sign (this does not hold when the weak and strong interactions are considered). The electric charge must then be an integer multiple of an elementary charge. : $q \in \mathbb{Z}$. Nuclei or ionized atoms can be represented by a single spinor, with the total charge.

There is no extension to deformable solids, because the particles must have the same charge. However one can consider matter fields.

Matter and anti-matter

Anti-matter is the topic of many Sci-fi stories, based on the idea that anti-matter annihilates with matter.

The representation of the momentum by spinors leads to consider the existence of two distinct classes of elementary particles, and this distinction does not involve the force fields. The CPT principle leads to the conclusion that particles and antiparticles have opposite charge, so, except for the neutrinos, particle and antiparticles have opposite EM charge. But it does not mean that particles and antiparticles annihilate with their opposite number : mesons are composed of a particle and an anti-particle. Leptons (electrons and neutrinos) do not associate with other particles, and they annihilate with their anti-particles. Other particles are composed of quarks, which have a different behavior, and the outcome differs (but can also leads to annihilation). More generally weak and strong interactions must be considered on these issues.

One issue which is the topic of big experiments is if antiparticles have the same gravitational charge as their associated particles. According to our representation - and the limited experimental results available - they do.

Matter fields

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 : $\psi : [0,T] \to Q[E \otimes F, \vartheta]$ 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. If their trajectories do not cross the particles can be represented as section of $J^1Q[E \otimes F, \vartheta]$.

Definition 89 *A **matter field** is a section $\psi \in \mathfrak{X}(Q[E \otimes F, \vartheta])$ which, at each point, represents the state 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 90 *$\mathfrak{X}(M)$ is the set of matter fields, $\mathfrak{X}(\psi_0)$ the set of matter fields corresponding to $\psi_0 \in E \otimes F$.*

A necessary condition to be a matter field is : $\langle \psi(m), \psi(m) \rangle = Ct$. 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 $\sigma(m) = \epsilon\sigma_w(m) \cdot \sigma_r(m)$. By choosing $a_w > 0$ then $\sigma_w(m)$ defines a field of trajectories, as for the spinors.

The quantization is done as for Spinors fields. The conservation of mass and charge is assured through ψ_0. However a density μ can be defined as for spinors fields, with the same continuity equation. A matter field can represent a collection of identical particles whose trajectories do not cross.

A matter field gives a section of the 1st jet bundle $J^1Q[E \otimes F, \vartheta]$:

$J^1\psi(m) = (m, \psi(m), \partial_\alpha \psi(m), \alpha = 0..3)$

and a section of $J^1Q[E \otimes F, \vartheta]$ can represent a matter field whose motion is not necessarily continuous :

$\delta\psi(m) = (m, \psi(m), \delta_\alpha \psi(m), \alpha = 0..3)$

Wave function

For a continuous matter field belonging to the Fréchet space :

$L^1 = L^1(M, Q[E \otimes F, \vartheta], \varpi_4) = \left\{ \psi \in \mathfrak{X}(Q[E \otimes F, \vartheta]) : \int_\Omega \|\psi(m)\| \varpi_4(m) < \infty \right\}$

the evaluation map : $\mathcal{E}(m) : L^1(\psi_0) \to E \otimes F :: \mathcal{E}(m)\psi = \psi(m)$ is continuous.

Proof. 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^1

$\langle \psi - \psi_n, \psi - \psi_n \rangle(m)$ is continuous, ≥ 0 on the open Ω so there are

$A_n = \min_{m \in \Omega} \langle \psi - \psi_n, \psi - \psi_n \rangle(m)$

$\int_\Omega A_n \varpi_4 \leq \int_\Omega \langle \psi - \psi_n, \psi - \psi_n \rangle \varpi_4$

$\Rightarrow A_n \to 0$

$\Rightarrow \psi_n(m) \to \psi(m)$ ■

Usually a collection of particles of different types is 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, depending on the properties which are measured. 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

In QTF "matter fields" are, mathematically, similar to the matter fields defined here : they are sections of associated vector bundles. In QTF the Geometry is that of Special Relativity and the fundamental states ψ_0 are actually represented explicitly as individual particles labeled by their flavor, and usually their right or left parts when chirality is involved.... Overall, our picture provides a representation which is consistent with Classic Physics and account for the usual features of Quantum Physics. The main difference comes from the interpretation of "virtuality".

In our picture a particle at any time occupies a unique spatial location. To each particle p is associated a map : $\psi_p : \mathbb{R} \to Q[E \otimes F, \vartheta] :: \psi_p(t)$ which can be seen as a trajectory in a given matter field : $\psi_p(m(t))$. A matter field, completed with a density, can also be used to represent collectively a collection of particles which follow similar trajectories, which do not cross, and, with a density μ, there is not necessarily a particle at a given point. In both cases the matter field is just a blue print for the specification of the variables : it is a virtual particle, in the meaning that its specifics (initial location and state, observer) are not incorporated. The matter field can be used in a model, and PDE provide general solutions which are then fitted to the initial conditions.

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 the mathematical expression of a tautology) solutions to circumvent the problem, but the main consequence is that to *each* particle is associated a 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 is valued 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.

5.3 CONNECTIONS

When a particle travels on a path $V = \frac{dq}{dt}$ the value of the field changes, and this variation is valued in the Lie algebra : $p(m)^{-1} \cdot p'(m) V = \delta p(m) \in T_1 U$. It acts on the particle : it changes its momentum (and thus its motion through the derivative) and this action depends both on the state ψ of the particle and on $\delta p(m)$. So it can be represented as a map : $D : J^1 Q[E \otimes F] \to J^1 Q[E \otimes F]$. It is assumed that this action is linear. Then the action can be represented either as a linear differential operator or a covariant derivative (which combines the operator and the derivation). In both cases the field is represented as a connection.

5.3.1 Connections in Mathematics

Our purpose is to define a derivative of sections on fiber bundles (Maths.27). Vectors on the tangent space to a fiber bundle split in a part related to the base M and the other to the fiber V :

$\varphi : M \times V \to P :: p = \varphi(m, u)$

$\varphi' : T_m M \times T_u V \to T_p P :: v_p = \varphi'_m(m, g) v_m + \varphi'_u(m, g) v_u$

The vertical space $V_p P = \{\pi'(p) v_p = 0\}$ of $T_p P$ does not depend on the trivialization and is isomorphic to the tangent space of V.

However the splitting between $\varphi'_m(m, g) v_m, \varphi'_u(m, g) v_u$ is not unique and depends on the trivialization.

A **connection** is a projection of v_p on the vertical space $V_p P$. It is a one form on P valued in the vertical bundle VP. So it enables to distinguish in a variation of p what can be imputed to a change of m (the location) and what can be imputed to a change of u (the field). 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 and depends on the field.

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.

All that holds for any fiber bundle, 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 covariant derivative acts on sections, so on spinor or matter fields, and involves the derivative. A covariant derivative along the velocity gives an action on the derivative with respect to t, and so an operator on $\frac{dS}{dt}, \frac{d\psi}{dt}$, and on maps $S(t), \psi(t)$ if the action is continuous.

The second way to define a connection is through differential operators acting on the first jet prolongation of vector bundles [3], that is on moments, and it does not assume that the maps are continuously differentiable.

$\nabla : J^1 P_V \to P_V \otimes TM^* ::$

$\nabla\left(m, z^t, z_\alpha^i, \alpha = 0...3\right) = \sum_{\alpha,\beta=0}^{3} \sum_{i=1}^{n} \left(z_\alpha^i + \sum_{j=1}^{p} \Gamma_{\alpha j}^{j} z^j(m) \right) e_i(m) \otimes d\xi^\alpha$

then there is no derivative involved : actually z_α^i accounts for it. This is useful for single particles, whose state is defined as a map $\mathbb{R} \to J^1 Q[E \otimes F] : z_\alpha^i \to \frac{dz}{dt}$.

[3] One can define covariant derivarive of order greater than 1 this way.

5.3.2 Connection for the force fields other than Gravity

Connection on the principal bundle P_U

Connection

Its tangent space 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 fundamental vectors are :

$\zeta : T_1U \rightarrow VP_U :: \zeta(\theta)(\varphi_U(m,g)) = \varphi'_{Ug}(m,g)\, L'_{g^{-1}} g(\theta)$ with $L'_g(h)$ the derivative of the left translation : $L_g : U \rightarrow U :: L_g(h) = g \cdot h$

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.

A **connection** is a tensor, a one form $\grave{\mathbf{A}} \in \mathbf{\Lambda}_1(TP_U; VP_U)$ on TP valued in VP :

$\grave{\mathbf{A}}(p)\left(\sum_{\alpha=0}^{3} v_m^\alpha \partial m_\alpha + \zeta(\theta)(p)\right) = \zeta\left(\theta + L'_{g-1} g \left(\sum_{\alpha=0}^{3} v_m^\alpha \grave{A}_\alpha(p)\right)\right)(p)$

The **connection form** $\widehat{\grave{A}}$ of $\grave{\mathbf{A}}$ is :

$\widehat{\grave{A}}(p) : T_pP_U \rightarrow T_1U : \grave{\mathbf{A}}(p)(v_p) = \zeta\left(\widehat{\grave{A}}(p)(v_p)\right)(p)$

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 $\grave{\mathbf{A}}$.[4]

Its value for any gauge on 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. :

$$\grave{A} \in \Lambda_1(M; T_1U) : TM \rightarrow T_1U :: \grave{A}(m) = \sum_{\alpha=0}^{3} \sum_{a=1}^{m} \grave{A}_\alpha^a(m)\, \overrightarrow{\theta}_a \otimes d\xi^\alpha \qquad (5.14)$$

$\grave{A}$ is a one form on TM, and transforms as such in a change of chart, but this is not a tensor in T_1U. In a change of gauge : $\mathbf{p}_U(m) \rightarrow \widetilde{\mathbf{p}}_U(m) = \mathbf{p}_U(m) \cdot \chi(m)^{-1}$, $\grave{A}$ changes with an *affine law*, which involves the derivative $\chi'(m)$ of the change of gauge :

$\grave{A}(m) \rightarrow \widetilde{\grave{A}}(m) = Ad_\chi\left(\grave{A}(m) - L'_{\chi^{-1}}(\chi)\, \chi'(m)\right)$

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 :

$$\nabla^U \mathbf{p}_g = \left(L'_{g^{-1}} g\right)(g'(m)) + \sum_{\alpha=0}^{3} Ad_{g^{-1}} \grave{A}_\alpha(m)\, d\xi^\alpha \in \Lambda_1(M, T_1U) \qquad (5.15)$$

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 :

$\mathbf{p}_U(m) \rightarrow \widetilde{\mathbf{p}}_U(m) = \mathbf{p}_U(m) \cdot \chi(m)^{-1}$

$\grave{A}(m) \rightarrow \widetilde{\grave{A}}(m) = Ad_\chi\left(\grave{A}(m) - L'_{\chi^{-1}}(\chi)\, \chi'(m)\right)$

$\nabla^U \mathbf{p}_g \rightarrow \widetilde{\nabla^U \mathbf{p}_g} = \nabla^U \mathbf{p}_g$

[4] For the precise definition of pull-back, push-forward, of tensors see Maths.16.1 and the Formulas in the Annex.

Covariant derivative 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 $P_U[F,\varrho]$, through $\varrho'(1)$ (Maths.27.4). The formula is, in the standard gauge $\mathbf{p}_1 = \varphi_U(m,1)$:

$$\nabla^F \phi = \sum_{\alpha=0}^{3} \left(\partial_\alpha \phi^i + \sum_{i=1}^{n} \left[\grave{A}_\alpha \right]_j^i \phi^j \right) \mathbf{f}_i(m) \otimes d\xi^\alpha \in \Lambda_1(M, P_U[F,\varrho]) \tag{5.16}$$

with the

Notation 91 $\left[\grave{A}_\alpha\right] = \sum_{a=1}^{m} \grave{A}_\alpha^a [\theta_a]$ *is a $n \times n$ matrix representing*
$\varrho'(1)\left(\sum_{a=1}^{m} \grave{A}_\alpha^a \overrightarrow{\theta}_a\right) = \sum_{a=1}^{m} \grave{A}_\alpha^a \varrho'(1)\left(\overrightarrow{\theta}_a\right) \in \mathcal{L}(F;F)$

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.

For the interactions with particles, 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 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 the group U is a real Lie group, and its Lie algebra is a real vector space, 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.

With the convention about the action ϱ :
$\varrho'(1)\left(\grave{A}_\alpha \overrightarrow{\theta}\right) = iq\grave{A}_\alpha$

The action of $U(1)$ depends on the charge of the particle and the covariant derivative reads :

$$\nabla_\alpha^F \psi = \partial_\alpha \psi + qi\grave{A}_\alpha \psi \tag{5.17}$$

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 $T_1 Spin(3,1)$.

The potential G of a principal connection $\mathbf{G}$ on P_G is a map : $G \in \Lambda_1(M; T_1 Spin(3,1))$.

$$\begin{gathered} G \in \Lambda_1(M; T_1 Spin(3,1)) : TM \to T_1 Spin(3,1) :: \\ G(m) = \sum_{a=1}^{6} \sum_{\alpha=0}^{3} G_\alpha^a(m) \overrightarrow{\kappa}_a \otimes d\xi^\alpha = \sum_{\alpha=0}^{3} \upsilon(G_{r\alpha}(m), G_{w\alpha}(m)) d\xi^\alpha \end{gathered} \tag{5.18}$$

$G_{r\alpha}(m), G_{w\alpha}(m)$ 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 :

$\mathbf{p}(m) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m) \cdot \chi(m)^{-1} : G(m) \to \widetilde{G}(m) = \mathbf{Ad}_\chi \left(G(m) - L'_{\chi^{-1}}(\chi)\chi'(m)\right)$

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, as above :

$$\begin{gathered} \nabla^G : \mathfrak{X}(P_G) \to \Lambda_1(M; T_1Spin) :: \\ \nabla^G \sigma = \sum_{\alpha=0}^3 \mathbf{Ad}_{\sigma^{-1}} \left(\partial_\alpha \sigma \cdot \sigma^{-1} + G_\alpha\right) d\xi^\alpha \end{gathered} \tag{5.19}$$

The covariant derivative is invariant in a change of gauge :

$\mathbf{p}(m) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m) \cdot \chi(m)^{-1}$

$\nabla^G \sigma \to \widetilde{\nabla^G \sigma} = \nabla^G \sigma$

The action of the connection on a section of the 1st jet extension $J^1Cl(TM)$ is then given by a differential operator :

$j^1\sigma = (m, \sigma, \upsilon(X_{r\alpha}, X_{w\alpha}), \alpha = 0..3) \in J^1Cl(TM) \to (m, \sigma, \mathbf{Ad}_{\sigma^{-1}}(\upsilon(X_{r\alpha}, X_{w\alpha}) + G_\alpha), \alpha = 0..3) \in J^1Cl(TM)$

or equivalently : $\nabla^G_\alpha j^1\sigma = \sum_{\alpha=0}^3 \mathbf{Ad}_{\sigma^{-1}}(\upsilon(X_{r\alpha}, X_{w\alpha}) + G_\alpha)\, d\xi^\alpha$

Let be a particle with continuous trajectory along V. Its motion is :

$\frac{d\sigma}{dt} \cdot \sigma^{-1} = \upsilon(X_r(t), X_w(t))$

with $\sigma = \sigma_w \cdot \sigma_r = (a_w + \upsilon(0, w)) \cdot (a_r + \upsilon(r, 0))$

$X_r \simeq -\frac{1}{2}\left(1 + \frac{3}{4}\frac{\|\overrightarrow{v}\|^2}{c^2}\right) j\left(\frac{\overrightarrow{v}}{c}\right)\left(\frac{d}{dt}\frac{\overrightarrow{v}}{c}\right) + \left[1 - \frac{1}{2}j\left(\frac{\overrightarrow{v}}{c}\right)j\left(\frac{\overrightarrow{v}}{c}\right)\right]\left(\frac{1}{a_r} + \frac{1}{2}j(r) + \frac{1}{4a_r}j(r)j(r)\right)\frac{dr}{dt}$

$X_w \simeq \left(1 + \frac{\|\overrightarrow{v}\|^2}{c^2} - \frac{1}{2}j\left(\frac{\overrightarrow{v}}{c}\right)j\left(\frac{\overrightarrow{v}}{c}\right)\right)\left(\frac{d}{dt}\frac{\overrightarrow{v}}{c}\right) + \left(1 + \frac{1}{2}\frac{\|\overrightarrow{v}\|^2}{c^2}\right) j\left(\frac{\overrightarrow{v}}{c}\right)\left(\frac{1}{a_r} + \frac{1}{2}j(r) + \frac{1}{4a_r}j(r)j(r)\right)\frac{dr}{dt}$

The action of the gravitational field on the motion is :

$\frac{d\sigma}{dt} \cdot \sigma^{-1} \to \frac{d\sigma}{dt} \cdot \sigma^{-1} + \sum_{\alpha=0}^3 V^\alpha \upsilon(G_{r\alpha}, G_{w\alpha})$

So, in the usual conditions :

the component $\widehat{G}_r = \sum_{\alpha=0}^3 V^\alpha \upsilon(G_{r\alpha}, 0)$ acts on the rotational motion :

$X_r \to \left(\frac{1}{a_r} + \frac{1}{2}j(r) + \frac{1}{4a_r}j(r)j(r)\right)\frac{dr}{dt} + \widehat{G}_r$

the component $\widehat{G}_w = \sum_{\alpha=0}^3 V^\alpha \upsilon(0, G_{w\alpha})$ acts on the translational motion (as an acceleration) :

$X_w \to \frac{d}{dt}\frac{\overrightarrow{v}}{c} + \widehat{G}_w$

Covariant derivative for the adjoint bundle

The adjoint bundle $P_G[T_1Spin(3,1), \mathbf{Ad}]$ is a vector bundle with the action $\mathbf{Ad}$ whose derivative at $g = 1$ is the adjoint action :

$(\mathbf{Ad}_g)'|_{g=1}(X)(u) = [X, u]$

so the covariant derivative reads, for any section $X = \upsilon(X_r(m), X_w(m)) \in \mathfrak{X}(P_G[T_1Spin(3,1), \mathbf{Ad}])$:

$\nabla X = \sum_{\alpha=0}^3 (\partial_\alpha X + [G_\alpha, X])\, d\xi^\alpha \in \Lambda_1(M; P_G[T_1Spin(3,1), \mathbf{Ad}])$

Covariant derivative for spinors

The covariant derivative reads for a section $\mathbf{S} \in \mathfrak{X}(P_G[E,\gamma C])$:

$$\nabla^S S = \sum_{\alpha=0}^{3} \left(\partial_\alpha S + \gamma C(G_\alpha) S\right) d\xi^\alpha = \sum_{\alpha=0}^{3} \left(\partial_\alpha S + \gamma C\left(\upsilon\left(G_{r\alpha}, G_{w\alpha}\right)\right) S\right) d\xi^\alpha \tag{5.20}$$

The connection is evaluated in the holonomic gauge : $\mathbf{S} = (\mathbf{p}(m), S(m)) = (\varphi_G(m,1), S(m))$.

It preserves the chirality.

In a change of gauge :

$\mathbf{p}(m) \rightarrow \widetilde{\mathbf{p}}(m) = \mathbf{p}(m) \cdot \chi(m)^{-1}$

a section on $\mathfrak{X}(P_G[E,\gamma C])$ transforms as : $\widetilde{S}(m) = \gamma C(\chi(m)) S(m)$

The covariant derivative transforms as a section of $P_G[E,\gamma C]$ so the operator reads: $\nabla^S : \mathfrak{X}(P_G[E,\gamma C]) \rightarrow *_1(M; \mathfrak{X}(P_G[E,\gamma C]))$

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. Here $\mathbf{Ad}$ acts on vectors as the matrix $h(s) \in SO(3,1)$ it is then more convenient to use the representation of $T_1 Spin(3,1)$ by matrices of $so(3,1)$:

$$[\Gamma_{M\beta}] = \sum_{a=1}^{6} G_\beta^a [\kappa_a] = \begin{bmatrix} 0 & G^1_{w\beta} & G^2_{w\beta} & G^3_{w\beta} \\ G^1_{w\beta} & 0 & -G^3_{r\beta} & G^2_{r\beta} \\ G^2_{w\beta} & G^3_{r\beta} & 0 & -G^1_{\beta} \\ G^3_{w\beta} & -G^2_{r\beta} & G^1_{r\beta} & 0 \end{bmatrix}$$

In a change of gauge :

$G(m) \rightarrow \widetilde{G}(m) = \mathbf{Ad}_\chi\left(G(m) - L'_{\chi^{-1}}(\chi)\chi'(m)\right)$

$\left[\widetilde{\Gamma}_{M\alpha}\right] = [h(s)]\left([\Gamma_{M\alpha}] - \left[h\left(s^{-1}\right)\right][h(s')]\right)$

The covariant derivative of a section $U \in \mathfrak{X}\left(P_G\left[\mathbb{R}^4, \mathbf{Ad}\right]\right)$ is then :

$$\nabla^M U = \sum_{\alpha i=0}^{3} \left(\partial_\alpha U^i + \sum_{j=0}^{3} [\Gamma_{M\alpha}(m)]^i_j U^j\right) \varepsilon_i(m) \otimes d\xi^\alpha \tag{5.21}$$

For any vector field W : $\nabla^M_W : \mathfrak{X}(TM) \rightarrow \mathfrak{X}(TM)$ is a linear map which preserves the scalar product of vectors : $\left\langle \nabla^M_W U, \nabla^M_W V \right\rangle = \langle U, V \rangle$

The action of the connection can be extended from vector fields to any tensor on TM, by linearity.

The Levi-Civita connection

Historically (Kobayashi and Nomizu, Lang,...) the theory of fiber bundles has been developed from the study of the tangent bundle TM and of the set of all linear bases (holonomic or not). TM can be considered as a principal bundle with group $GL(4,\mathbb{R})$ and an associated vector bundle whose sections are vector fields. In this context Maths.16.4), the connections (called affine connections) are bilinear operators acting on vector fields : $\nabla \in \mathcal{L}^2(\mathfrak{X}(TM), \mathfrak{X}(TM); \mathfrak{X}(TM))$. They read in holonomic basis of a chart :

$\nabla V = \sum_\beta \left(\partial_\beta V^\alpha + \sum_\gamma \Gamma^\alpha_{\beta\gamma} V^\gamma\right) \partial\xi^\beta \otimes d\xi_\alpha$

with Christoffel symbols $\Gamma^\alpha_{\beta\gamma}(m)$ which change in a change of chart in a complicated way (similar to the potential). Their action is extended to tensors by linearity. 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. There is a unique connection which is both metric and symmetric, called the Levi-Civita connection. It reads :

$\Gamma^{\alpha}_{\beta\gamma} = \frac{1}{2}\sum_{\eta} g^{\alpha\eta}\left(\partial_\beta g_{\gamma\eta} + \partial_\gamma g_{\beta\eta} - \partial_\eta g_{\beta\gamma}\right)$

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.

In the "modern" theory of fiber bundles, in a principal bundle we can have any group, and it can be associated to almost any vector space. But the associated vector bundles inherit natural properties : because the representation $\left[\mathbb{R}^4, \mathbf{Ad}\right]$ is unitary, the bases (such as ε_i) of $P_G\left[\mathbb{R}^4, \mathbf{Ad}\right]$ are orthonormal, there is a scalar product and the metric is defined through the tetrad. As a consequence the linear connections inherit also special properties : any linear, principal connection $\mathbf{G}$ on P_G induces a linear connection on $P_G\left[\mathbb{R}^4, \mathbf{Ad}\right]$ which is metric. Its covariant derivative ∇^M can be written with Christoffel symbol Γ_M in the basis ε_i or translated into the holonomic basis $\left(\partial\xi_\alpha\right)_{\alpha=0}^3$ of any chart to give an affine connection with Christoffel coefficients $\widehat{\Gamma}^{\gamma}_{\alpha\beta}$:

$\widehat{\Gamma}^{\gamma}_{\alpha\beta} = P^{\gamma}_i\left(\partial_\alpha P'^i_\beta + \Gamma^i_{M\alpha j} P'^j_\beta\right)$

In matrix form :

$\widehat{\Gamma}^{\gamma}_{\alpha\beta} = \left[\widehat{\Gamma}_\alpha\right]^{\gamma}_{\beta}, \Gamma^i_{M\alpha j} = \left[\Gamma_{M\alpha}\right]^i_j,$

$\left[\Gamma_{M\alpha}\right] = \sum_{a=1}^6 G_{a\alpha}\left[\kappa_a\right]$

$\left[\widehat{\Gamma}_\alpha\right] = [P]\left(\left[\partial_\alpha P'\right] + \left[\Gamma_{M\alpha}\right][P']\right) \Leftrightarrow \left[\Gamma_{M\alpha}\right] = \left([P']\left[\widehat{\Gamma}_\alpha\right] - \left[\partial_\alpha P'\right]\right)[P]$

Any affine connection deduced this way from a principal connection is *necessarily metric*, but it is *not necessarily symmetric* (Maths.2191).

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 or tensors 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. The condition is :

$\forall\alpha, \beta, \gamma :$

$\left[\widehat{\Gamma}_\alpha\right]^{\gamma}_{\beta} = \left([P]\left(\left[\partial_\alpha P'\right] + \left[\Gamma_{M\alpha}\right][P']\right)\right)^{\gamma}_{\beta} = \left[\widehat{\Gamma}_\beta\right]^{\gamma}_{\alpha} = \left([P]\left(\left[\partial_\beta P'\right] + \left[\Gamma_{M\beta}\right][P']\right)\right)^{\gamma}_{\alpha}$

The Levi-Civita connection is the natural choice in a theory of gravitation based on the metric, such as the Einstein's theory, but other choices are possible and have been considered. In the formalism of affine connections they lead to great complications, meanwhile in the fiber formalism this is to impose the condition that the connection is symmetric, which is always possible at any time, which brings complications. These issues are addressed with the review of Einstein's theory in the following.

5.3.4 The total connection

Action of the fields

Proposition 92 *There are on Q a connection defined by the potentials*

$G \in \Lambda_1\left(M; T_1 Spin\left(3,1\right)\right) : TM \to T_1 Spin(3,1) ::$

$G\left(m\right) = \sum_{\alpha=0}^3 \upsilon\left(G_{r\alpha}\left(m\right), G_{w\alpha}\left(m\right)\right) d\xi^\alpha$

$\grave{A} \in \Lambda_1\left(M; T_1 U\right) : TM \to T_1 U :: \grave{A}\left(m\right) = \sum_{\alpha=0}^3 \sum_{a=1}^m \grave{A}^a_\alpha\left(m\right)\theta_a \otimes dm^\alpha$

The expression of the covariant derivative differs for a matter field - which is a section of a vector bundle - and for a single particle.

Matter field

A matter field is represented by a section $\psi \in Q[E \otimes F, \vartheta]$. Its covariant derivative is :

$\nabla : Q[E \otimes F, \vartheta] \to TM^* \otimes Q[E \otimes F, \vartheta] :: \nabla\psi = \sum_{\alpha=0}^{3} \sum_{i=1}^{4} \sum_{j=1}^{n} [\nabla_\alpha \psi]_j^i \, e_i \otimes f_j \otimes d\xi^\alpha$

$[\nabla_\alpha \psi] = [\partial_\alpha \psi] + [\gamma C(G_\alpha)][\psi] + [\psi]\left[\varrho'(1)\left(\grave{A}_\alpha\right)\right]$

$\psi = \vartheta(\sigma, \varkappa)\psi_0$

$\nabla_\alpha \psi = [\gamma C(\partial_\alpha \sigma)][\psi_0][\varrho(\varkappa)] + [\gamma C(\sigma)][\psi_0][\varrho'(\varkappa)(\partial_\alpha \varkappa)]$

$+[\gamma C(G_\alpha)][\gamma C(\sigma)][\psi_0][\varrho(\varkappa)] + [\gamma C(\sigma)][\psi_0][\varrho(\varkappa)]\left[\varrho'(1)\left(\grave{A}_\alpha\right)\right]$

$[\gamma C(\partial_\alpha \sigma)][\psi_0][\varrho(\varkappa)] = [\gamma C(\sigma)]\left[\gamma C\left(\sigma^{-1} \cdot \partial_\alpha \sigma\right)\right][\psi_0][\varrho(\varkappa)] = \vartheta(\sigma, \varkappa)\left(\left[\gamma C\left(\sigma^{-1} \cdot \partial_\alpha \sigma\right)\right][\psi_0]\right)$

$[\gamma C(\sigma)][\psi_0][\varrho'(\varkappa)(\partial_\alpha \varkappa)] = \vartheta(\sigma, \varkappa)\left([\psi_0][\varrho'(\varkappa)(\partial_\alpha \varkappa)]\left[\varrho\left(\varkappa^{-1}\right)\right]\right)$

$[\varrho'(\varkappa)(\partial_\alpha \varkappa)]\left[\varrho\left(\varkappa^{-1}\right)\right] = [\varrho(\varkappa)]\left[\varrho'(1) L'_{\varkappa^{-1}}\varkappa(\partial_\alpha \varkappa)\right]\left[\varrho\left(\varkappa^{-1}\right)\right] = Ad_\varkappa \left[\varrho'(1) L'_{\varkappa^{-1}}\varkappa(\partial_\alpha \varkappa)\right]$

$= \left[\varrho'(1) Ad_\varkappa L'_{\varkappa^{-1}}\varkappa(\partial_\alpha \varkappa)\right] = \left[\varrho'(1) R'_{\varkappa^{-1}}\varkappa L'_\varkappa 1 L'_{\varkappa^{-1}}\varkappa(\partial_\alpha \varkappa)\right]$

$= \left[\varrho'(1) R'_{\varkappa^{-1}}\varkappa(\partial_\alpha \varkappa)\right] = \left[R'_{\varkappa^{-1}}\varkappa(\partial_\alpha \varkappa)\right]$ (Maths.1900,23.2.1)

$[\gamma C(G_\alpha)][\gamma C(\sigma)][\psi_0][\varrho(\varkappa)]$

$= [\gamma C(\sigma)]\left[\gamma C\left(\sigma^{-1}\right)\right][\gamma C(G_\alpha)][\gamma C(\sigma)][\psi_0][\varrho(\varkappa)] = \vartheta(\sigma, \varkappa)\left([\gamma C(\mathbf{Ad}_{\sigma^{-1}} G_\alpha)][\psi_0]\right)$

$[\gamma C(\sigma)][\psi_0][\varrho(\varkappa)]\left[\varrho'(1)\left(\grave{A}_\alpha\right)\right]$

$= [\gamma C(\sigma)][\psi_0][\varrho(\varkappa)]\left[\varrho'(1)\left(\grave{A}_\alpha\right)\right]\left[\varrho\left(\varkappa^{-1}\right)\right][\varrho(\varkappa)]$

$= [\gamma C(\sigma)][\psi_0]\left[Ad_\varkappa \varrho'(1)\left(\grave{A}_\alpha\right)\right][\varrho(\varkappa)] = [\gamma C(\sigma)][\psi_0]\left[\varrho'(1) Ad_\varkappa\left(\grave{A}_\alpha\right)\right][\varrho(\varkappa)]$

$= \vartheta(\sigma, \varkappa)\left([\psi_0]\left[\varrho'(1) Ad_\varkappa\left(\grave{A}_\alpha\right)\right]\right) = \vartheta(\sigma, \varkappa)\left([\psi_0]\left[Ad_\varkappa\left(\grave{A}_\alpha\right)\right]\right)$

$[\nabla_\alpha \psi]$

$= \vartheta(\sigma, \varkappa)\left(\left[\gamma C\left(\sigma^{-1} \cdot \partial_\alpha \sigma\right)\right][\psi_0] + [\gamma C(\mathbf{Ad}_{\sigma^{-1}} G_\alpha)][\psi_0] + [\psi_0]\left[R'_{\varkappa^{-1}}\varkappa(\partial_\alpha \varkappa)\right] + [\psi_0]\left[Ad_\varkappa\left(\grave{A}_\alpha\right)\right]\right)$

$[\nabla_\alpha \psi]$

$= \vartheta(\sigma, \varkappa)\left(\left[\gamma C\left(\sigma^{-1} \cdot \partial_\alpha \sigma + \mathbf{Ad}_{\sigma^{-1}} G_\alpha\right)\right][\psi_0] + [\psi_0]\left[R'_{\varkappa^{-1}}\varkappa(\partial_\alpha \varkappa) + Ad_\varkappa\left(\grave{A}_\alpha\right)\right]\right)$

$\sigma^{-1} \cdot \partial_\alpha \sigma + \mathbf{Ad}_{\sigma^{-1}} G_\alpha = \mathbf{Ad}_{\sigma^{-1}}\left(\partial_\alpha \sigma \cdot \sigma^{-1} + G_\alpha\right) = \nabla_\alpha^G \sigma$

$[\nabla_\alpha \psi] = \vartheta(\sigma, \varkappa)\left(\left[\gamma C\left(\nabla_\alpha^G \sigma\right)\right][\psi_0] + [\psi_0]\left[R'_{\varkappa^{-1}}\varkappa(\partial_\alpha \varkappa) + Ad_\varkappa\left(\grave{A}_\alpha\right)\right]\right)$

Usually this is the value of the covariant derivative along the trajectory which is considered. So we have, along the trajectory of a particle which follows the integral curves of the vector field V associated to the section ψ :

$\vartheta\left(\sigma^{-1}, \varkappa^{-1}\right)\nabla_V \psi$

$= \left[\gamma C\left(\nabla_V^G \sigma\right)\right][\psi_0] + [\psi_0]\left[R'_{\varkappa^{-1}}\varkappa\left(\frac{d\varkappa}{dt}\right) + Ad_\varkappa\left(\sum_{\alpha=0}^{3} V^\alpha \grave{A}_\alpha\right)\right]$

The motion is continuous and the particle keeps its characteristics (that is ψ_0) : $\frac{d\varkappa}{dt} = 0$ thus :

$[\nabla_V \psi] = \vartheta(\sigma, \varkappa)\left(\left[\gamma C\left(\nabla_\alpha^G \sigma\right)\right][\psi_0] + [\psi_0]\left[Ad_\varkappa\left(\grave{A}_\alpha\right)\right]\right)$

and $\nabla_V \psi \in \mathfrak{X}(Q[E \otimes F])$.

The vector field V is defined through the section σ by :

$U(m) = -\frac{c}{\langle \mathbf{Ad}_\sigma \varepsilon_0, \varepsilon_0 \rangle_{Cl}} \mathbf{Ad}_\sigma \varepsilon_0$

$V(q(t)) = \frac{dq}{dt} = c\varepsilon_0 + \overrightarrow{v} = \sum_{\alpha,i=0}^{3} [P]_j^\alpha [U]^j \partial\xi_\alpha$

thus the tetrad is involved in the computation of :

$\frac{d\psi}{dt} = \sum_{\alpha=0}^{3} V^\alpha(q(t))\partial_\alpha \psi(q(t)) = \sum_{j=0}^{3} [U]^j \sum_{\alpha=0i}^{3} [P]_j^\alpha \partial_\alpha \psi(q(t))$

$\widehat{G} = \upsilon\left(\widehat{G}_r(t), \widehat{G}_w(t)\right) = \sum_{j=0}^{3} [U]^j \sum_{\alpha=0i}^{3} [P]_j^\alpha \upsilon(G_{r\alpha}(q(t)), G_{w\alpha}(q(t)))$

$\widehat{\grave{A}}(t) = \sum_{j=0}^{3} [U]^j \sum_{\alpha=0}^{3} [P]_j^\alpha \grave{A}_\alpha(q(t))$

$$\begin{gathered}[\nabla_\alpha \psi] = \vartheta(\sigma, \varkappa)\left(\left[\gamma C\left(\mathbf{Ad}_{\sigma^{-1}}\left(\partial_\alpha \sigma \cdot \sigma^{-1} + G_\alpha\right)\right)\right][\psi_0] + [\psi_0]\left[Ad_\varkappa\left(\grave{A}_\alpha\right)\right]\right) \\ [\nabla_V \psi] = \sum_{j=0}^{3} [U]^j \sum_{\alpha=0}^{3} [P]_j^\alpha [\nabla_\alpha \psi]\end{gathered} \tag{5.22}$$

The extension to a section $j^1\psi$ of the first jet bundle $J^1Q[E\otimes F,\vartheta]$ is immediate : the connection acts then as a differential operator :

$j^1\psi(m)=(m,\psi(m),\delta_\alpha\psi(m),\alpha=0..3)$

$\left[\nabla_\alpha j^1\psi\right]=\vartheta(\sigma,\varkappa)\left(\left[\gamma C\left(\mathbf{Ad}_{\sigma^{-1}}\left(\upsilon\left(X_{r\alpha},X_{w\alpha}\right)+G_\alpha\right)\right)\right]\left[\psi_0\right]+\left[\psi_0\right]\left[Ad_\varkappa\left(\grave{A}_\alpha\right)\right]\right)$

Single particle

It is represented by a map :

$\mathcal{M}:\mathbb{R}\to J^1Q[E\otimes F,\vartheta]::\mathcal{M}(t)=(q(t),\psi=\vartheta(\sigma,\varkappa)\psi_0,\delta\psi=\vartheta(\upsilon(X_r,X_w)\cdot\sigma,\varkappa)\psi_0)$

There is no one form ∇_α, but a differential operator on the 1st jet bundle :

$\nabla_V:C\left(\mathbb{R};J^1Q[E\otimes F,\vartheta]\right)\to C\left(\mathbb{R};J^1Q[E\otimes F,\vartheta]\right)$

along the trajectory.

$\nabla_V\psi=[\delta\psi]+\left[\gamma C\left(\sum_{\alpha=0}^3V^\alpha G_\alpha\right)\right][\psi]+[\psi]\left[\varrho'(1)\left(\sum_{\alpha=0}^3V^\alpha\sum_{a=1}^m\grave{A}_\alpha^a\overrightarrow{\theta}_a\right)\right]$

and with the same computation as above :

$\nabla_V\psi=\vartheta(\sigma,\varkappa)\left(\left[\gamma C\left(\mathbf{Ad}_{\sigma^{-1}}\left(\upsilon(X_r,X_w)+\left(\sum_{\alpha=0}^3V^\alpha G_\alpha\right)\right)\right)\right][\psi_0]+[\psi_0]\left[Ad_\varkappa\sum_{\alpha=0}^3V^\alpha\sum_{a=1}^m\grave{A}_\alpha^a\overrightarrow{\theta}_a\right]\right)$

The velocity depends on σ (and not its derivative) through :

$U(t)=-\frac{c}{\left\langle\mathbf{Ad}_{\sigma(t)}\varepsilon_0,\varepsilon_0\right\rangle_{Cl}}\mathbf{Ad}_\sigma\varepsilon_0$

$V^\alpha(t)=\sum_{i=0}^3\left[P'(q(t))\right]_i^\alpha U^i(t)$

then :

$\widehat{G}(t)=\upsilon\left(\widehat{G}_r(t),\widehat{G}_w(t)\right)=\sum_{j=0}^3[U(t)]^j\sum_{\alpha=0i}^3[P(q(t))]_j^\alpha\,\upsilon\left(G_{r\alpha}(q(t)),G_{w\alpha}(q(t))\right)$

$\widehat{\grave{A}}(t)=\sum_{j=0}^3[U(t)]^j\sum_{\alpha=0}^3[P(q(t))]_j^\alpha\,\grave{A}_\alpha(q(t))$

$$[\nabla_V\psi]=\vartheta(\sigma,\varkappa)\left(\left[\gamma C\left(\mathbf{Ad}_{\sigma^{-1}}\left(\upsilon(X_r,X_w)+\widehat{G}\right)\right)\right][\psi_0]+[\psi_0]\left[Ad_\varkappa\widehat{\grave{A}}(t)\right]\right)\qquad(5.23)$$

In a continuous motion : $\upsilon(X_r,X_w)=\frac{d\sigma}{dt}\cdot\sigma^{-1}$.

Energy of a particle

The variation of kinetic energy on a trajectory $V=\sum_{\alpha=0}^3V^\alpha\partial\xi_\alpha$ is :

$\delta K=\frac{1}{M_p}\frac{1}{i}\langle\psi,\delta\psi\rangle=-\frac{1}{2}\epsilon M_pk_0^t\operatorname{Re}\left(\mathbf{Ad}_{\sigma^{-1}}\upsilon(X_r,X_w)\right)$

The energy exchanged by the particle with the fields can be defined by :

$\frac{1}{M_p}\frac{1}{i}\left\langle\psi,[\gamma C(\delta G)][\psi]+[\psi]\left[\varrho'(1)\left(\delta\grave{A}\right)\right]\right\rangle$

and the variation of the total energy of the particle is :

$\delta E=\frac{1}{M_p}\frac{1}{i}\left\langle\psi,\delta\psi+[\gamma C(\delta G)][\psi]+[\psi]\left[\varrho'(1)\left(\delta\grave{A}\right)\right]\right\rangle$

$=\frac{1}{M_p}\frac{1}{i}\left\langle\vartheta(\sigma,\varkappa)\psi_0,\vartheta(\sigma,\varkappa)\left(\left[\gamma C\left(\nabla_\alpha^G\sigma\right)\right][\psi_0]+[\psi_0]\left[Ad_\varkappa\left(\grave{A}_\alpha\right)\right]\right)\right\rangle$

$=\frac{1}{M_p}\frac{1}{i}\left\langle\psi_0,\left[\gamma C\left(\nabla_V^G\sigma\right)\right][\psi_0]\right\rangle+\frac{1}{M_p}\frac{1}{i}\left\langle\psi_0,[\psi_0]\left[Ad_\varkappa\left(\widehat{\grave{A}}\right)\right]\right\rangle$

$=\delta K+\frac{1}{M_p}\frac{1}{i}\left\langle\psi_0,\left[\gamma C\left(\mathbf{Ad}_{\sigma^{-1}}\left(\widehat{G}\right)\right)\right][\psi_0]\right\rangle+\frac{1}{M_p}\frac{1}{i}\left\langle\psi_0,[\psi_0]\left[Ad_\varkappa\widehat{\grave{A}}\right]\right\rangle$

$\frac{1}{M_p}\frac{1}{i}\left\langle\psi_0,\left[\gamma C\left(\mathbf{Ad}_{\sigma^{-1}}\left(\widehat{G}\right)\right)\right][\psi_0]\right\rangle=-\epsilon\frac{M_p}{2}k_0^t\operatorname{Re}\mathbf{Ad}_{\sigma^{-1}}\left(\widehat{G}\right)$

$\frac{1}{M_p}\frac{1}{i}\left\langle\psi_0,[\psi_0]\left[Ad_\varkappa\widehat{\grave{A}}\right]\right\rangle=\frac{1}{M_p}\frac{1}{i}Tr\left([\psi_0]^*\gamma_0[\psi_0]\left[Ad_\varkappa\widehat{\grave{A}}\right]\right)$

$=\frac{1}{M_p}\frac{1}{i}\sum_{a=1}^m\left(Ad_\varkappa\widehat{\grave{A}}\right)^aTr\left([\psi_0]^*\gamma_0[\psi_0][\theta_a]\right)$

$\overline{Tr[\psi_0]^*\gamma_0[\psi_0][\theta_a]}=Tr\left([\psi_0]^*\gamma_0[\psi_0][\theta_a]\right)^*=Tr[\theta_a]^*[\psi_0]^*[\gamma_0][\psi_0]=-Tr[\theta_a][\psi]^*[\gamma_0][\psi]=-Tr[\psi]^*[\gamma_0][\psi][\theta_a]$

Thus : $\frac{1}{i}Tr\left([\psi_0]^*\gamma_0[\psi_0][\theta_a]\right)=\operatorname{Im}Tr\left([\psi_0]^*\gamma_0[\psi_0][\theta_a]\right)$

There is a vector k_c similar to k characteristic of the charges of the particle :

$\frac{1}{M_p}\frac{1}{i}\left\langle \psi_0, [\psi_0]\left[Ad_\varkappa \widehat{A}\right]\right\rangle = \frac{1}{M_p}\frac{1}{i}\sum_{a=1}^m \left(Ad_\varkappa \widehat{A}\right)^a Tr\left([\psi_0]^* \gamma_0 [\psi_0][\theta_a]\right)$

$= \frac{1}{M_p}\sum_{a=1}^m \left(Ad_\varkappa \widehat{A}\right)^a \operatorname{Im} Tr\left([\psi_0]^* \gamma_0 [\psi_0][\theta_a]\right) = -\epsilon \frac{M_p}{2} k_c^t \left[Ad_\varkappa \widehat{A}\right]$

$a = 1...m : k_c^a = -2\epsilon \frac{1}{M_p^2}\frac{1}{i}\left\langle \psi_0, [\psi_0]\left[Ad_\varkappa \widehat{A}\right]\right\rangle = -2\epsilon \frac{1}{M_p^2} \operatorname{Im} Tr\left([\psi_0]^* \gamma_0 [\psi_0][\theta_a]\right)$

The expression for k_c depends on the units, to be consistent, for the EM field :

$\varrho'(1)\left(\grave{A}_\alpha \overrightarrow{\theta}\right) = iq\grave{A}_\alpha = [\theta]\grave{A}_\alpha$

$k_{EM} = -2\epsilon \frac{1}{M_p^2} \operatorname{Im} Tr\left([\psi_0]^* \gamma_0 [\psi_0] iq\right) = -2\epsilon \frac{1}{M_p^2} q \operatorname{Re} Tr\left([\psi_0]^* \gamma_0 [\psi_0]\right) = -2q$

$$\delta E = \frac{1}{M_p}\frac{1}{i}\langle \psi, \nabla_V \psi\rangle = -\frac{1}{2}\epsilon M_p \left\{k_0^t \operatorname{Re} \mathbf{Ad}_{\sigma^{-1}}\left(\upsilon(X_r, X_w) + \widehat{G}\right) + k_c^t\left(Ad_\varkappa \widehat{A}\right)\right\} \tag{5.24}$$

This quantity is the balance of the energy exchanged by the particle along its trajectory. We will see in the following that, at equilibrium : $\delta E = 0$. An increase in the kinetic energy is balanced by an energy received from the fields, and conversely a decrease of the kinetic energy implies a transfer of energy to the fields. For a free particle : $\frac{1}{i}\langle \psi, \nabla_V \psi\rangle = \delta K$ which is minimum for a continuous motion.

5.3.5 Geodesics

There are several concepts of geodesics, parallel transport, lift of a curve, which are related but distinct. We will see here the concepts related to the parallel transport of a vector along a curve by a connection. It can be implemented with any connection defined over the tangent bundle TM of the base manifold, and is of particular interest for the connection induced by the gravitational field.

Parallel transport of a vector by a connection

Let C be a curve defined by a path $p : \mathbb{R} \to M : p(\tau)$ with $p(0) = a$, and a vector $v \in T_a M$. The vector, **parallel transported** by the connection along C, is given by a map :

$U : \mathbb{R} \to T_{p(\tau)} M : U(\tau)$ such that : $\nabla^M_{\frac{dp}{d\tau}} U(\tau) = 0, U(0) = v$

thus we have the differential equation with $U(\tau) = \sum_{i=0}^3 U^i(\tau)\varepsilon_i(p(\tau))$

$\nabla^M_{\frac{dp}{d\tau}} U(\tau) = \sum_{i=0}^3 \left(\frac{d}{dt}U^i + \sum_{\alpha,j=0}^3 \Gamma_M(p(\tau))^i_{\alpha j} U^j \left(\frac{dp}{d\tau}\right)^\alpha\right)\varepsilon_i(p(t)) = 0$

$\frac{dU^i}{d\tau} + \sum_{\alpha j=0}^3 \Gamma_M(p(\tau))^i_{\alpha j} U^j \left(\frac{dp}{d\tau}\right)^\alpha = 0$

Geodesic

A **geodesic** is a path such that its tangent is parallel transported by the connection :

$p : \mathbb{R} \to M : p(\tau)$ with $p(0) = a$

$U(\tau) = \frac{dp}{d\tau} = \sum_{i=0}^3 U^i(\tau)\varepsilon_i(p(\tau)) = \sum_{k,\alpha=0}^3 U^k(\tau) P_k^\alpha(p(\tau))\partial\xi_\alpha$

$\frac{dU^i}{d\tau} + \sum_{\alpha jk=0}^3 \Gamma_M(p(\tau))^i_{\alpha j} U^j(\tau) U^k(\tau) P_k^\alpha(p(\tau)) = 0$

or in matrix form :

$\left[\frac{dU}{d\tau}\right] + \sum_\alpha ([\Gamma_{M\alpha}][U])([P][U])^\alpha = 0$

The scalar product $\langle U, U\rangle$ is constant :

$\frac{d}{d\tau}\langle U, U\rangle = \frac{d}{d\tau}\left([U]^t[\eta][U]\right)$

$= \left[\frac{dU}{d\tau}\right]^t [\eta][U] + [U]^t[\eta]\left[\frac{dU}{d\tau}\right]$

$= -\sum_\alpha ([P][U])^\alpha [U]^t [\Gamma_{M\alpha}]^t [\eta][U] - \sum_\alpha ([P][U])^\alpha [U]^t [\eta]([\Gamma_{M\alpha}][U])$

$= -\sum_\alpha ([P][U])^\alpha [U]^t \left([\Gamma_{M\alpha}]^t[\eta] + [\eta][\Gamma_{M\alpha}]\right)[U] = 0$

Field of Geodesics

A **field of geodesics** is a vector field U such that it is parallel transported along its integral curves $q(\tau) = \Phi_V(\tau, x)$. So a field of geodesics has a constant length : $\langle U, U\rangle = Ct$ which can be null.

In the standard chart of any observer :

$q(\tau) = \varphi_o(t(\tau), \xi(\tau))$

$U \in \mathfrak{X}\left(P_G\left[\mathbb{R}^4, \mathbf{Ad}\right]\right) \sim V \in \mathfrak{X}(TM)$ with $V^\alpha = \sum_{i=0}^3 U^i P_i^\alpha : [U] = [P'][V]$

$\left[\frac{dU}{d\tau}\right] + \sum_\alpha ([\Gamma_{M\alpha}][U])([P][U])^\alpha = 0$

$\Leftrightarrow$

$\left[\frac{dP'}{d\tau}\right][V] + [P']\left[\frac{dV}{d\tau}\right] + \sum_\alpha ([\Gamma_{M\alpha}][P'][V])([P][P'][V])^\alpha = 0$

$[P]\left[\frac{dP'}{d\tau}\right][V] + \left[\frac{dV}{d\tau}\right] + \sum_\alpha [P][\Gamma_{M\alpha}][P'][V][V]^\alpha = 0$

$\left[\frac{dP'}{d\tau}\right] = \sum_{\alpha=0}^3 V^\alpha [\partial_\alpha P']$

$\left[\frac{dV}{d\tau}\right] + \sum_\alpha V^\alpha [P][\partial_\alpha P'][V] + [P][\Gamma_{M\alpha}][P'][V][V]^\alpha = 0$

$\left[\frac{dV}{d\tau}\right] + \sum_\alpha V^\alpha [P]([\partial_\alpha P'] + [\Gamma_{M\alpha}][P'])[V] = 0$

that is in the holonomic basis :

$\left[\frac{dV}{d\tau}\right] + \sum_\alpha V^\alpha \left[\widehat{\Gamma}_\alpha\right][V] = 0$

with $\left[\widehat{\Gamma}_\alpha\right] = [P]([\partial_\alpha P'] + [\Gamma_{M\alpha}][P'])$

A curve is lifted from M to a fiber bundle P by imposing that its tangent in TP is null with a connection : it belongs to a horizontal vector field. So one can say that a field of geodesics is the projection on TM of a field of projectable horizontal vector fields on the fiber bundle P.

To any, non null, future oriented, vector field V one can associate a section of $\sigma \in P_W$ such that

$V(m) = \sqrt{-\langle V(m), V(m)\rangle_{TM}}\mathbf{Ad}_{\sigma(m)}\varepsilon_0(m)$

and we have the following :

Theorem 93 *For a given observer, fields of geodesics are represented by sections $\sigma \in \mathfrak{X}(P_G)$ such that $\nabla_U^G \sigma \in T_1 Spin(3)$. They are solutions of the differential equation :*

$$\frac{dw}{d\tau} = [j(w)]\widehat{G}_r + \left(-a_w + \frac{1}{4a_w} j(w) j(w)\right)\widehat{G}_w \tag{5.25}$$

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) The scalar product is constant along a geodesic :

$\langle V(m), V(m)\rangle_{TM} = -k^2$

In the tetrad :

$\frac{V}{k} = U = \mathbf{Ad}_\sigma \varepsilon_0$

For the sections of $P_G\left[\mathbb{R}^4, \mathbf{Ad}\right]$ the covariant derivative reads :

$\nabla_U^M U = \frac{dU}{d\tau} + \sum_{\alpha=0}^3 V^\alpha [\upsilon(G_{r\alpha}, G_{w\alpha}), U]$

because $(\mathbf{Ad}_\sigma)'|_{\sigma=1} = ad$ and the condition reads :

$\nabla_U^M U = \frac{dU}{d\tau} + \left[\upsilon\left(\widehat{G}_r, \widehat{G}_w\right), U\right] = 0$

with $\sum_{\alpha=0}^3 V^\alpha \upsilon(G_{r\alpha}, G_{w\alpha}) = \upsilon\left(\widehat{G}_r, \widehat{G}_w\right)$

ii) $\frac{dU}{d\tau} = \frac{d}{d\tau}\mathbf{Ad}_\sigma \varepsilon_0 = \frac{d}{d\tau}\left(\sigma \cdot \varepsilon_0 \cdot \sigma^{-1}\right)$

$= \frac{d\sigma}{d\tau} \cdot \varepsilon_0 \cdot \sigma^{-1} - \sigma \cdot \varepsilon_0 \cdot \sigma^{-1} \cdot \frac{d\sigma}{d\tau} \cdot \sigma^{-1}$

$= \mathbf{Ad}_\sigma \left[\sigma^{-1} \cdot \frac{d\sigma}{d\tau}, \varepsilon_0\right]$

$\mathbf{Ad}_\sigma \left[\sigma^{-1} \cdot \frac{d\sigma}{d\tau}, \varepsilon_0\right] + \left[\upsilon\left(\widehat{G}_r, \widehat{G}_w\right) \cdot U\right] = 0$

$\left[\mathbf{Ad}_\sigma \left(\sigma^{-1} \cdot \frac{d\sigma}{d\tau}\right), \mathbf{Ad}_\sigma \varepsilon_0\right] + \left[\upsilon\left(\widehat{G}_r, \widehat{G}_w\right), \mathbf{Ad}_\sigma \varepsilon_0\right] = 0$

$\left[\mathbf{Ad}_{\sigma}\left(\sigma^{-1}\cdot\frac{d\sigma}{d\tau}+\mathbf{Ad}_{\sigma^{-1}}\upsilon\left(\widehat{G}_r,\widehat{G}_w\right)\right),\mathbf{Ad}_{\sigma}\varepsilon_0\right]=0$
$\left[\mathbf{Ad}_{\sigma}\left(\nabla_U^G\sigma\right),\mathbf{Ad}_{\sigma}\varepsilon_0\right]=0$
$\left[\nabla_U^G\sigma,\varepsilon_0\right]=0$
The only elements of $Cl\left(3,1\right)$ which commute with ε_0 belong to $T_1Spin\left(3\right)$
iii) σ is defined by : $\sigma=a_w+\upsilon\left(0,w\right)$
$\sigma^{-1}\cdot\frac{d\sigma}{d\tau}=\upsilon(\frac{1}{2}j\left(w\right)\frac{dw}{d\tau},\frac{1}{4a_w}\left(-j\left(w\right)j\left(w\right)+4\right)\frac{dw}{d\tau})$
$\mathbf{Ad}_{\sigma^{-1}}\upsilon\left(\widehat{G}_r,\widehat{G}_w\right)=\left[\begin{array}{cc}\left[1-\frac{1}{2}j\left(w\right)j\left(w\right)\right] & \left[a_wj\left(w\right)\right]\\ -\left[a_wj\left(w\right)\right] & \left[1-\frac{1}{2}j\left(w\right)j\left(w\right)\right]\end{array}\right]\left[\begin{array}{c}\widehat{G}_r\\ \widehat{G}_w\end{array}\right]$
So geodesic fields are associated to the sections such that :
$\frac{1}{a_w}\left(1-\frac{1}{4}j\left(w\right)j\left(w\right)\right)\frac{dw}{d\tau}-a_w\left[j\left(w\right)\right]\widehat{G}_r+\left[1-\frac{1}{2}j\left(w\right)j\left(w\right)\right]\widehat{G}_w=0$
By left multiplication with w^t :
$w^t\frac{dw}{d\tau}+a_ww^t\widehat{G}_w=0$
$w^t\frac{dw}{d\tau}=4a_w\frac{da_w}{dt}=-a_ww^t\widehat{G}_w$
$\frac{da_w}{d\tau}=-\frac{1}{4}w^t\widehat{G}_w$
The equation becomes :
$\left(-\frac{1}{4}ww^t+\frac{1}{4}w^tw+1\right)\frac{dw}{d\tau}-a_w^2\left[j\left(w\right)\right]\widehat{G}_r+a_w\left[1-\frac{1}{2}\left(ww^t-w^tw\right)\right]\widehat{G}_w=0$
$-\left(\frac{1}{4}\left(w^t\frac{dw}{d\tau}\right)+\frac{1}{2}a_ww^t\widehat{G}_w\right)w+\left(\frac{1}{4}4\left(a_w^2-1\right)+1\right)\frac{dw}{d\tau}-a_w^2\left[j\left(w\right)\right]\widehat{G}_r+a_w\left(1+\frac{1}{2}w^tw\right)\widehat{G}_w=0$
$\left(a_w\frac{da_w}{d\tau}\right)w+a_w^2\frac{dw}{d\tau}-a_w^2\left[j\left(w\right)\right]\widehat{G}_r+a_w\left(2a_w^2-1\right)\widehat{G}_w=0$
$\frac{da_w}{d\tau}w+a_w\frac{dw}{d\tau}-a_w\left[j\left(w\right)\right]\widehat{G}_r+\left(2a_w^2-1\right)\widehat{G}_w=0$
$-\frac{1}{4}ww^t\widehat{G}_w+a_w\frac{dw}{d\tau}-a_w\left[j\left(w\right)\right]\widehat{G}_r+\left(2a_w^2-1\right)\widehat{G}_w=0$
$a_w\frac{dw}{d\tau}-a_w\left[j\left(w\right)\right]\widehat{G}_r+\left(\left(2a_w^2-1\right)-\frac{1}{4}\left(j\left(w\right)j\left(w\right)+4\left(a_w^2-1\right)\right)\right)\widehat{G}_w=0$
$a_w\frac{dw}{d\tau}-a_w\left[j\left(w\right)\right]\widehat{G}_r+\left(a_w^2-\frac{1}{4}j\left(w\right)j\left(w\right)\right)\widehat{G}_w=0$
$\frac{dw}{d\tau}=\left[j\left(w\right)\right]\widehat{G}_r+\left(-a_w+\frac{1}{4a_w}j\left(w\right)j\left(w\right)\right)\widehat{G}_w$ ■

There are other definitions of geodesic curves, in particular as curve with an extremal length. It holds on any metric space, and so on any manifold endowed with a metric. A classic demonstration proves that a curve of extremal length is necessarily a curve along which the tangent is transported, and so a geodesic as understood here, 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.

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 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 far away, 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 are represented together 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 in 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.

In a realist point of view, field are physical entities, as well as particles. Their properties are represented by variables, which give the value of the measures which can be done about them. These variables, as mathematical objects, can be defined over any abstract domain, such as M or $\mathbb{R}$. However, as representing physical entities, their value is defined only if it can be measured - this is where one retrieves the criterion of QM. And, on this point, particles and fields are very different.

Particles, by definition, occupy a unique location at a given time : this is their main property and actually it is linked to the concept of location itself. From this property one deduces that they travel on their world line : so they occupy physically only one location at each time, parametrized by a single scalar, their proper time τ. When their characteristics are represented by variables X, the domain of these variables is $\mathbb{R}$, but the value of X is not fixed for every τ : the value of X is fixed only if it can be measured and this depends on the observer. For a given observer there is a relation between his proper time and the proper time of each particle, and so X is defined only if $\tau > \tau_P(t)$. Before $\tau_P(t)$ the variable X has a definite value (which has or could have been measured) and for $\tau > \tau_P(t)$ the variable has not yet a definite value. The Principle of Causality, which requires that these restrictions are consistent for all observers, leads then to the existence of the Lorentz metric, as shown in the Chapter 3.

Fields, by definition exist everywhere, or more precisely the properties of a given field can be measured at different locations in space and time : the variables Y representing the field have for domain M itself. However the limitation imposed by the possibility to be measured still holds : for a given observer the distinction between past, present and future is clear, and a field cannot be measured in the future, so even if Y is defined over M as a mathematical entity, its value, as a physical object is not defined everywhere. For a given observer O the partition of M between a domain M_o^- where Y has a fixed value, M_o^+ where it has no fixed value, and the border $\partial_o M$ where the propagation occurs, is given by $\Omega_3(t)$. However, if the field is a physical entity, it should exist a partition which does not depend on the observer. There should exist a variable $s \in \mathbb{R}$, the phase, similar to the proper time of a particle, and some function $F : M \to \mathbb{R}$ such that $F(m) = s$ tells "when" the propagation has occurred at m. The function F defines a foliation of M by hypersurfaces $W(s)$ which are the front of the propagation. The causal structure implies then that s does not depend of the observer, and that the time t of any observer is related (up to an additive constant) to s. This is the opposite of the particle case.

This is the picture of the usual cosmological models. M is just an abstract object, but it represents a physical entity, a container, whose content is not frozen. And s can be seen as a universal time : the time of the Universe (or its age...). It partitions the Universe between all that has already happened, and what has to happen. When one attempts to represent the whole Universe one cannot escape the issue of the observer. The function F defines a universal chart : $W(s)$ are space like hypersurfaces, and $gradF'(m)$ defines a vector field of time like, future oriented, vectors normal to $W(s)$. The "universal observer" uses this chart, and his proper time is just s. The model is then consistent. All physical entities "live" in the same spatial universe $W(s)$ which moves with s increasing. But we do not come back to the Galilean Geometry : $W(s)$ is a Riemannian 3 dimensional surface, not a plane and the metric varies with s (this is the expansion of the universe). And of course the only field which is considered is gravity, because on average the distribution of positive and negative charges is null.

But, if the concept of fields as a physical entity sustains the usual cosmological models, it is not easy to conciliate them with our Physics, notably because "time" has not the same value. The charts that we can use are conventional : their unique purpose is to locate a point, from phenomena that we measure at *our* location. We can use the direction of a far away star to fix a vector of a basis, but that does not imply that the point so located is precisely on the border $W(s)$. We have no way to know if $s = t$. We can compare the rates at which work two clocks located at different points, but the observer cannot tell if the rate of his own clock changes with time, he can only assume that it stays constant. However, because the measures of lengths and time rely, practically, on fields, our charts reflect actually their propagation. 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.

In this Section we introduce the main concepts and variables which represent the propagation of fields.

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

Main features of the tangent space to a principal fiber bundle

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 $\sum_{\alpha=0}^{3} v_m^\alpha \partial m_\alpha$ is a projectable vector field iff v_m^α does not depend on g.

Connection and horizontal vectors

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_U valued in VP_U. For a principal connection its value depend on the potential $\grave{A}$:

$\grave{\mathbf{A}}(\mathbf{p}(m))(\varphi'_m(m,1)\, v_m + \zeta(\theta)(\mathbf{p}(m))) = \zeta\left(\theta + \sum_\alpha \grave{A}_\alpha(m)\, v_m^\alpha\right)(\mathbf{p}(m))$

The vertical vector bundle $VP_U = \ker \pi'$ depends only on the principal bundle structure. Similarly for each connection there is a vector bundle, the **horizontal bundle** $HP_U = \ker \grave{\mathbf{A}}$, which is a vector subbundle HP_U of TP_U depending on the connection, and :

$TP_U = HP_U \oplus VP_U$

$\pi'_U (VP_U) = 0$

$\grave{\mathbf{A}} (HP_U) = 0$

$\dim VP_U = \dim T_1 U = m$

$\dim HP_U = \dim TP_U - \dim VP_U = \dim M = 4$

$HP_U = \left\{ \sum_{\alpha=0}^{3} v_m^\alpha \partial m_\alpha + \zeta(\theta)(p) : \theta + Ad_{g^{-1}} \sum_{\alpha=0}^{3} \grave{A}_\alpha (m) v_m^\alpha = 0 \right\}$

The vectors of HP_U are called horizontal.

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 :

$\lambda = \sum_{a=1}^{m} \sum_{\{\alpha_1,...\alpha_r\}=0}^{3} \lambda_{\alpha_1....\alpha_r}^a dm^{\alpha_1} \wedge .... \wedge dm^{\alpha_r} \otimes \overrightarrow{f}_a$

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 :

$\chi (p) :: T_p P_U \to H_p P_U :: \chi (p)(v_p) = v_p - \grave{\mathbf{A}} (p)(v_p)$

$\chi (p) \left(\sum_\alpha v_m^\alpha \partial m_\alpha + \zeta (\theta)(p) \right) = \sum_\alpha v_m^\alpha \partial m_\alpha - \zeta \left(Ad_{g^{-1}} \grave{A} (m) v_m \right)(p)$

χ is a projection on the horizontal bundle :

$v_p \in V_p P_U : \grave{\mathbf{A}} (p)(v_p) = v_p \Rightarrow \chi (p)(v_p) = 0$

$v_p \in HP_U : \chi (p)(v_p) = v_p$

The horizontal lift of a vector field $V \in \mathfrak{X} (TM)$:

$\chi_L : \mathfrak{X} (TM) \to \mathfrak{X} (HP_U) :: \chi_L (p)(V) = \varphi'_P (m, g) \left(V(m), -\grave{\mathbf{A}} (p) V(m) \right)$

and $W = \chi_L (p)(V)$ is a vector field projectable on V : $\pi'_U (p)(\chi_L (p)(V)) = V(\pi_U (p))$ such that $\grave{\mathbf{A}} (W) = 0$

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 :

$\chi^* (p) : \Lambda_r (T_p P; V) \to \Lambda_r (H_p P; V) :: \chi^* (p) \omega (p)(v_1, ..., v_r) = \omega (p)(\chi (p) v_1, ..., \chi (p) v_r)$

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 $\chi^* \widehat{\grave{A}} = 0$.

Derivative of a tensor

The set of tensors on a manifold valued in a fixed vector space is an algebra $\mathcal{T}$, with the tensorial product as internal operation. A derivative on a manifold is along a vector field (or along a curve), and the derivative of a tensor is an operator $D : TM \times \mathcal{T} \to \mathcal{T}$ called a derivation, which meets the properties (Maths.16.2.1) :

- it does not change the nature of the tensor, thus if $T \in \Lambda_r (TM; F) : D_V (T) \in \Lambda_r (TM; F), D(T) \in \Lambda_{r+1} (TM; F)$,
- it is linear with respect to $V : D_{V+W} (T) = D_V (T) + D_W (T)$
- it is a linear operator on $\mathcal{T}$
- it follows the Leibniz rule with respect to the tensorial product : $D_V (T \otimes T') = D_V (T) \otimes T' + T \otimes D_V (T')$
- it commutes with the trace operator, and the contraction of tensors.

The only general operator which meets these criteria is the Lie derivative (Maths.16.2.2). Using the flow of the vector field V , by pull back or push forward one can bring the tensors in the same vector space and compute the quantities :

$\lim_{h\to 0}\Delta_R(h)=\lim_{h\to 0}\frac{1}{h}\left(\Phi_V(s+h,p)^* T(p)-\Phi_V(s,p)^* T(p)\right)$
$\lim_{h\to 0}\Delta_L(h)=\lim_{h\to 0}\frac{1}{h}\left(\Phi_V(s,p)^* T(p)-\Phi_V(s-h,p)^* T(p)\right)$
If the tensor is continuously differentiable, then the Lie derivative is :
$\pounds_V T=\frac{d}{ds}\Phi_V(s,p)^* T(p)|_{s=0}=\lim_{h\to 0}\Delta_R(h)=\lim_{h\to 0}\Delta_L(h)$
It can be extended to any tensor valued in a *fixed* vector space, and it holds for any manifold.

Notice that one can have a right and a left derivative which have a different value : we have a discontinuity, and we will come back to this possibility in the last chapter.

A principal connection is defined by the connection form $\widehat{\grave{A}}\in\Lambda_1(TP_U;T_1U)$ which is a tensor valued in the fixed vector space T_1U :

$\widehat{\grave{A}}(p):T_pP_U\to T_1U::\grave{\mathbf{A}}(p)(v_p)=\zeta\left(\widehat{\grave{A}}(p)(v_p)\right)(p)$

and one can compute its Lie derivative $\pounds_W\widehat{\grave{A}}\in\Lambda_1(TP_U;T_1U)$ along a vector field $W\in\mathfrak{X}(TP_U)$:

$\pounds_W\widehat{\grave{A}}=\frac{d}{ds}\Phi_W(s,p)^*\widehat{\grave{A}}(p)|_{s=0}\in\Lambda_1(TP_U;T_1U)$

To the Lie derivative is associated the fundamental vector field $\zeta\left(\pounds_W\widehat{\grave{A}}(v_p)\right)(p)$, 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 :

$\zeta\left(\pounds_W\widehat{\grave{A}}(v_p)\right)(p)\to\zeta\left(\widetilde{\pounds_W\widehat{\grave{A}}(v_p)}\right)(p)=\rho'_p(p,\chi(m))\,\zeta\left(\pounds_W\widehat{\grave{A}}(v_p)\right)(p)$

This is a change of gauge :
$\widetilde{\mathbf{p}}=\widetilde{\varphi}_U(m,1)=\varphi_U(m,\chi(m))=\mathbf{p}\cdot\chi(m)$
the measure changes as :

$\zeta\left(\widetilde{\pounds_W\widehat{\grave{A}}(v_p)}\right)(p)=Ad_{\chi(m)}\zeta\left(\pounds_W\widehat{\grave{A}}(v_p)\right)(p(m))$

So $\zeta\left(\pounds_W\widehat{\grave{A}}(v_p)\right)$ 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 $\mathbf{p}(m)^*\pounds\widehat{\grave{A}}\in\Lambda_1(TM;P_U[T_1U,Ad])$

The strength of the field

However to define a derivation we want to keep the link with the base M. To do this :

- we use a given section $\mathbf{p}\in\mathfrak{X}(P_U)$ to go from M to P_U : $\mathbf{p}(m)\in P_U$.
- we lift a vector field V (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 $\grave{\mathbf{A}}(p)(W)=0$. The affine parameter s is the same along the integral curves of V,W.

The Lie derivative of the tensor $\widehat{\grave{A}}$ along $V,\mathbf{p}$ is then :
$\pounds_{\chi_L(V)}\widehat{\grave{A}}(p(m))=\frac{d}{ds}\Phi_{\chi_L(v)}(s,p(m))^*\widehat{\grave{A}}(p(m))|_{s=0}$
$\mathbf{p}(m)^*\widehat{\grave{A}}(p(m))=\grave{A}(m)$
The definition is then consistent and one can go from M to P_U.
The Lie derivative and the exterior differential are related (Maths.1531) :
$\pounds_W\widehat{\grave{A}}=d\left(i_W\widehat{\grave{A}}\right)+i_Wd\widehat{\grave{A}}$ where d is the exterior differential on M.
but, because $W=\chi_L(V)$ is horizontal :
$i_W\widehat{\grave{A}}=0$
$\chi_*(W)=W$
$\Rightarrow\pounds_W\widehat{\grave{A}}=i_Wd\widehat{\grave{A}}=i_W\chi^*d\grave{A}$

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 result holds for any vector field $V \in \mathfrak{X}(TM)$ and the **strength of the field** is defined as :

$$\mathcal{F}_A(m) = -\mathbf{p}^*(m)\,\pounds\widehat{\grave{A}} = -\mathbf{p}^*(m)\,\chi^* d\grave{A} \in \Lambda_2(M; T_1U) \tag{5.26}$$

with the standard gauge $\mathbf{p}(m) = \varphi_U(m,1)$.

It has the following expression :

$$\mathcal{F}_A = \sum_{a=1}^{m}\left(d\left(\sum_{\alpha=0}^{3}\grave{A}^a_\alpha d\xi^\alpha\right) + \sum_{\alpha\beta}\left[\grave{A}_\alpha, \grave{A}_\beta\right] d\xi^\alpha \wedge d\xi^\beta\right) \otimes \overrightarrow{\theta}_a \tag{5.27}$$

where d is the exterior differential on TM and $[]$ is the bracket in T_1U.

Equivalently with ordered indices :

$$\mathcal{F}_A = \sum_{a=1}^{m}\sum_{\{\alpha,\beta\}}\left(\mathcal{F}^a_{A\alpha\beta} d\xi^\alpha \wedge d\xi^\beta\right) \otimes \overrightarrow{\theta}_a \in \Lambda_2(M; T_1U) \tag{5.28}$$

and in components :

$$\mathcal{F}^a_{A\alpha\beta} = \partial_\alpha \grave{A}^a_\beta - \partial_\beta \grave{A}^a_\alpha + 2\left[\grave{A}_\alpha, \grave{A}_\beta\right]^a \tag{5.29}$$

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 $\left(\overrightarrow{\theta}_a\right)_{a=1}^{m}$) the group U acts through the map Ad. In a change of gauge $\mathcal{F}_A$ changes as :

$$\begin{gathered}\mathbf{p}_U(m) = \varphi_{P_U}(m,1) \to \widetilde{\mathbf{p}}_U(m) = \mathbf{p}_U(m)\cdot\varkappa(m)^{-1} : \\ \mathcal{F}_{A\alpha\beta} \to \widetilde{\mathcal{F}}_{A\alpha\beta}(m) = Ad_{\varkappa(m)}\mathcal{F}_{A\alpha\beta}\end{gathered} \tag{5.30}$$

so that $\mathcal{F}_A$ can be seen as a 2 form on TM valued in the **adjoint bundle** $P_U[T_1U, Ad]$. This gives a more geometrical meaning to the concept, and we will see that these relations are crucial in the definition of the lagrangian.

Connection, potential and strength

The two quantities come from the connection, but they have different physical meaning and mathematical properties.

The potential is not a geometric quantity, this is a function (and not a tensor) defined on M and valued in T_1U, the component of $\widehat{\grave{A}}$ in the standard gauge $\mathbf{p}(m) = \varphi_U(m,1)$. The strength $\mathcal{F}$ is a tensor, a 2 form on M valued in the Lie algebra. It is a special derivative of the connection. It can be computed from the potential, but the converse is not true. This is a classic issue : if $\mathcal{F}$ is the strength of the potential $\grave{A}$ then $\grave{A}+H$ will provide the same strength $\mathcal{F}$ if $dH + 2\sum_{\alpha,\beta}[H_\alpha, H_\beta] = 0$. In Electrodynamics this issue is solved by imposing additional constraints to the potential, using the "gauge freedom". So $\mathcal{F}$ can be considered as a different variable.

The question which arises is then of the choice of the "right variable" in a model.

In a Theory based upon fields existing everywhere and propagating on one hand, and particles located at a geometric point on the other hand, the vacuum exists almost everywhere. In the vacuum the mechanism at work is the interaction of the field with itself, which is represented through the lagrangian. As we will see in the next chapter, in order to be consistent, the lagrangian must be defined through $\mathcal{F}_{\alpha\beta}$ and not the potential or its derivatives. So there is a dominant variable, which is the value of the field in the vacuum, and this variable is $\mathcal{F}$. Moreover, this variable must be continuous because the propagation in the vacuum by self interaction assumes continuity. $\mathcal{F}$ is a continuous variable.

A field is measured by its impact on known particles. Particles are never immobile, they travel on their world line, the value of the connection changes and the measure of the field is the measure of the *variation* of the connection along the trajectory of particles. So eventually the measure of a field is given by the Lie derivative of the connection, that is by its strength $\mathcal{F}$. This is obvious in the usual expression of the Lorentz law, using the electric and magnetic fields, which are components of $\mathcal{F}_{EM}$.

To sum up : the potential is used to represent the interaction of the field with particles. The strength $\mathcal{F}$ is used to represent the propagation of the field.

Curvature

There is another introduction of the same concept, through the curvature, which is more usual but less immediate.

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 : $\Omega (p) = \zeta \left(\widehat{\Omega} (p) \right) (p)$

$\widehat{\Omega} \in \Lambda_2 (TP_U; T_1 U) : \widehat{\Omega} (p) = -Ad_{g^{-1}} \left(\sum_{a=1}^m \sum_{\alpha,\beta=0}^3 \left(\partial_\alpha \grave{A}_\beta^a + \left[\grave{A}_\alpha, \grave{A}_\beta \right] \right) \right) dm^\alpha \wedge dm^\beta \otimes \overrightarrow{\theta}_a$

where the bracket $\left[\grave{A}_\alpha, \grave{A}_\beta \right]$ is the bracket in the Lie algebra $T_1 U$.

For any r form ϖ on TP_U valued in a fixed vector space the exterior covariant derivative associated to the connection is the map :

$\nabla_e : \Lambda_r (TP_U; F) \to \Lambda_{r+1} (TP_U; F) :: \nabla_e \omega = \chi^* (d\omega)$

where $d\omega$ is the exterior differential on TM (the components along dg^a have vanished).

$\widehat{\Omega} = \nabla_e \widehat{\grave{A}}$

$\mathcal{F}_A$ 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 : $\nabla_e \widehat{\Omega} = 0$.

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 $T_1 Spin(3,1)$ which reads with the basis $(\overrightarrow{\kappa}_a)_{a=1}^6$:

$$\begin{gathered} \mathcal{F}_G = \sum_{a=1}^6 \left(dG^a + \sum_{\alpha\beta=0}^3 [G_\alpha, G_\beta]^a d\xi^\alpha \wedge d\xi^\beta \right) \otimes \overrightarrow{\kappa}_a \\ \mathcal{F}_G = \sum_{a=1}^6 \sum_{\{\alpha,\beta\}=0}^3 \mathcal{F}_{G\alpha\beta}^a d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\kappa}_a \\ \mathcal{F}_G = \sum_{a=1}^6 \sum_{\{\alpha,\beta\}=0}^3 \left(\partial_\alpha G_\beta^a - \partial_\beta G_\alpha^a + 2 [G_\alpha, G_\beta]^a \right) d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\kappa}_a \end{gathered} \tag{5.31}$$

where d is the exterior differential on TM and $[]$ is the bracket in $T_1 Spin(3,1)$.[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,
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 :
$a = 1,2,3 : \mathcal{F}^a_{G\alpha\beta} = \mathcal{F}^a_{r\alpha\beta}$
$a = 4,5,6 : \mathcal{F}^a_{G\alpha\beta} = \mathcal{F}^a_{w\alpha\beta}$
with the signature (3,1) :

$$\begin{array}{c} \mathcal{F}_G = \sum_{\{\alpha,\beta\}=0}^{3} \upsilon\left(\mathcal{F}_{r\alpha\beta}, \mathcal{F}_{w\alpha\beta}\right) d\xi^\alpha \wedge d\xi^\beta \\ \mathcal{F}_{r\alpha\beta} = \upsilon\left(\partial_\alpha G_{r\beta} - \partial_\beta G_{r\alpha} + 2\left(j\left(G_{r\alpha}\right)G_{r\beta} - j\left(G_{w\alpha}\right)G_{w\beta}\right), 0\right) \\ \mathcal{F}_{w\alpha\beta} = \upsilon\left(0, \partial_\alpha G_{w\beta} - \partial_\beta G_{w\alpha} + 2\left(j\left(G_{w\alpha}\right)G_{r\beta} + j\left(G_{r\alpha}\right)G_{w\beta}\right)\right) \end{array} \tag{5.32}$$

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 $Spin\,(3,1)$ acts through the map $\mathbf{Ad}$, and the action is given by 6×6 matrices seen previously. In a change of gauge on the principal bundle the strength changes as :

$$\begin{array}{c} \mathbf{p}_G(m) = \varphi_G(m,1) \to \widetilde{\mathbf{p}}_G(m) = \mathbf{p}_G(m) \cdot s(m)^{-1} : \\ \mathcal{F}_{G\alpha\beta} \to \widetilde{\mathcal{F}}_{G\alpha\beta}(m) = \mathbf{Ad}_{s(m)} \mathcal{F}_{G\alpha\beta} \\ \upsilon\left(\widetilde{\mathcal{F}}_{r\alpha\beta}, \widetilde{\mathcal{F}}_{w\alpha\beta}\right) = \mathbf{Ad}_{s(m)} \upsilon\left(\mathcal{F}_{r\alpha\beta}, \mathcal{F}_{w\alpha\beta}\right) \end{array} \tag{5.33}$$

and the strength can be seen as valued in the adjoint bundle $P_G\left[T_1 Spin(3,1), \mathbf{Ad}\right]$.
It is convenient to use the complex notation :
$a = 1,2,3$:
$G^a_\beta = G^a_{r\beta} + iG^a_{w\beta}$
$\mathcal{F}^a_{G\alpha\beta} = \mathcal{F}^a_{r\alpha\beta} + i\mathcal{F}^a_{w\alpha\beta}$

$$\mathcal{F}_{G\alpha\beta} = \partial_\alpha G_\beta - \partial_\beta G_\alpha + 2j\left(G_\alpha\right)G_\beta \tag{5.34}$$

5.4.2 Algebra of two forms

Computations with two-forms are an arduous process. We will use some notations and tools which make it easier.

Rotational and transversal components

The first tool is based on the decomposition of any scalar two form according to its components.

A second order tensor has 16 components, a two form only 6. 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. It is convenient to use a precise order of the indices. One can always write :
$\mathcal{F} = \mathcal{F}^r + \mathcal{F}^w$
with

[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.

$\mathcal{F}^r = \mathcal{F}_{32} d\xi^3 \wedge d\xi^2 + \mathcal{F}_{13} d\xi^1 \wedge d\xi^3 + \mathcal{F}_{21} d\xi^2 \wedge d\xi^1$
$\mathcal{F}^w = \mathcal{F}_{01} d\xi^0 \wedge d\xi^1 + \mathcal{F}_{02} d\xi^0 \wedge d\xi^2 + \mathcal{F}_{03} d\xi^0 \wedge d\xi^3$
and we will denote the 1×3 row matrices :

$$[\mathcal{F}^r] = \left[\begin{array}{ccc} \mathcal{F}_{32} & \mathcal{F}_{13} & \mathcal{F}_{21} \end{array} \right]; [\mathcal{F}^w] = \left[\begin{array}{ccc} \mathcal{F}_{01} & \mathcal{F}_{02} & \mathcal{F}_{03} \end{array} \right] \tag{5.35}$$

This is very similar to what is done with the EM field : one distinguishes an electric and a magnetic field, which are represented by orthogonal vectors. Notice that we have the same ordering as in the Lie algebra $T_1 Spin(3,1)$. This is just to be consistent and easier to use.

With this notation it is easy to write the usual operations in a matrix form.
$\mathcal{F} = \mathcal{F}^r + \mathcal{F}^w; K = K^r + K^w$:
$\mathcal{F} \wedge K = -\left([\mathcal{F}^r] [K^w]^t + [\mathcal{F}^w] [K^r]^t \right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

Any 2-form can also be written in matrix form where the indices α, β are the rows and columns of the matrix :

$$[\mathcal{F}_{\alpha\beta}]^{\alpha=0...3}_{\beta=0...3} = \begin{bmatrix} 0 & \mathcal{F}_{01} & \mathcal{F}_{02} & \mathcal{F}_{03} \\ \mathcal{F}_{10} & 0 & \mathcal{F}_{12} & \mathcal{F}_{13} \\ \mathcal{F}_{20} & \mathcal{F}_{21} & 0 & \mathcal{F}_{23} \\ \mathcal{F}_{30} & \mathcal{F}_{31} & \mathcal{F}_{32} & 0 \end{bmatrix}_{4\times4}$$

$$[\mathcal{F}] = \left[\begin{array}{cccc} 0 & [\mathcal{F}^w]_1 & [\mathcal{F}^w]_2 & [\mathcal{F}^w]_3 \\ -[\mathcal{F}^w]_1 & 0 & -[\mathcal{F}^r]_3 & [\mathcal{F}^r]_1 \\ -[\mathcal{F}^w]_2 & [\mathcal{F}^r]_3 & 0 & -[\mathcal{F}^r]_2 \\ -[\mathcal{F}^w]_3 & -[\mathcal{F}^r]_1 & [\mathcal{F}^r]_2 & 0 \end{array} \right] = \left[\begin{array}{cc} 0 & [\mathcal{F}^w]_{1\times3} \\ -([\mathcal{F}^w])^t_{3\times1} & j([\mathcal{F}^r])_{3\times3} \end{array} \right]$$

$[\mathcal{F}]^t = -[\mathcal{F}]$

Impact of a change of chart

In a change of chart the holonomic basis $\partial\xi_\alpha \to \partial\eta_\alpha$ with the jacobian $[J] = \left[\frac{\partial\eta^\alpha}{\partial\xi^\beta}\right]$. The components of a 2 form change as :

$\mathcal{F}_{\alpha\beta} \to \widetilde{\mathcal{F}}_{\alpha\beta} = \sum_{\{\lambda\mu\}=0}^3 \mathcal{F}_{\lambda\mu} \det [K]^{\{\lambda\mu\}}_{\{\alpha\beta\}}$
where $[K]$ is the inverse of the jacobian $[K] = [J]^{-1}$
$\widetilde{\mathcal{F}}_{\alpha\beta} = \sum_{\{\lambda\mu\}=0}^3 \mathcal{F}_{\lambda\mu} \left(K^\lambda_\alpha K^\mu_\beta - K^\lambda_\beta K^\mu_\alpha \right)$
$= \frac{1}{2} \sum_{\lambda\mu=0}^3 \mathcal{F}_{\lambda\mu} \left(K^\lambda_\alpha K^\mu_\beta - K^\lambda_\beta K^\mu_\alpha \right) = \frac{1}{2} \sum_{\gamma\eta=0}^3 [K]^\lambda_\alpha [\mathcal{F}]^\lambda_\mu [K]^\mu_\beta - [K]^\lambda_\beta [\mathcal{F}]^\lambda_\mu [K]^\mu_\alpha$
$\widetilde{\mathcal{F}}_{\alpha\beta} = \frac{1}{2} \left([K]^t [\mathcal{F}] [K] \right)^\alpha_\beta - \left([K]^t [\mathcal{F}] [K] \right)^\beta_\alpha = \frac{1}{2} \left(\left([K]^t [\mathcal{F}] [K] \right) - \left([K]^t [\mathcal{F}] [K] \right)^t \right)^\alpha_\beta$
$= \left([K]^t [\mathcal{F}] [K] \right)^\alpha_\beta$
$\left[\widetilde{\mathcal{F}}\right] = [K]^t [\mathcal{F}] [K]$

We will meet several times this kind of formula, so it is useful to give a more detailed computation.

Using the notation : $[K] = \left[\begin{array}{cc} K^0_0 & [K^0]_{1\times3} \\ [K_0]_{3\times1} & [k]_{3\times3} \end{array} \right]$

$$\left[\begin{array}{cc} 0 & \left[\widetilde{\mathcal{F}}^w\right] \\ -\left[\widetilde{\mathcal{F}}^w\right]^t & j\left(\left[\widetilde{\mathcal{F}}^r\right]\right) \end{array} \right] = \left[\begin{array}{cc} K^0_0 & [K_0]^t \\ [K^0]^t & [k]^t \end{array} \right] \left[\begin{array}{cc} 0 & [\mathcal{F}^w] \\ -[\mathcal{F}^w]^t & j([\mathcal{F}^r]) \end{array} \right] \left[\begin{array}{cc} K^0_0 & [K^0] \\ [K_0] & [k] \end{array} \right]$$

then :
$K^0_0 [\mathcal{F}^w] [K_0] - K^0_0 [K_0]^t [\mathcal{F}^w]^t + [K_0]^t j([\mathcal{F}^r]) [K_0] = 0$
$\left[\widetilde{\mathcal{F}}^w\right] = [\mathcal{F}^w] K^0_0 [k] - [K_0]^t_{1\times3} [\mathcal{F}^w]^t_{3\times1} [K^0]_{1\times3} + [K_0]^t j([\mathcal{F}^r]) [k]$
$= [\mathcal{F}^w] K^0_0 [k] - [\mathcal{F}^w] [K_0] [K^0] - [\mathcal{F}^r] j([K_0]) [k]$

$j\left(\left[\widetilde{\mathcal{F}}^r\right]\right) = \left[K^0\right]^t_{3\times1}[\mathcal{F}^w]_{1\times3}[k]_{3\times3} - [k]^t[\mathcal{F}^w]^t_{3\times1}\left[K^0\right]_{1\times3} + [k]^t j([\mathcal{F}^r])[k]$
$= \left[K^0\right]^t([\mathcal{F}^w][k]) - ([\mathcal{F}^w][k])^t\left[K^0\right] + [k]^t j([\mathcal{F}^r])[k]$
$= j([\mathcal{F}^w][k])\, j\left(\left[K^0\right]\right) + ([\mathcal{F}^w][k])\left[K^0\right]^t - j\left(\left[K^0\right]\right) j([\mathcal{F}^w][k]) - ([\mathcal{F}^w][k])\left[K^0\right]^t + [k]^t j([\mathcal{F}^r])[k]$
$j([\mathcal{F}^w][k])\, j\left(\left[K^0\right]\right) - j\left(\left[K^0\right]\right) j([\mathcal{F}^w][k]) = j\left([\mathcal{F}^w][k]\, j\left(\left[K^0\right]\right)\right)$
$[k]^t j([\mathcal{F}^r])[k] = j\left([r]^t\left[k^{-1}\right]^t\right)\det k$
with $j\left([r]^t[M]\right) = \left([M]^{-1}\right) j(r)\left([M]^{-1}\right)^t \det M$
$j\left(\left[\widetilde{\mathcal{F}}^r\right]\right) = j\left([\mathcal{F}^w][k]\, j\left(\left[K^0\right]\right)\right) + j\left([\mathcal{F}^r]\left([k]^{-1}\right)^t\right)\det k$
$= j\left([\mathcal{F}^w][k]\, j\left(\left[K^0\right]\right) + [\mathcal{F}^r]\left([k]^{-1}\right)^t\det k\right)$
$\left[\widetilde{\mathcal{F}}^r\right] = [\mathcal{F}^r]\left([k]^{-1}\right)^t\det k + [\mathcal{F}^w][k]\, j\left(\left[K^0\right]\right)$

$$\begin{gathered}\left[\widetilde{\mathcal{F}}\right] = [K]^t[\mathcal{F}][K]\\ \begin{bmatrix}\left[\widetilde{\mathcal{F}}^r\right] & \left[\widetilde{\mathcal{F}}^w\right]\end{bmatrix} = \begin{bmatrix}[\mathcal{F}^r] & [\mathcal{F}^w]\end{bmatrix}[L_K]\\ [L_K] = \begin{bmatrix}\left([k]^{-1}\right)^t\det k & -j([K_0])[k]\\ [k]\, j\left(\left[K^0\right]\right) & K^0_0[k] - [K_0]\left[K^0\right]\end{bmatrix}\end{gathered} \tag{5.36}$$

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 :
$\begin{bmatrix}\left[\widetilde{\mathcal{F}}^r\right] & \left[\widetilde{\mathcal{F}}^w\right]\end{bmatrix} = \begin{bmatrix}[\mathcal{F}^r] & [\mathcal{F}^w]\end{bmatrix}\begin{bmatrix}(\det k)\left[k^{-1}\right]^t & 0\\ 0 & [k]\end{bmatrix}$
We have a generalization of these relations, which is useful in computations.

Any linear map on the vectorial space of 2 forms which is the extension of a linear map on the dual bundle TM^* (Maths.426) is expressed as in a change of chart :
$\widetilde{\mathcal{F}}_{\alpha\beta} = \sum_{\lambda\mu=0}^3 \mathcal{F}_{\lambda\mu}K^\lambda_\alpha K^\mu_\beta = \sum_{\{\lambda\mu\}=0}^3 \mathcal{F}_{\lambda\mu}\left(K^\lambda_\alpha K^\mu_\beta - K^\lambda_\beta K^\mu_\alpha\right) = \left([K]^t[\mathcal{F}][K]\right)^\alpha_\beta$
$\left[\widetilde{\mathcal{F}}\right] = [K]^t[\mathcal{F}][K]$
So, if we have the product of two linear maps :
$\left[\widetilde{\mathcal{F}}_1\right] = [K_1]^t[\mathcal{F}][K_1]$
$\left[\widetilde{\mathcal{F}}_2\right] = [K_2]^t\left[\widetilde{\mathcal{F}}_1\right][K_2] = [K_2]^t[K_1]^t[\mathcal{F}][K_1][K_2]$
then :

$$[L_{K_1K_2}] = [L_{K_1}][L_{K_2}] \tag{5.37}$$

The inverse operation $\widetilde{\mathcal{F}}_{\alpha\beta} \to \mathcal{F}_{\alpha\beta}$ is given by :
$\left[\widetilde{\mathcal{F}}\right] = [K]^t[\mathcal{F}][K] \Leftrightarrow [\mathcal{F}] = \left([K]^{-1}\right)^t\left[\widetilde{\mathcal{F}}\right][K]^{-1}$ with $[K]^{-1}$ so that $[L_K]^{-1} = [L_{K^{-1}}]$.
For instance, the inverse operation in a change of chart is given by $[K]^{-1}$ and $[L_{K^{-1}}] = [L_K]^{-1}$.

The 2 form $\mathcal{F}^* = \sum_{\{\alpha\beta\}}\mathcal{F}^{\alpha\beta}d\xi^\alpha\wedge d\xi^\beta$ defined by lifting the indices with the metric :
$\mathcal{F}^{\alpha\beta} = \sum_{\lambda\mu=0}^3 g^{\alpha\lambda}g^{\beta\mu}\mathcal{F}_{\lambda\mu} = \sum_{\lambda\mu=0}^3\left[g^{-1}\right]^\alpha_\lambda[\mathcal{F}]^\lambda_\mu\left[g^{-1}\right]^\mu_\beta = \left(\left[g^{-1}\right][\mathcal{F}]\left[g^{-1}\right]\right)^\alpha_\beta$
has for matrix :

$$[\mathcal{F}^*] = \left[\mathcal{F}^{\alpha\beta}\right]^{\alpha=0..3}_{\beta=0..3} = [g]^{-1}[\mathcal{F}][g]^{-1} \tag{5.38}$$

and :
$\begin{bmatrix}\mathcal{F}^{*r} & \mathcal{F}^{*w}\end{bmatrix} = \begin{bmatrix}\mathcal{F}^r & \mathcal{F}^w\end{bmatrix}[L_{g^{-1}}]$

Expression in the orthonormal basis

Any 2 form $\mathcal{F} \in \Lambda_2(M;\mathbb{R})$ can be expressed in the orthonormal basis $\left(\varepsilon^i\right)_{i=0}^3$:

$\mathcal{F} = \sum_{\alpha\beta} \mathcal{F}_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta = \sum_{ij} F_{ij} \varepsilon^i \wedge \varepsilon^j$ with $\varepsilon^i = \sum_{\alpha=0}^3 P'^i_\alpha d\xi^\alpha$

with the obvious notations $[F^r], [F^w]$. This is similar to a change of chart with $[P] = [J], [P'] = [K]$

$\left[\begin{array}{cc} [F^r] & [F^w] \end{array}\right] = \left[\begin{array}{cc} [\mathcal{F}^r] & [\mathcal{F}^w] \end{array}\right] [L_{P'}] \Leftrightarrow \left[\begin{array}{cc} [\mathcal{F}^r] & [\mathcal{F}^w] \end{array}\right] = \left[\begin{array}{cc} [F^r] & [F^w] \end{array}\right] [L_P]$

Extension to fiber bundles

The notation can be extended to the strength of the field, 2-forms valued in the Lie algebras.

Gravitational field

$[\mathcal{F}^r_G]_{6\times 3} = [\mathcal{F}^r]^{a=1..6}$

$[\mathcal{F}^w_G]_{6\times 3} = [\mathcal{F}^w]^{a=1..6}$

$[\mathcal{F}]_{6\times 6} = \left[\begin{array}{cc} \mathcal{F}^r_r & \mathcal{F}^w_r \\ \mathcal{F}^r_w & \mathcal{F}^w_w \end{array}\right] = \left[\mathcal{F}^a_{G\alpha\beta}\right]$

with the 3×3 matrices :

$[\mathcal{F}^r_r]_{3\times 3} = \left[\begin{array}{ccc} \mathcal{F}^1_{G32} & \mathcal{F}^1_{G13} & \mathcal{F}^1_{G21} \\ \mathcal{F}^2_{G32} & \mathcal{F}^2_{G13} & \mathcal{F}^2_{G21} \\ \mathcal{F}^3_{G32} & \mathcal{F}^3_{G13} & \mathcal{F}^3_{G21} \end{array}\right]$

$[\mathcal{F}^w_r]_{3\times 3} = \left[\begin{array}{ccc} \mathcal{F}^1_{G01} & \mathcal{F}^1_{G02} & \mathcal{F}^1_{G03} \\ \mathcal{F}^2_{G01} & \mathcal{F}^2_{G02} & \mathcal{F}^2_{G02} \\ \mathcal{F}^3_{G01} & \mathcal{F}^3_{G02} & \mathcal{F}^3_{G03} \end{array}\right]$

$[\mathcal{F}^r_w]_{3\times 3} = \left[\begin{array}{ccc} \mathcal{F}^4_{G32} & \mathcal{F}^4_{G13} & \mathcal{F}^4_{G21} \\ \mathcal{F}^5_{G32} & \mathcal{F}^5_{G13} & \mathcal{F}^5_{G21} \\ \mathcal{F}^6_{G32} & \mathcal{F}^6_{G13} & \mathcal{F}^6_{G21} \end{array}\right]$

$[\mathcal{F}^w_w]_{3\times 3} = \left[\begin{array}{ccc} \mathcal{F}^4_{G01} & \mathcal{F}^4_{G02} & \mathcal{F}^4_{G03} \\ \mathcal{F}^5_{G01} & \mathcal{F}^5_{G02} & \mathcal{F}^5_{G03} \\ \mathcal{F}^6_{G01} & \mathcal{F}^6_{G02} & \mathcal{F}^6_{G03} \end{array}\right]$

and in complex notation :

$[\mathcal{F}^r_G] = [\mathcal{F}^r_w] + i\,[\mathcal{F}^r_w]$

$[\mathcal{F}^w_G] = [\mathcal{F}^w_w] + i\,[\mathcal{F}^w_w]$

The strength of the gravitational field reads then with the complex notation :

$[G]_{3\times 3} = \left[\begin{array}{ccc} G^1_1 + iG^4_1 & G^1_2 + iG^4_2 & G^1_3 + iG^4_3 \\ G^2_1 + iG^5_1 & G^2_2 + iG^5_2 & G^2_3 + iG^5_3 \\ G^3_1 + iG^6_1 & G^3_2 + iG^6_2 & G^3_3 + iG^6_3 \end{array}\right]$

$[G_0]_{3\times 1} = \left[\begin{array}{c} G^1_0 + iG^4_0 \\ G^2_0 + iG^5_0 \\ G^3_0 + iG^6_0 \end{array}\right]$

$\mathcal{F}_{G\alpha\beta} = \partial_\alpha G_\beta - \partial_\beta G_\alpha + 2j(G_\alpha) G_\beta = (dG)_{\alpha\beta} + 2j(G_\alpha) G_\beta$

$[\mathcal{F}^r_G] = [dG^r] + 2\left[\begin{array}{ccc} j(G_3) G_2 & j(G_1) G_3 & j(G_2) G_1 \end{array}\right] = [dG^r] - 2(\det[G])[G]^{-1}$

with $\left[\begin{array}{ccc} j([M]_2)[M]_3 & j([M]_3)[M]_1 & j([M]_1)[M]_2 \end{array}\right] = (\det M)\left[M^{-1}\right]^t$

$[dG^r] = \left[\begin{array}{ccc} \partial_3 G_2 - \partial_2 G_3 & \partial_1 G_3 - \partial_3 G_1 & \partial_2 G_1 - \partial_1 G_2 \end{array}\right]_{3\times 3}$

$[\mathcal{F}^w_G] = [dG^w] + 2\left[\begin{array}{ccc} j(G_0) G_1 & j(G_0) G_2 & j(G_0) G_3 \end{array}\right] = [dG^w] + 2[j(G_0)][G]$

$[dG^w] = \left[\begin{array}{ccc} \partial_0 G_1 - \partial_1 G_0 & \partial_0 G_2 - \partial_2 G_0 & \partial_0 G_3 - \partial_3 G_0 \end{array}\right]_{3\times 3}$

$$\left[\begin{array}{cc} \mathcal{F}^r_G & \mathcal{F}^w_G \end{array}\right]_{3\times 6} = \left[\begin{array}{cc} [dG^r] - 2(\det[G])[G]^{-1} & [dG^w] + 2[j(G_0)][G] \end{array}\right] \quad (5.39)$$

Other fields :

$[\mathcal{F}]_{m\times 6}=\left[\begin{array}{cc}\mathcal{F}_A^r & \mathcal{F}_A^w\end{array}\right]=\left[\mathcal{F}_{A\alpha\beta}^a\right]$

with $m\times 3$ matrices :

$[\mathcal{F}_A^r]_{m\times 3}=\left[\begin{array}{ccc}\mathcal{F}_{A32}^1 & \mathcal{F}_{A13}^1 & \mathcal{F}_{A21}^1\\ ... & ... & ...\\ \mathcal{F}_{A32}^m & \mathcal{F}_{G13}^m & \mathcal{F}_{G21}^m\end{array}\right]$

$[\mathcal{F}_A^w]_{m\times 3}=\left[\begin{array}{ccc}\mathcal{F}_{A01}^1 & \mathcal{F}_{A02}^1 & \mathcal{F}_{A03}^1\\ ... & ... & ...\\ \mathcal{F}_{A01}^m & \mathcal{F}_{A02}^m & \mathcal{F}_{A03}^m\end{array}\right]$

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.19.1.2). G_r is a bilinear symmetric form, which does not depend on a chart, is non degenerate and definite positive if g is Riemannian.

$G_r(\lambda,\mu)=\sum_{\{\alpha_1...\alpha_r\}\{\beta_1...\beta_r\}}\lambda_{\alpha_1...\alpha_r}\mu_{\beta_1...\beta_r}\det\left[g^{-1}\right]^{\{\alpha_1...\alpha_r\}}_{\{\beta_1...\beta_r\}}$

So for 2 forms :

$G_2(\mathcal{F},K)=\sum_{\{\alpha\beta\}\{\lambda\mu\}}\mathcal{F}_{\alpha\beta}K_{\lambda\mu}\det\left[g^{-1}\right]^{\{\alpha\beta\}}_{\{\lambda\mu\}}=\sum_{\{\alpha\beta\}\{\lambda\mu\}}\mathcal{F}_{\alpha\beta}K_{\lambda\mu}\left(h_{\alpha\lambda}h_{\beta\mu}-h_{\alpha\mu}h_{\beta\lambda}\right)$

with denoting $[h]=[g]^{-1}$

A straightforward computation gives with the matricial notation for $[\mathcal{F}],[K]$:

$\sum_{\alpha\beta\lambda\mu}\mathcal{F}_{\alpha\beta}K_{\lambda\mu}\left(h_{\alpha\lambda}h_{\beta\mu}-h_{\alpha\mu}h_{\beta\lambda}\right)$

$=4\sum_{\{\alpha\beta\}\{\lambda\mu\}}\mathcal{F}_{\alpha\beta}K_{\lambda\mu}\left(h_{\alpha\lambda}h_{\beta\mu}-h_{\alpha\mu}h_{\beta\lambda}\right)$

$=\sum_{\alpha\beta\lambda\mu}[\mathcal{F}]^\alpha_\beta[K]^\lambda_\mu\left([h]^\alpha_\lambda[h]^\beta_\mu-[h]^\alpha_\mu[h]^\beta_\lambda\right)$

$=\sum_{\alpha\beta\lambda\mu}-\left([h][K][h]\right)^\alpha_\beta[\mathcal{F}]^\beta_\alpha-[K]^\lambda_\mu\left([h][\mathcal{F}][h]\right)^\mu_\lambda$

$=-Tr\left([h][K][h][\mathcal{F}]\right)-Tr\left([K][h][\mathcal{F}][h]\right)$

$=-2Tr\left([\mathcal{F}][h][K][h]\right)$

$G_2(\mathcal{F},K)=-\frac{1}{2}Tr\left([\mathcal{F}][h][K][h]\right)$

The 2 form $[\mathcal{F}^*]=\left[\mathcal{F}^{\alpha\beta}\right]^{\alpha=0..3}_{\beta=0..3}$ has for matrix $[\mathcal{F}^*]=[h][\mathcal{F}][h]$ so $G_2(\mathcal{F},K)$ can also be written :

$$G_2(\mathcal{F},K)=-\frac{1}{2}Tr\left([\mathcal{F}][g]^{-1}[K][g]^{-1}\right)=-\frac{1}{2}Tr\left([\mathcal{F}][K^*]\right) \quad (5.40)$$

In the standard chart $[g]^{-1}=\left[\begin{array}{cc}-1 & 0\\ 0 & [g_3]^{-1}\end{array}\right]$

$G_2(\mathcal{F},\mathcal{F})=[\mathcal{F}^w][g_3]^{-1}[\mathcal{F}^w]^t+[\mathcal{F}^r][g_3][\mathcal{F}^r]\det[g_3]^{-1}$

and G_2 is definite positive : $G_2(\mathcal{F},\mathcal{F})=0\Leftrightarrow\mathcal{F}=0$

Hodge duality

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 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 :

$$\forall\lambda,\mu\in\Lambda_2(M;\mathbb{R}):*\lambda\wedge\mu=G_2(\lambda,\mu)\varpi_4=G_2(\mu,\lambda)\varpi_4 \quad (5.41)$$

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. Writing $\mathcal{F} = \mathcal{F}^r + \mathcal{F}^w$ then : $*\mathcal{F} = *\mathcal{F}^r + *\mathcal{F}^w$

$$\begin{array}{rl} *\mathcal{F}^r = & -\left(\mathcal{F}^{01}d\xi^3\wedge d\xi^2 + \mathcal{F}^{02}d\xi^1\wedge d\xi^3 + \mathcal{F}^{03}d\xi^2\wedge d\xi^1\right)\det P' \\ *\mathcal{F}^w = & -\left(\mathcal{F}^{32}d\xi^0\wedge d\xi^1 + \mathcal{F}^{13}d\xi^0\wedge d\xi^2 + \mathcal{F}^{21}d\xi^0\wedge d\xi^3\right)\det P' \\ *\mathcal{F}^{\alpha\beta} = & -\mathcal{F}^{\alpha\beta}\det P' = -\left(\sum_{\lambda\mu=0}^{3} g^{\alpha\lambda}g^{\beta\mu}\mathcal{F}_{\lambda\mu}\right)\det P' \end{array} \tag{5.42}$$

The components of the parts $*\mathcal{F}^r, *\mathcal{F}^w$ are *exchanged* and the indices are lifted 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 $d\xi^\alpha \wedge d\xi^\beta$:

We will use the 2 notations introduced previously.

i) With respect to the basis $d\xi^\alpha \wedge d\xi^\beta$ the decomposition of $*\mathcal{F}$ in the 2 components $*\mathcal{F}^r, *\mathcal{F}^w$

$$\begin{array}{c} *\mathcal{F} = *\mathcal{F}^r + *\mathcal{F}^w \\ \left[*\mathcal{F}^r\right] = \left[\begin{array}{ccc} *\mathcal{F}^{01} & *\mathcal{F}^{02} & *\mathcal{F}^{03} \end{array}\right] = -\left[\begin{array}{ccc} \mathcal{F}^{01} & \mathcal{F}^{02} & \mathcal{F}^{03} \end{array}\right](\det P') \\ \left[*\mathcal{F}^w\right] = \left[\begin{array}{ccc} *\mathcal{F}^{32} & *\mathcal{F}^{13} & *\mathcal{F}^{21} \end{array}\right] = -\left[\begin{array}{ccc} \mathcal{F}^{32} & \mathcal{F}^{13} & \mathcal{F}^{21} \end{array}\right](\det P') \end{array} \tag{5.43}$$

In this 1×4 matricial representation the column index refers to the index of the basis $d\xi^\alpha$.

ii) The matrix of the components $\left[*\mathcal{F}^{\alpha\beta}\right]_{\beta=0..3}^{\alpha=0..3}$

$$[*\mathcal{F}]_{4\times4} = \left[\begin{array}{cccc} 0 & *\mathcal{F}^{01} & *\mathcal{F}^{02} & *\mathcal{F}^{03} \\ -*\mathcal{F}^{01} & 0 & -*\mathcal{F}^{21} & *\mathcal{F}^{13} \\ -*\mathcal{F}^{02} & *\mathcal{F}^{21} & 0 & -*\mathcal{F}^{32} \\ -*\mathcal{F}^{03} & -*\mathcal{F}^{13} & *\mathcal{F}^{32} & 0 \end{array}\right]$$

In this 4×4 matricial representation the row and column refer to the indices α, β of $*\mathcal{F}^{\alpha\beta}$ as in $[\mathcal{F}]$

$$[*\mathcal{F}]_{4\times4} = \left[\begin{array}{cc} 0 & [*\mathcal{F}^r] \\ -[*\mathcal{F}^r] & j([*\mathcal{F}^w]) \end{array}\right]$$

The Hodge dual is easily computed using the matrix :

$$[\mathcal{F}^*] = \left[\mathcal{F}^{*\alpha\beta}\right] = \left[\begin{array}{cccc} 0 & \mathcal{F}^{01} & \mathcal{F}^{02} & \mathcal{F}^{03} \\ -\mathcal{F}^{01} & 0 & -\mathcal{F}^{21} & \mathcal{F}^{13} \\ -\mathcal{F}^{02} & \mathcal{F}^{21} & 0 & -\mathcal{F}^{32} \\ -\mathcal{F}^{03} & -\mathcal{F}^{13} & \mathcal{F}^{32} & 0 \end{array}\right] = [g]^{-1}[\mathcal{F}][g]^{-1}$$

So : $[*\mathcal{F}] = -[\mathcal{F}^*]\det P'$

$$[*\mathcal{F}] = \left[\begin{array}{cc} 0 & [*\mathcal{F}^r] \\ -[*\mathcal{F}^r]^t & j([*\mathcal{F}^w]) \end{array}\right] = -[\mathcal{F}^*]\det P' = -[g]^{-1}[\mathcal{F}][g]^{-1}\det P' \tag{5.44}$$

$\left[\begin{array}{cc} \mathcal{F}^{*r} & \mathcal{F}^{*w} \end{array}\right] = \left[\begin{array}{cc} \mathcal{F}^r & \mathcal{F}^w \end{array}\right][L_H]$

with $[H]_{3\times3} = \left[g^{\alpha\beta}\right] = [g]^{-1} = \left[\begin{array}{cc} [H]_0^0 & [H]^0 \\ [H]_0 & [h] \end{array}\right]$;

$[H] = [H]^t \Rightarrow [H]_{3\times1} = [H_0], [H]^t = \left[H^0\right] = [H_0]^t$

$[L_H] = \left[\begin{array}{cc} [h]^{-1}\det h & -j(H_0)[h] \\ [h]j(H_0) & H_0^0[h] - [H_0][H_0]^t \end{array}\right]$

so that :

$[\mathcal{F}^{*r}] = [\mathcal{F}^r][h]^{-1}\det h + [\mathcal{F}^w][h]j(H_0)$

$[\mathcal{F}^{*w}] = -[\mathcal{F}^r]j(H_0)[h] + [\mathcal{F}^w]\left(H_0^0[h] - [H_0][H_0]^t\right)$

$\left[\begin{array}{cc} [*\mathcal{F}^r] & [*\mathcal{F}^w] \end{array}\right] = -\left[\begin{array}{cc} [\mathcal{F}^{*w}] & [\mathcal{F}^{*r}] \end{array}\right]\det P'$

$[*\mathcal{F}^r] = \left([\mathcal{F}^r]j(H_0)[h] - [\mathcal{F}^w]\left(H_0^0[h] - [H_0][H_0]^t\right)\right)\det P'$

$[*\mathcal{F}^w] = -\left([\mathcal{F}^r] [h]^{-1} \det h + [\mathcal{F}^w] [h] j(H_0) \right) \det P'$

$\begin{bmatrix} [*\mathcal{F}^r] & [*\mathcal{F}^w] \end{bmatrix} = \begin{bmatrix} [\mathcal{F}^r] & [\mathcal{F}^w] \end{bmatrix} \begin{bmatrix} j(H_0)[h] & -[h]^{-1} \det h \\ -\left(H_0^0 [h] - [H_0][H_0]^t \right) & -[h] j(H_0) \end{bmatrix} \det P'$

$\begin{bmatrix} [*\mathcal{F}^r] & [*\mathcal{F}^w] \end{bmatrix} = \begin{bmatrix} [\mathcal{F}^r] & [\mathcal{F}^w] \end{bmatrix} \left[\widetilde{L}_{g^{-1}} \right] \det P'$

with $\left[\widetilde{L}_{g^{-1}} \right] = \begin{bmatrix} j(H_0)[h] & -[h]^{-1} \det h \\ -\left(H_0^0 [h] - [H_0][H_0]^t \right) & -[h] j(H_0) \end{bmatrix}$

or equivalently, to keep the same notation and properties as above with $\left[L_{g^{-1}} \right]$:

$[*\mathcal{F}^r] = -[\mathcal{F}^{*w}] \det P'$

$[*\mathcal{F}^w] = -[\mathcal{F}^{*r}] \det P'$

$\begin{bmatrix} [*\mathcal{F}^w] & [*\mathcal{F}^r] \end{bmatrix} = -\begin{bmatrix} [\mathcal{F}^{*r}] & [\mathcal{F}^{*w}] \end{bmatrix} \det P' = -\begin{bmatrix} \mathcal{F}^r & \mathcal{F}^w \end{bmatrix} [L_H] \det P'$

$$\begin{gathered} \begin{bmatrix} [*\mathcal{F}^w] & [*\mathcal{F}^r] \end{bmatrix} = -\begin{bmatrix} [\mathcal{F}^r] & [\mathcal{F}^w] \end{bmatrix} [L_H] \det P' \\ [L_H] = \begin{bmatrix} [h]^{-1} \det h & -j(H_0)[h] \\ [h] j(H_0) & H_0^0 [h] - [H_0][H_0]^t \end{bmatrix} \end{gathered} \tag{5.45}$$

In the standard chart : $H = 0, [H]_0^0 = -1 : [H] = \begin{bmatrix} -1 & 0 \\ 0 & [g_3]^{-1} \end{bmatrix}$

$[*\mathcal{F}^r] = [\mathcal{F}^w] [h] \det P'$

$[*\mathcal{F}^w] = -[\mathcal{F}^r] [h]^{-1} \det h \det P'$

$\det P' = \det Q'$

$\det h = \det [g_3]^{-1} = \det [Q] [Q]^t = (\det Q)^2$

$$\begin{gathered} [*\mathcal{F}^r] = [\mathcal{F}^w] [g_3]^{-1} \det Q' \\ [*\mathcal{F}^w] = -[\mathcal{F}^r] [g_3] \det Q \end{gathered} \tag{5.46}$$

For the forms $(\mathcal{F}_G^a)_{a=1..6}$, $(\mathcal{F}_A^a)_{a=1..m}$ we can compute the Hodge dual for each component $(\mathcal{F}_G^a)_{a=1..6}$, $(\mathcal{F}_A^a)_{a=1..m}$ using the formulas above :

$\begin{bmatrix} [*\mathcal{F}^{a,w}] & [*\mathcal{F}^{a,r}] \end{bmatrix} = -\begin{bmatrix} [\mathcal{F}^{a,r}] & [\mathcal{F}^{a,w}] \end{bmatrix} [L_H] \det P'$

For the gravitational field :

$\begin{bmatrix} [*\mathcal{F}_r^w] & [*\mathcal{F}_r^r] \end{bmatrix} = -\begin{bmatrix} [\mathcal{F}_r^r] & [\mathcal{F}_r^w] \end{bmatrix} [L_H] \det P'$

$\begin{bmatrix} [*\mathcal{F}_w^w] & [*\mathcal{F}_w^r] \end{bmatrix} = -\begin{bmatrix} [\mathcal{F}_w^r] & [\mathcal{F}_w^w] \end{bmatrix} [L_H] \det P'$

or in complex formalism :

$\begin{bmatrix} [*\mathcal{F}_G^w] & [*\mathcal{F}_G^r] \end{bmatrix} = -\begin{bmatrix} [\mathcal{F}_G^r] & [\mathcal{F}_G^w] \end{bmatrix} [L_H] \det P'$

and for the other fields :

$\begin{bmatrix} [*\mathcal{F}_A^w] & [*\mathcal{F}_A^r] \end{bmatrix} = -\begin{bmatrix} [\mathcal{F}_A^r] & [\mathcal{F}_A^w] \end{bmatrix} [L_H] \det P'$

For any r form we have : $**\lambda_r = -(-1)^{r(4-r)} \lambda$ so that for 2 forms the map : $* : \Lambda_2(M; \mathbb{R}) \to \Lambda_2(M; \mathbb{R}) :: *\lambda_2$ is such that : $**\lambda_2 = -\lambda_2$ so :

$\begin{bmatrix} [**\mathcal{F}^w] & [**\mathcal{F}^r] \end{bmatrix} = \begin{bmatrix} [*\mathcal{F}^r] & [*\mathcal{F}^w] \end{bmatrix} \left[\widetilde{L}_H \right] \det P'$

$\begin{bmatrix} [*\mathcal{F}^r] & [*\mathcal{F}^w] \end{bmatrix} = \begin{bmatrix} [\mathcal{F}^r] & [\mathcal{F}^w] \end{bmatrix} \left[\widetilde{L}_H \right] \det P'$

$\left[\widetilde{L}_H \right]^2 (\det P')^2 = -I_6$

$\left[\widetilde{L}_H \right]^{-1} = \left[\widetilde{L}_H \right] \det g$

Computing $\left[\widetilde{L}_H \right]^2 = -I_6 (\det P)^2$ we get the relation between the components of $[L_H]$:

$$[H]_0^0 = [H_0]^t [h]^{-1} [H_0] + (\det g) \det [h]^{-1} \tag{5.47}$$

Consider a change of chart with some matrix $[K]$ as defined above :

$\left[\begin{array}{cc}\left[\widetilde{\mathcal{F}}^r\right] & \left[\widetilde{\mathcal{F}}^w\right]\end{array}\right] = \left[\begin{array}{cc}[\mathcal{F}^r] & [\mathcal{F}^w]\end{array}\right][L_K]$

$\left[\begin{array}{cc}[*\mathcal{F}^w] & [*\mathcal{F}^r]\end{array}\right] = -\left[\begin{array}{cc}[\mathcal{F}^r] & [\mathcal{F}^w]\end{array}\right][L_H]\det P'$

$\left[\begin{array}{cc}\left[*\widetilde{\mathcal{F}}^w\right] & \left[*\widetilde{\mathcal{F}}^r\right]\end{array}\right] = -\left[\begin{array}{cc}\left[\widetilde{\mathcal{F}}^r\right] & \left[\widetilde{\mathcal{F}}^w\right]\end{array}\right]\left[L_{\widetilde{H}}\right]\det \widetilde{P}'$

The matrix of the metric changes as : $[g] \to [\widetilde{g}] = [K]^t [g][K]$ so

$[H] = [g]^{-1} \to \left[\widetilde{H}\right] = [K]^{-1}[H]\left([K]^{-1}\right)^t$

$[L_H] \to \left[L_{\widetilde{H}}\right] = [L_{K^{-1}}][L_H]\left[L_{(K^{-1})^t}\right]$

$\det[g] \to \det[g]\det[K]^2$

$(\det P')^2 = -\det[g] \to -\det[g]\det[K]^2$

$(\det P') \to \det[P']\det[K]$

$\left[\begin{array}{cc}\left[*\widetilde{\mathcal{F}}^w\right] & \left[*\widetilde{\mathcal{F}}^r\right]\end{array}\right] = -\left[\begin{array}{cc}\left[\widetilde{\mathcal{F}}^r\right] & \left[\widetilde{\mathcal{F}}^w\right]\end{array}\right][L_{K^{-1}}][L_H]\left[L_{(K^{-1})^t}\right]\det[P']\det[K]$

We can then compute another formula for the scalar product of 2 forms. Take any two scalar 2 forms $\mathcal{F}, K$ and their decomposition as above, a straightforward computation gives :

$*\mathcal{F}^w \wedge K^w = 0$

$*\mathcal{F}^w \wedge K^r = \left(\mathcal{F}^{32}K_{32} + \mathcal{F}^{13}K_{13} + \mathcal{F}^{21}K_{21}\right)\varpi_4$

$\mathcal{F}^w \wedge K^r = \left([\mathcal{F}^r][K^w]^t\right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$*\mathcal{F}^r \wedge K^w = \left(\mathcal{F}^{01}K_{01} + \mathcal{F}^{02}K_{02} + \mathcal{F}^{03}K_{03}\right)\varpi_4$

$*\mathcal{F}^r \wedge K^r = 0$

$G_2(\mathcal{F}^w, K^w) = G_2(\mathcal{F}^r, K^r) = 0$

$G_2(\mathcal{F}^w, K^r) = \left(\mathcal{F}^{32}K_{32} + \mathcal{F}^{13}K_{13} + \mathcal{F}^{21}K_{21}\right)$

$= -\frac{1}{\det P'}[*\mathcal{F}^w][K^r]^t$

$= \left\{[\mathcal{F}^r][h]^{-1}\det h + [\mathcal{F}^w][h]\,j\left([H]^0\right)\right\}[K^r]^t$

$G_2(\mathcal{F}^r, K^w) = \left(\mathcal{F}^{01}K_{01} + \mathcal{F}^{02}K_{02} + \mathcal{F}^{03}K_{03}\right)$

$= -\frac{1}{\det P'}[*\mathcal{F}^r][K^w]^t$

From there, because G_2 is bilinear :

$G_2(\mathcal{F}, K)$

$= G_2(\mathcal{F}^r + \mathcal{F}^w, K^r + K^w)$

$= G_2(\mathcal{F}^r, K^w) + G_2(\mathcal{F}^w, K^r)$

$= \left(\mathcal{F}^{32}K_{32} + \mathcal{F}^{13}K_{13} + \mathcal{F}^{21}K_{21} + \mathcal{F}^{01}K_{01} + \mathcal{F}^{02}K_{02} + \mathcal{F}^{03}K_{03}\right)$

$= \sum_{\{\alpha\beta\}}\mathcal{F}^{\alpha\beta}K_{\alpha\beta}$

$$G_2(\mathcal{F}, K) = -\frac{1}{\det P'}\left([*\mathcal{F}^w][K^r]^t + [*\mathcal{F}^r][K^w]^t\right) = \sum_{\{\alpha\beta\}}\mathcal{F}^{\alpha\beta}K_{\alpha\beta} = \frac{1}{2}\sum_{\alpha\beta=0}^{3}\mathcal{F}^{\alpha\beta}K_{\alpha\beta} \qquad (5.48)$$

5.4.3 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 $\overrightarrow{B}$ are represented by vectors of $\mathbb{R}^3$ with an orthonormal basis :

$\overrightarrow{E} = \sum_{i=1}^3 E^i\varepsilon_i = \sum_{i=1}^3\sum_{\beta=0}^3 E^i P_i^\beta \partial\xi_\beta$

$\overrightarrow{B} = \sum_{i=1}^3 B^i\varepsilon_i = \sum_{i=1}^3\sum_{\beta=0}^3 B^i P_i^\beta \partial\xi_\beta$

To $\overrightarrow{E}, \overrightarrow{B}$ one can associate one forms in the 3 dimensional tangent space to $\Omega_3(t)$:

$E^* = \sum_{\alpha,\beta=1}^3 g_{\alpha\beta}E^\beta d\xi^\alpha$

$B^* = \sum_{\alpha,\beta=1}^{3} g_{\alpha\beta} B^\beta d\xi^\alpha$
Then :
$\mathcal{F}_{EM}^w = d\xi^0 \wedge E^* = \sum_{\alpha,\beta=1}^{3} g_{\alpha\beta} E^\beta d\xi^0 \wedge d\xi^\alpha$
$[\mathcal{F}_{EM}^w] = \left[\sum_{\beta=1}^{3} g_{\alpha\beta} E^\beta\right] = [E]^t [g_3]$
In $T\Omega_3(t)$ one can compute the Hodge dual of B^* which is a $3-1=2$ form :
$*B^* = (\det Q') \sum_{\alpha,\beta=1}^{3} (-1)^{\alpha+1} g^{\alpha\beta} (B^*)_\beta \, d\xi^1 \wedge ..\widehat{d\xi^\alpha}..d\xi^3$
$= (\det Q') \sum_{\gamma=1}^{3} -B^1 d\xi^3 \wedge d\xi^2 - B^2 d\xi^1 \wedge d\xi^3 - B^3 d\xi^2 \wedge d\xi^1$
Then :
$\mathcal{F}^r = - * B^*$
$[\mathcal{F}_{EM}^r] = -[B]^t \det Q'$
And :

$$\begin{gathered}\mathcal{F}_{EM} = d\xi^0 \wedge E^* - *H^* \det Q \\ \mathcal{F}_{EM} = \sum_{\alpha,\beta=1}^{3} g_{\alpha\beta} E^\beta d\xi^0 \wedge d\xi^\alpha + B^1 d\xi^3 \wedge d\xi^2 + B^2 d\xi^1 \wedge d\xi^3 + B^3 d\xi^2 \wedge d\xi^1 \\ [\mathcal{F}_{EM}^r] = -[B]^t \det Q' \\ [\mathcal{F}_{EM}^w] = [E]^t [g_3]\end{gathered} \tag{5.49}$$

The matrix of $\mathcal{F}_{EM}$ is :
$[\mathcal{F}_{EM}] = \begin{bmatrix} 0 & [E^*]^t \\ E^* & [j(B)] \end{bmatrix}$
From which :
$[*\mathcal{F}_{EM}^r] = [E]^t \det Q'$
$[*\mathcal{F}_{EM}^w] = [B]^t [g_3]$

The only geometric quantity is the strength $\mathcal{F}_{EM}$, and the vectors E, B are just components in an orthonormal basis. The split between electric and magnetic components depends on the chart, that is on the observer, as it is well known in Electrodynamics : an EM field can look as a pure electric or magnetic field for another observer.

The Pointing vector is : $S = \frac{1}{\mu_0} E \times B = -\frac{1}{\mu_0} j\left([E][g_3]^{-1}\right)[B]^t \det Q'$ where μ_0 is the vacuum permeability.

The potential $\grave{A}$ is a one form : $\grave{A} = \sum_\alpha \grave{A}_\alpha d\xi^\alpha$:
$\mathcal{F}_{\alpha\beta} = \sum_\alpha \left(\partial_\alpha \grave{A}_\beta - \partial_\beta \grave{A}_\alpha\right) d\xi^\alpha \wedge d\xi^\beta$
$\left[d\grave{A}^r\right] = [\mathcal{F}_{EM}^r] = -[B]^t \det Q'$
$\left[d\grave{A}^w\right] = [\mathcal{F}_{EM}^w] = [E]^t [g_3]$
thus $\left[d\grave{A}^r\right]$ is related to the magnetic field, and $\left[d\grave{A}^w\right]$ to the electric field.

5.4.4 Scalar curvature

In GR another definition of curvature is commonly used, and it is necessary to see how these concepts are related.

Riemann Tensor

The strength $\mathcal{F}_G$ is defined on P_G. The **Riemann curvature** is the tensor, on the associated vector bundle $P_G\left[\mathbb{R}^4, \mathbf{Ad}\right]$:

$R = \sum_{\{\alpha\beta\}ij=0}^{3} \sum_{a=1}^{6} \mathcal{F}_{\alpha\beta}^a [\kappa_a]_j^i \, d\xi^\alpha \wedge d\xi^\beta \otimes \varepsilon_i(m) \otimes \varepsilon^j(m)$

where $[\kappa_a]$ is the matrix of the basis of $so(3,1)$. This is a 2-form on M valued in the linear morphisms $\pounds(T_mM;T_mM)$, expressed in the tetrad. It is convenient to denote the 4×4 matrix $[\mathcal{F}_{\alpha\beta}] = \sum_{a=1}^{6}\mathcal{F}_{\alpha\beta}^{a}[\kappa_a]$.

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 $T_1Spin(3,1)$ in the matrix algebra $so(3,1)$.

We have :

$[R_{\alpha\beta}]_j^i = \partial_\alpha \Gamma^i_{M\beta j} - \partial_\beta \Gamma^i_{M\alpha j} + \sum_{k=0}^{3}\left(\Gamma^i_{M\alpha k}\Gamma^k_{M\beta j} - \Gamma^i_{M\beta k}\Gamma^k_{M\alpha j}\right)$

where $[\Gamma_{M\alpha}] = \sum_{a=1}^{6} G^a_\alpha [\kappa_a]$.

By construct this quantity is covariant (in a change of chart on M) and in a change of gauge on P_G : $\widetilde{R} = R$.

Using

$\varepsilon_i(m) = \sum_{\gamma=0}^{3} P_i^\gamma \partial\xi_\gamma$

$\varepsilon^j(m) = \sum_{\eta=0}^{3} P'^j_\eta d\xi^\eta$

it can be expressed in the chart :

$R = \sum_{\{\alpha\beta\}\gamma\eta}([P][\mathcal{F}_{G\alpha\beta}][P'])^\gamma_\eta d\xi^\alpha \wedge d\xi^\beta \otimes \partial\xi_\gamma \otimes d\xi^\eta$

For any common affine connection the **Riemann tensor** is the tensor :

$\widehat{R} = \sum_{\{\alpha\beta\}}\sum_{\gamma\eta}\widehat{R}^\gamma_{\alpha\beta\eta} d\xi^\alpha \wedge d\xi^\beta \otimes \partial\xi_\gamma \otimes d\xi^\eta$

where $\left[\widehat{R}_{\alpha\beta}\right]_{4\times4} = \left[\partial_\alpha\widehat{\Gamma}_\beta\right] - \left[\partial_\beta\widehat{\Gamma}_\alpha\right] + \left[\widehat{\Gamma}_\alpha\right]\left[\widehat{\Gamma}_\beta\right] - \left[\widehat{\Gamma}_\beta\right]\left[\widehat{\Gamma}_\alpha\right]$ and $\left[\widehat{\Gamma}_\alpha\right]^\beta_\gamma = \left[\widehat{\Gamma}^\gamma_{\alpha\beta}\right]_{4\times4}$ denotes the Christofell form in matrix form.

With any principal connection on P_G one can define an affine connection on TM :

$\widehat{\Gamma}^\gamma_{\alpha\beta} = \left[\widehat{\Gamma}_\alpha\right]^\gamma_\beta = ([P]([\partial_\alpha P'] + [\Gamma_{M\alpha}][P']))^\gamma_\beta$

and one can check that

$$\left[\widehat{R}_{\alpha\beta}\right] = [R_{\alpha\beta}] = [P][\mathcal{F}_{G\alpha\beta}][P'] \Leftrightarrow [\mathcal{F}_{G\alpha\beta}] = [P'][R_{\alpha\beta}][P] \tag{5.50}$$

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 :

$R = \sum_{\{\alpha\beta\}ij}\sum_{a=1}^{6}\mathcal{F}^a_{G\alpha\beta}[\kappa_a]^i_j d\xi^\alpha \wedge d\xi^\beta \otimes \varepsilon_i(m) \otimes \varepsilon^j(m)$

$= \sum_{\{\alpha\beta\}}\sum_{a=1}^{6}\mathcal{F}^a_{G\alpha\beta}([P][\kappa_a][P'])^\gamma_\eta d\xi^\alpha \wedge d\xi^\beta \otimes \partial\xi_\gamma \otimes d\xi^\eta$

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

$R_{\alpha\beta\gamma\eta} = \sum_\lambda R^\lambda_{\alpha\beta\gamma} g_{\lambda\eta} = \sum_\lambda ([P][\mathcal{F}_{G\alpha\beta}][P'])^\lambda_\gamma g_{\lambda\eta} = \left([P']^t[\eta][\mathcal{F}_{G\alpha\beta}][P']\right)^\eta_\gamma$

$= \left(\left([P']^t[\eta][\mathcal{F}_{G\alpha\beta}][P']\right)^t\right)^\gamma_\eta = \left([P']^t[\mathcal{F}_{G\alpha\beta}]^t[\eta][P']\right)^\gamma_\eta$

$= -\left([P']^t[\eta][\mathcal{F}_{G\alpha\beta}][P']\right)^\gamma_\eta = -R_{\alpha\beta\eta\gamma}$

Thus this symmetry is not specific to the Levi-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. 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 :

$\sum_{\beta\gamma} [P]^{\alpha}_{k} [\mathcal{F}_{G\alpha\beta}]^{k}_{l} [P']^{l}_{\eta} [P]^{\beta}_{i} \varepsilon^{i} \otimes [P]^{\eta}_{j} \varepsilon^{j} = \sum_{\alpha\beta j} \left([P] [\mathcal{F}_{G\alpha\beta}]\right)^{\alpha}_{j} d\xi^{\beta} \otimes \varepsilon^{j}$

$\sum_{\alpha\gamma} [P]^{\beta}_{k} [\mathcal{F}_{G\alpha\beta}]^{k}_{l} [P']^{l}_{\eta} [P]^{\alpha}_{i} \varepsilon^{i} \otimes [P]^{\eta}_{j} \varepsilon^{j} = \sum_{\beta\gamma} \left([P] [\mathcal{F}_{G\alpha\beta}]\right)^{\beta}_{j} d\xi^{\alpha} \otimes \varepsilon^{j}$

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 lifting 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 :

$Ric = \sum_{\alpha\beta} Ric_{\alpha\beta} d\xi^{\alpha} \otimes d\xi^{\beta} = \sum_{\alpha\beta} \sum_{\gamma} \left([P] [\mathcal{F}_{G\alpha\gamma}] [P']\right)^{\gamma}_{\beta} d\xi^{\alpha} \otimes d\xi^{\beta}$

$[Ric]^{\alpha}_{\beta} = \sum_{a=1}^{6} \sum_{\alpha\beta} \left([\mathcal{F}^{a}] [P] [\kappa_{a}] [P']\right)^{\alpha}_{\beta}$

$$Ric = \sum_{\alpha\beta=0}^{3} \sum_{a=1}^{6} \left([\mathcal{F}^{a}] [P] [\kappa_{a}] [P']\right)^{\alpha}_{\beta} d\xi^{\alpha} \otimes d\xi^{\beta} \tag{5.51}$$

and the scalar curvature :

$\mathbf{R} = \sum_{\alpha\beta\gamma} g^{\alpha\gamma} [R_{\alpha\beta}]^{\beta}_{\gamma} = \sum_{\alpha\beta\gamma} g^{\alpha\gamma} \left([P] [\mathcal{F}_{G\alpha\beta}] [P']\right)^{\beta}_{\gamma}$ and with $[g]^{-1} = [P] [\eta] [P]^{t}$

$\mathbf{R} = \sum_{\alpha\beta\gamma} \left([P] [\eta] [P]^{t}\right)^{\gamma}_{\alpha} \left([P] [\mathcal{F}_{G\alpha\beta}] [P']\right)^{\beta}_{\gamma} = \sum_{a=1}^{6} \sum_{\alpha\beta} [\mathcal{F}^{a}]^{\alpha}_{\beta} \left([P] [\kappa_{a}] [\eta] [P]^{t}\right)^{\beta}_{\alpha}$

$$\mathbf{R} = \sum_{a=1}^{6} Tr\left([P]^{t} [\mathcal{F}^{a}] [P] [\kappa_{a}] [\eta]\right) \tag{5.52}$$

We have seen previously (change of chart) expressions such as $[P]^{t} [\mathcal{F}^{a}] [P]$. With :

$[P] = \begin{bmatrix} P^{0}_{0} & [P^{0}]_{1\times 3} \\ [P_{0}]_{3\times 1} & [Q]_{3\times 3} \end{bmatrix}$

$[P]^{t} [\mathcal{F}^{a}] [P] = \begin{bmatrix} 0 & \left[\widetilde{\mathcal{F}^{w}}\right]^{a}_{1\times 3} \\ -\left(\left[\widetilde{\mathcal{F}^{w}}\right]^{a}\right)^{t}_{3\times 1} & j\left(\left[\widetilde{\mathcal{F}^{r}}\right]^{a}\right)_{3\times 3} \end{bmatrix}$

$\Leftrightarrow \begin{bmatrix} \left[\widetilde{\mathcal{F}^{r}}\right]^{a} & \left[\widetilde{\mathcal{F}^{w}}\right]^{a} \end{bmatrix} = \begin{bmatrix} [\mathcal{F}^{r}]^{a} & [\mathcal{F}^{w}]^{a} \end{bmatrix} [L_{P}]$

$$[L_P] = \begin{bmatrix} [Q']^t \det Q & -j([O_0])[Q] \\ [Q]j([P^0]) & P_0^0[Q] - [P_0][P^0] \end{bmatrix}$$
$$\left[\widetilde{\mathcal{F}}^r\right]^a = [\mathcal{F}^r]^a [Q']^t \det [Q] + [\mathcal{F}^w]^a [Q] j([P^0])$$
$$\left[\widetilde{\mathcal{F}}^w\right]^a = -[\mathcal{F}^r]^a j([P_0])[Q] + [\mathcal{F}^w]^a \left(P_0^0 [Q] - [P_0][P^0]\right)$$
$a = 1,2,3$:
$$[\kappa_a][\eta] = \begin{bmatrix} 0 & 0 \\ 0 & j(\varepsilon_a) \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & j(\varepsilon_a) \end{bmatrix}$$
$$[P]^t [\mathcal{F}_r^a][P][\kappa_a][\eta] = \begin{bmatrix} 0 & \left[\widetilde{\mathcal{F}_r^w}\right]^a j(\varepsilon_a) \\ 0 & j\left(\left[\widetilde{\mathcal{F}_r^r}\right]^a\right) j(\varepsilon_a) \end{bmatrix}$$
$$Tr\left([P]^t[\mathcal{F}^a][P][\kappa_a][\eta]\right) = Trj\left(\left[\widetilde{\mathcal{F}_r^r}\right]^a\right) j(\varepsilon_a) = -2\left[\widetilde{\mathcal{F}_r^r}\right]^a [\varepsilon_a]$$
$a = 4,5,6$:
$$[\kappa_a][\eta] = \begin{bmatrix} 0 & [\varepsilon_a]^t \\ \varepsilon_a & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} 0 & [\varepsilon_a]^t \\ -\varepsilon_a & 0 \end{bmatrix}$$
$$[P]^t [\mathcal{F}_w^a][P][\kappa_a][\eta] = \begin{bmatrix} -\left[\widetilde{\mathcal{F}_w^w}\right]^a [\varepsilon_a] & 0 \\ -j\left(\left[\widetilde{\mathcal{F}_w^r}\right]^a\right)[\varepsilon_a] & 0 \end{bmatrix}$$
$$Tr\left([P]^t[\mathcal{F}^a][P][\kappa_a][\eta]\right) = Tr\left(-\left[\widetilde{\mathcal{F}_w^w}\right]^a [\varepsilon_a]\right) = -\left[\widetilde{\mathcal{F}_w^w}\right]^a [\varepsilon_a]$$
$$\mathbf{R} = \sum_{a=1}^6 Tr\left([P]^t[\mathcal{F}^a][P][\kappa_a][\eta]\right) = -\sum_{a=1}^6 \left(2\left[\widetilde{\mathcal{F}_r^r}\right]^a [\varepsilon_a] + \left[\widetilde{\mathcal{F}_w^w}\right]^a [\varepsilon_a]\right)$$
$$= \sum_{a=1}^6 \{-2\left([\mathcal{F}_r^r]^a [Q']^t \det[Q] + [\mathcal{F}_r^w]^a [Q] j([P^0])\right)[\varepsilon_a]$$
$$-\left(-[\mathcal{F}_w^r]^a j([P_0])[Q] + [\mathcal{F}_w^w]^a \left(P_0^0[Q] - [P_0][P^0]\right)\right)[\varepsilon_a]\}$$

$$\mathbf{R} = Tr\left\{-2[\mathcal{F}_r^r][Q']^t \det[Q] - 2[\mathcal{F}_r^w][Q]j([P^0]) + [\mathcal{F}_w^r] j([P_0])[Q] - [\mathcal{F}_w^w]\left(P_0^0[Q] - [P_0][P^0]\right)\right\}$$

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.

In the standard chart : $\mathbf{R} = -Tr\left(2[\mathcal{F}_r^r]\left[Q'\right]^t (\det Q) + [\mathcal{F}_w^w][Q]\right)$: only $[\mathcal{F}_r^r], [\mathcal{F}_w^w]$ are involved, which reduces significantly the interest of the scalar curvature to account for $\mathcal{F}_G$.

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.

5.4.5 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 the balance of energy in 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. 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.

We have already a scalar product for scalar forms, we need to extend it to forms valued in the Lie algebras.

Scalar products on the Lie algebras

The strength can be seen as a section of the associated vector bundles $P_G\left[T_1 Spin(3,1), \mathbf{Ad}\right]$,

$P_U\left[T_1 U, Ad\right]$ 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 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.

Scalar product for the gravitational field

The scalar product on $T_1 Spin(3,1)$, induced by the scalar product on the Clifford algebra, is, up to a constant, the Killing form :

$\langle \upsilon(r,w), \upsilon(r',w') \rangle_{Cl(3,1)} = \frac{1}{4}\left(r^t r' - w^t w'\right)$

$a = 1,2,3 : \mathcal{F}^a_{G\alpha\beta} = \mathcal{F}^a_{r\alpha\beta}$

$a = 4,5,6 : \mathcal{F}^a_{G\alpha\beta} = \mathcal{F}^a_{w\alpha\beta}$

For fixed indices $\alpha, \beta, \lambda, \mu$:

$\left\langle \mathcal{F}_{G\alpha\beta}(m), \mathcal{F}'_{G\lambda\mu}(m) \right\rangle_{Cl} = \left\langle \upsilon\left(\mathcal{F}_{r\alpha\beta}, \mathcal{F}_{w\alpha\beta}\right), \upsilon\left(\mathcal{F}'_{r\lambda\mu}, \mathcal{F}'_{w\lambda\mu}\right) \right\rangle_{Cl}$

$= \frac{1}{4}\left(\sum_{a=1}^{3} \mathcal{F}^a_{G\alpha\beta}\mathcal{F}'^a_{G\lambda\mu} - \sum_{a=4}^{6} \mathcal{F}^a_{G\alpha\beta}\mathcal{F}'^a_{G\lambda\mu}\right)$

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 $T_1 U$ is a real vector space the scalar product is a *bilinear symmetric* form.

So we will assume that :

Proposition 94 *There is a definite positive scalar product on the Lie algebra $T_1 U$, defined by a bilinear symmetric form preserved by the adjoint map, that we will denote $\langle\rangle_{T_1 U}$. The basis $\left(\overrightarrow{\theta}_a\right)_{a=1}^m$ of $T_1 U$ 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 $\Lambda_2(M; T_1 U)$ is then defined, pointwise, as

$$\left\langle \mathcal{F}_{A\alpha\beta}(m), \mathcal{F}'_{A\lambda\mu}(m) \right\rangle_{T_1 U} = \sum_{a=1}^{m} \mathcal{F}^a_{A\alpha\beta}(m)\, \mathcal{F}'^a_{A\lambda\mu}(m) \qquad (5.53)$$

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

$\langle \mathcal{F}, K\rangle_G = \left\langle \sum_{a=1}^{6}\sum_{\{\alpha,\beta\}=0}^{3} \mathcal{F}^a_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\kappa}_a, \sum_{b=1}^{6}\sum_{\{\alpha,\beta\}=0}^{3} K^b_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\kappa}_b \right\rangle$

$= \sum_{a,b=1}^{6} \left\langle \sum_{\{\alpha,\beta\}=0}^{3} \mathcal{F}^a_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta, \sum_{\{\alpha,\beta\}=0}^{3} K^b_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta \right\rangle_{TM} \langle \overrightarrow{\kappa}_a, \overrightarrow{\kappa}_b\rangle_{Cl}$

$= \frac{1}{4}\sum_{a,b=1}^{3} \left\langle \sum_{\{\alpha,\beta\}=0}^{3} \mathcal{F}^a_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta, \sum_{\{\alpha,\beta\}=0}^{3} K^a_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta \right\rangle_{TM}$

$-\frac{1}{4}\sum_{a,b=4}^{6} \left\langle \sum_{\{\alpha,\beta\}=0}^{3} \mathcal{F}^a_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta, \sum_{\{\alpha,\beta\}=0}^{3} K^a_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta \right\rangle_{TM}$

Which can be expressed equivalently :

$\langle \mathcal{F}, K\rangle_G = \frac{1}{4}\left(G_2(\mathcal{F}_r, K_r) - G_2(\mathcal{F}_w, K_w)\right)$

$= \frac{1}{4}\sum_{\{\alpha\beta\}} \mathcal{F}_r^{\alpha\beta} K_{r\alpha\beta} - \mathcal{F}_w^{\alpha\beta} K_{w\alpha\beta}$

$= \frac{1}{4}\frac{1}{\det P'}\left([*\mathcal{F}_w^w][K_w^r]^t + [*\mathcal{F}_w^r][K_w^w]^t - \left([*\mathcal{F}_r^w][K_r^r]^t + [*\mathcal{F}^r][K^w]^t\right)\right)$

$= -\frac{1}{8} Tr\left([\mathcal{F}_r][g]^{-1}[K_r][g]^{-1} - [\mathcal{F}_w][g]^{-1}[K_w][g]^{-1}\right)$

With the complex notation : $\mathcal{F}_G = \mathcal{F}_r + i\mathcal{F}_w$

$\langle \mathcal{F}, K\rangle_G = -\frac{1}{4}\frac{1}{\det P'}\operatorname{Re}\left([*\mathcal{F}^w][K^r]^t + [*\mathcal{F}^r][K^w]^t\right)$

$= -\frac{1}{8}\operatorname{Re} Tr\left([\mathcal{F}][g]^{-1}[K][g]^{-1}\right)$

$$\langle \mathcal{F}, K\rangle_G = \frac{1}{4}\sum_{a=1}^{3}\sum_{\{\alpha\beta\}} \mathcal{F}_r^{a\alpha\beta} K^a_{r\alpha\beta} - \mathcal{F}_w^{a\alpha\beta} K^a_{w\alpha\beta} = \frac{1}{8}\sum_{a=1}^{3}\sum_{\alpha\beta=0}^{3} \mathcal{F}_r^{a\alpha\beta} K^a_{r\alpha\beta} - \mathcal{F}_w^{a\alpha\beta} K^a_{w\alpha\beta}$$
$$\langle \mathcal{F}, K\rangle_G = -\frac{1}{4}\frac{1}{\det P'}\operatorname{Re}\left([*\mathcal{F}^w][K^r]^t + [*\mathcal{F}^r][K^w]^t\right) = -\frac{1}{8}\operatorname{Re} Tr\left([\mathcal{F}][g]^{-1}[K][g]^{-1}\right) \tag{5.54}$$

and because the scalar product is symmetric :

$\langle \mathcal{F}, K\rangle_G = -\frac{1}{4\det P'}\operatorname{Re} Tr\left([*K^w][\mathcal{F}^r]^t + [*K^r][\mathcal{F}^w]^t\right)$

For the other fields

$\left\langle \sum_{a=1}^{m}\sum_{\{\alpha,\beta\}=0}^{3} \mathcal{F}^a_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\theta}_a, \sum_{b=1}^{m}\sum_{\{\alpha,\beta\}=0}^{3} K^b_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\theta}_b \right\rangle$

$= \sum_{a,b=1}^{m} \left\langle \sum_{\{\alpha,\beta\}=0}^{3} \mathcal{F}^a_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta, \sum_{\{\alpha,\beta\}=0}^{3} K^b_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta \right\rangle \left\langle \overrightarrow{\theta}_a, \overrightarrow{\theta}_b \right\rangle_{T_1U}$

$= \sum_{a=1}^{m} \left\langle \sum_{\{\alpha,\beta\}=0}^{3} \mathcal{F}^a_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta, \sum_{\{\alpha,\beta\}=0}^{3} K^a_{\alpha\beta} d\xi^\alpha \wedge d\xi^\beta \right\rangle_{TM}$

Which can be expressed equivalently :

$\langle \mathcal{F}, K\rangle_A = \sum_{a=1}^{m} G_2(\mathcal{F}^a, K^a)$

$= \sum_{a=1}^{m}\sum_{\{\alpha\beta\}} \mathcal{F}^{a\alpha\beta} K^a_{\alpha\beta}$

$= -\frac{1}{\det P'}\sum_{a=1}^{m}\left([*\mathcal{F}^{aw}][K^{ar}]^t + [*\mathcal{F}^{ar}][K^{aw}]^t\right)$

$= -\frac{1}{\det P'} Tr\left([*\mathcal{F}^w][K^r]^t + [*\mathcal{F}^r][K^w]^t\right)$

$= -\frac{1}{2}\sum_{a=1}^{m} Tr\left([\mathcal{F}^a][g]^{-1}[K^a][g]^{-1}\right)$

$$\langle \mathcal{F}, K\rangle_A = \sum_{a=1}^{m}\sum_{\{\alpha\beta\}} \mathcal{F}^{a\alpha\beta} K^a_{\alpha\beta} = \frac{1}{2}\sum_{a=1}^{m}\sum_{\alpha\beta=0}^{3} \mathcal{F}^{a\alpha\beta} K^a_{\alpha\beta}$$
$$\langle \mathcal{F}, K\rangle_A = -\frac{1}{\det P'} Tr\left([*\mathcal{F}^w][K^r]^t + [*\mathcal{F}^r][K^w]^t\right) = -\frac{1}{2}\sum_{a=1}^{m} Tr\left([\mathcal{F}^a][g]^{-1}[K^a][g]^{-1}\right) \tag{5.55}$$

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 in a standard chart. Then :

$G_2(\mathcal{F}^a, K^a) = [\mathcal{F}^{a,w}][g_3]^{-1}[K^{a,w}]^t + [\mathcal{F}^{a,r}][g_3][K^{a,r}]^t \det[g_3]^{-1}$

$$\begin{gathered}\langle \mathcal{F}, K\rangle_G = \frac{1}{4}\sum_{a=1}^{3}\operatorname{Re} Tr\left([\mathcal{F}^w][g_3]^{-1}[K^w]^t + [\mathcal{F}^r][g_3][K^r]^t \det[g_3]^{-1}\right)\\ \langle \mathcal{F}, K\rangle_A = Tr\,[\mathcal{F}_A^w][g_3]^{-1}[K^w]^t + [\mathcal{F}_A^r][g_3][K_A^r]^t \det[g_3]^{-1}\end{gathered} \tag{5.56}$$

Because $[g_3]$ is definite positive, and the scalar product on T_1U is also definite positive, then the scalar product is definite positive on the space of $\mathcal{F}_A$:

$$\langle \mathcal{F}_A, \mathcal{F}_A \rangle \geq 0; \langle \mathcal{F}_A, \mathcal{F}_A \rangle_A = 0 \Leftrightarrow \mathcal{F}_A = 0 \tag{5.57}$$

For the EM field :

$$\langle \mathcal{F}_{EM}, \mathcal{F}_{EM} \rangle = [E]^t [g_3] [E] + [B] [g_3] [B]^t \tag{5.58}$$

But this is not the case for the gravitational field.

From the computation above we have :
$\langle \mathcal{F}, K \rangle_G \varpi_4 = \frac{1}{4} \sum_{a=1}^{3} *\mathcal{F}_r^a \wedge K_r^a - *\mathcal{F}_w^a \wedge K_w^a = \frac{1}{4} \sum_{a=1}^{3} *K_r^a \wedge \mathcal{F}_r^a - *K_w^a \wedge \mathcal{F}_w^a$
Similarly :
$\langle \mathcal{F}, K \rangle_A \varpi_4 = \sum_{a=1}^{m} *\mathcal{F}_A^a \wedge K_A^a = \sum_{a=1}^{m} *K_A^a \wedge \mathcal{F}_A^a$

Identity

We have a useful property which is more general, and holds for all the fields:

Theorem 95 *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 :*

$$\forall X, Y, Z \in T_1U : \langle X, [Y, Z] \rangle = \langle [X, Y], Z \rangle \tag{5.59}$$

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
$X = \sum_{a=1}^{6} \sum_{\{\alpha,\beta\}=0}^{3} X_{\alpha\beta}^a d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\kappa}_a,$
$Y = \sum_{a=1}^{6} Y_\alpha^a d\xi^\alpha \otimes \overrightarrow{\kappa}_a,$
$Z = \sum_{a=1}^{6} Z_\alpha^a d\xi^\alpha \otimes \overrightarrow{\kappa}_a$
$[Y, Z] = \sum_{a=1}^{6} \sum_{\{\alpha,\beta\}=0}^{3} [Y_\alpha, Z_\beta]^a d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\kappa}_a$
$\langle X, [Y, Z] \rangle_G$
$= \frac{1}{4} \sum_{a=1}^{3} \sum_{\{\alpha\beta\}} X_r^{a\alpha\beta} [Y_\alpha, Z_\beta]_r^a - X_w^{a\alpha\beta} [Y_\alpha, Z_\beta]_w^a$
$= \sum_{\{\alpha\beta\}} \left\langle X^{\alpha\beta}, [Y_\alpha, Z_\beta] \right\rangle_{Cl}$
$= \sum_{\{\alpha\beta\}} \left\langle \left[X^{\alpha\beta}, Y_\alpha\right], Z_\beta \right\rangle_{Cl}$

$$\langle X, [Y, Z] \rangle_G = \sum_{\{\alpha\beta\}} \left\langle X^{\alpha\beta}, [Y_\alpha, Z_\beta] \right\rangle_{Cl} = \sum_{\{\alpha\beta\}} \left\langle \left[X^{\alpha\beta}, Y_\alpha\right], Z_\beta \right\rangle_{Cl} \tag{5.60}$$

For the other fields :
$X = \sum_{a=1}^{6} \sum_{\{\alpha,\beta\}=0}^{3} X_\alpha^a d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\theta}_a, Y = \sum_{a=1}^{6} Y_\alpha^a d\xi^\alpha \otimes \overrightarrow{\theta}_a, Z = \sum_{a=1}^{6} Z_\alpha^a d\xi^\alpha \otimes \overrightarrow{\theta}_a$
$\langle X, [Y, Z] \rangle_A$
$= \sum_{a=1}^{m} \sum_{\{\alpha\beta\}} X^{a\alpha\beta} [Y_\alpha, Z_\beta]_r^a$
$= \sum_{\{\alpha\beta\}} \left\langle X^{\alpha\beta}, [Y_\alpha, Z_\beta] \right\rangle_{T_1U} = \sum_{\{\alpha\beta\}} \left\langle \left[X^{\alpha\beta}, Y_\alpha\right], Z_\beta \right\rangle_{T_1U}$

$$\langle X, [Y, Z] \rangle_A = \sum_{\{\alpha\beta\}} \left\langle X^{\alpha\beta}, [Y_\alpha, Z_\beta] \right\rangle_{T_1U} = \sum_{\{\alpha\beta\}} \left\langle \left[X^{\alpha\beta}, Y_\alpha\right], Z_\beta \right\rangle_{T_1U} \tag{5.61}$$

Norm for the strength of the field

There is a norm on the spaces of strengths, when expressed in the standard chart :

$\|\mathcal{F}_A(m)\|^2 = \langle \mathcal{F}_A(m), \mathcal{F}_A(m)\rangle_A$

$\|\mathcal{F}_G(m)\|^2 = \|\mathcal{F}_r(m)\|^2 + \|\mathcal{F}_w(m)\|^2$

$\|\mathcal{F}_A\| = \int_\omega \|\mathcal{F}_A(m)\| \varpi_4(m)$

$\|\mathcal{F}_G\| = \int_\omega \|\mathcal{F}_G(m)\| \varpi_4(m)$

If Ω is a relatively compact open of M, the spaces :

$L^2(\Omega, T_1 Spin(3,1), \varpi_4) : \int_\omega \|\mathcal{F}_G(m)\|^2 \varpi_4(m) < \infty$

$L^2(\Omega, T_1 U, \varpi_4) : \int_\omega \|\mathcal{F}_A(m)\|^2 \varpi_4(m) < \infty$

where ω is any compact of Ω, are Fréchet spaces.

Energy

Proposition 96 *The **energy density of the fields**, with respect to the volume form ϖ_4, is, up to a constant,*

for the gravitational field : $\langle \mathcal{F}_G, \mathcal{F}_G\rangle_G$

for the other fields : $\langle \mathcal{F}_A, \mathcal{F}_A\rangle_A$

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.

Conservation of energy

If is assumed that fields interact with particles, and with themselves, but not between each other. However energy is a common variable, so in the vacuum the balance of energy due to the variation of the energy of each field should be even. For a system *comprised only of fields* and a given observer, the conservation of energy means that for the observer :

$\mathcal{E}(t) = \int_{\Omega(t)} \langle \mathcal{F}, \mathcal{F}\rangle \varpi_3 = Ct = \int_{\Omega(t)} i_{\varepsilon_0}(\langle \mathcal{F}, \mathcal{F}\rangle \varpi_4)$

where $\langle \mathcal{F}, \mathcal{F}\rangle$ is the sum of the energy of each field, as defined above.

Consider the manifold $\Omega([t_1, t_2])$ with borders $\Omega(t_1), \Omega(t_2)$:

$\mathcal{E}(t_2) - \mathcal{E}(t_1) = \int_{\partial\Omega([t_1,t_2])} i_{\varepsilon_0}(\langle \mathcal{F}, \mathcal{F}\rangle \varpi_4) = \int_{\Omega([t_1,t_2])} d(i_{\varepsilon_0}\langle \mathcal{F}, \mathcal{F}\rangle \varpi_4)$

with the Lie derivative $\pounds_{\varepsilon_0}(\langle \mathcal{F}, \mathcal{F}\rangle \varpi_4)$:

$d(i_{\varepsilon_0}\langle \mathcal{F}, \mathcal{F}\rangle \varpi_4) = \pounds_{\varepsilon_0}(\langle \mathcal{F}, \mathcal{F}\rangle \varpi_4) - i_{\varepsilon_0} d(\langle \mathcal{F}, \mathcal{F}\rangle \varpi_4)$

$i_{\varepsilon_0} d(\langle \mathcal{F}, \mathcal{F}\rangle \varpi_4) = i_{\varepsilon_0}(d\langle \mathcal{F}, \mathcal{F}\rangle \wedge \varpi_4) + i_{\varepsilon_0}\langle \mathcal{F}, \mathcal{F}\rangle d\varpi_4 = i_{\varepsilon_0}(d\langle \mathcal{F}, \mathcal{F}\rangle \wedge \varpi_4) = 0$

$d(i_{\varepsilon_0}\langle \mathcal{F}, \mathcal{F}\rangle \varpi_4) = \pounds_{\varepsilon_0}(\langle \mathcal{F}, \mathcal{F}\rangle \varpi_4)$

$= (\pounds_{\varepsilon_0}\langle \mathcal{F}, \mathcal{F}\rangle)\varpi_4 + \langle \mathcal{F}, \mathcal{F}\rangle \pounds_{\varepsilon_0}\varpi_4$

$= \frac{1}{c}\frac{\partial}{\partial t}\langle \mathcal{F}, \mathcal{F}\rangle \varpi_4 + \langle \mathcal{F}, \mathcal{F}\rangle (div\varepsilon_0)\varpi_4$

$div\varepsilon_0 = \sum_{\alpha=0}^3 \frac{\partial \varepsilon_\alpha^\alpha}{\partial \xi^\alpha} + \frac{1}{2}\varepsilon_0^\alpha \sum_{\beta\gamma=0}^3 g^{\beta\gamma}\partial_\alpha g_{\beta\gamma} = \frac{1}{2}\sum_{\beta\gamma=0}^3 g^{\beta\gamma}\partial_0 g_{\beta\gamma}$

$= \sum_{\alpha=0}^3 \frac{\partial \varepsilon_\alpha^\alpha}{\partial \xi^\alpha} + \frac{1}{2}\frac{1}{\det g}\varepsilon_0^\alpha \partial_\alpha(\det g) = \frac{1}{2}\frac{1}{\det g}\partial_0(\det g)$

$d(i_{\varepsilon_0}\langle \mathcal{F}, \mathcal{F}\rangle \varpi_4) = \frac{1}{c}\frac{\partial}{\partial t}\langle \mathcal{F}, \mathcal{F}\rangle \varpi_4 + \langle \mathcal{F}, \mathcal{F}\rangle \frac{1}{2}\frac{1}{\det g}\partial_0(\det g)\varpi_4$

$\mathcal{E}(t_2) - \mathcal{E}(t_1) = \int_{\Omega([t_1,t_2])} \left(\frac{\partial}{\partial t}\langle \mathcal{F}, \mathcal{F}\rangle + \langle \mathcal{F}, \mathcal{F}\rangle \frac{1}{2}\frac{1}{\det g}\partial_0(\det g)\right)\varpi_4 =$

$\int_{\Omega(t)} \left(\frac{\partial}{\partial t}\langle \mathcal{F}, \mathcal{F}\rangle + \langle \mathcal{F}, \mathcal{F}\rangle \frac{1}{2}\frac{1}{\det g}\partial_0(\det g)\right)\varpi_3 = Ct$

The conservation of energy implies for the observer :

$$\frac{\partial}{\partial t}\langle \mathcal{F}, \mathcal{F}\rangle + \langle \mathcal{F}, \mathcal{F}\rangle \frac{1}{2}\frac{1}{\det g}\partial_0(\det g) = 0 \tag{5.62}$$

which is similar to the continuity equation for the density of particles.

5.4.6 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 has 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. And its implementation for a 4 dimensional manifold is quite easy.

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 :

$\forall u_1,...,u_{2n} \in \mathfrak{X}(TM) :: \widehat{\varphi}_n(\mathcal{F})(u_1,...,u_{2n})$

$= \frac{1}{(2n)!}\sum_{\sigma\in\mathfrak{S}(2n)} \epsilon(\sigma)\varphi\left(\mathcal{F}\left(u_{\sigma(1)},u_{\sigma(2)}\right),...\mathcal{F}\left(u_{\sigma(2n-1)},u_{\sigma(2n)}\right)\right)$

$\mathcal{F}(u_p,u_q) = \sum_{a=1}^{\dim T_1G}\sum_{\alpha,\beta=1}^{\dim M} \mathcal{F}_{\alpha\beta}^a u_p^\alpha u_q^\beta \overrightarrow{\kappa}_a$

$\varphi_n(\kappa_1,...,\kappa_n) = \sum_{a_1..a_n=1}^{\dim T_1G} \varphi_{i_1...i_n}\kappa_1^{a_1}...\kappa_n^{a_n}$

$\widehat{\varphi}_n(\mathcal{F})$

$= \sum_{\beta_1\beta_2...\beta_{2n}=1}^{\dim M}\left(\frac{1}{(2n)!}\sum_{a_1..a_n=1}^{\dim T_1G}\varphi_{a_1...a_n}\sum_{\sigma\in\mathfrak{S}(2n)}\epsilon(\sigma)\mathcal{F}_{\beta_{\sigma(1)}\beta_{\sigma(2)}}^{a_1}...\mathcal{F}_{\beta_{\sigma(2n-3)}\beta_{\sigma(2n)}}^{a_n}\right) d\xi^{\beta_1}...\wedge d\xi^{\beta_{2n}}$

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 $P_G\left[\mathbb{R}^4, \mathbf{Ad}\right]$

$R = \sum_{\{\alpha\beta\}ij} [\mathcal{F}_{\alpha\beta}]^i_j \, d\xi^\alpha \wedge d\xi^\beta \otimes \varepsilon_i(m) \otimes \varepsilon^j(m)$ and $[\mathcal{F}_{\alpha\beta}] = \sum_{a=1}^6 \mathcal{F}^a_{\alpha\beta}[\kappa_a]$.

The issue is then to compute the maps $\varphi \in I(T_1G, Ad, G)$. 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 $f(X) = \det(I - t[X])$ (Kobayashi p.298), and polarization (Kolar p.218) but we will proceed here with 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 :

$\forall g \in G, X \in T_1G :: \varphi_1(\kappa) = \sum_{a=1}^{\dim T_1G} \varphi_a \kappa^a = \varphi_1(Ad_g \kappa) = \sum_{a,b=1}^{\dim T_1G} \varphi_a [Ad_g]^a_b \kappa^b$

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

$\varphi_2(X, Y) = [X]^t [\varphi_2][Y]$

with a symmetric matrix $[\varphi_2]$. 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 of

$X_{\beta_1\beta_2\beta_3\beta_4} = \sum_{a=1}^3 \mathcal{F}^a_{r\sigma(0)\sigma(1)} \mathcal{F}^a_{r\sigma(2)\sigma(3)} - \mathcal{F}^a_{w\sigma(0)\sigma(1)} \mathcal{F}^a_{w\sigma(2)\sigma(3)}$

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 : $\langle \mathcal{F}_{01}, \mathcal{F}_{32} \rangle = \sum_{a,b=1}^{\dim T_1G} \varphi_{ab} \mathcal{F}^a_{01} \mathcal{F}^b_{32}$

$(\langle \mathcal{F}_{01}, \mathcal{F}_{32} \rangle + \langle \mathcal{F}_{02}, \mathcal{F}_{13} \rangle + \langle \mathcal{F}_{03}, \mathcal{F}_{21} \rangle)\, d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$
$= \sum_{a,b=1}^{\dim T_1 G} \varphi_{ab} \left(\mathcal{F}_{01}^a \mathcal{F}_{32}^b + \mathcal{F}_{02}^a \mathcal{F}_{13}^b + \mathcal{F}_{03}^a \mathcal{F}_{21}^b \right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$
$= \sum_{a,b=1}^{\dim T_1 G} \varphi_{ab} \left([\mathcal{F}^{ar}] \left[\mathcal{F}^{bw}\right]^t + [\mathcal{F}^{aw}] \left[\mathcal{F}^{br}\right]^t \right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$
$= -\sum_{a,b=1}^{\dim T_1 G} \varphi_{ab} \mathcal{F}^a \wedge \mathcal{F}^b$

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 :

$$\langle \mathcal{F}_{01}, \mathcal{F}_{32} \rangle + \langle \mathcal{F}_{02}, \mathcal{F}_{13} \rangle + \langle \mathcal{F}_{03}, \mathcal{F}_{21} \rangle = 0 \tag{5.63}$$

The identity reads for the gravitational field, with $\langle \mathcal{F}_{G01}, \mathcal{F}_{G32} \rangle = [\mathcal{F}_{r01}]^t [\mathcal{F}_{r32}] - [\mathcal{F}_{w01}]^t [\mathcal{F}_{w32}]$
$\sum_{a=1}^3 \mathcal{F}_{r01}^a \mathcal{F}_{r32}^a + \mathcal{F}_{r02}^a \mathcal{F}_{r13}^a + \mathcal{F}_{r03}^a \mathcal{F}_{r21}^a - \mathcal{F}_{w01}^a \mathcal{F}_{w32}^a - \mathcal{F}_{w02}^a \mathcal{F}_{w13}^a - \mathcal{F}_{w03}^a \mathcal{F}_{w21}^a = 0$
or :

$$\left[Tr \left([\mathcal{F}_r^r]^t [\mathcal{F}_r^w] - [\mathcal{F}_w^r]^t [\mathcal{F}_w^w] \right) = 0 \right] \tag{5.64}$$

In complex notation :
$[\mathcal{F}_G^r] = [\mathcal{F}_r^r] + i [\mathcal{F}_w^r], [\mathcal{F}_G^w] = [\mathcal{F}_r^w] + i [\mathcal{F}_w^w]$
$[\mathcal{F}_r^r]^t [\mathcal{F}_r^w] - [\mathcal{F}_w^r]^t [\mathcal{F}_w^w] = \operatorname{Re} [\mathcal{F}_G^r]^t \operatorname{Re} [\mathcal{F}_G^w] - \operatorname{Im} [\mathcal{F}_G^r]^t \operatorname{Im} [\mathcal{F}_G^w] = \operatorname{Re} [\mathcal{F}_G^r]^t [\mathcal{F}_G^w]$

$$Tr \operatorname{Re} [\mathcal{F}_G^r]^t [\mathcal{F}_G^w] = 0 \tag{5.65}$$

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 :

$$Tr \left([\mathcal{F}_A^r]^t [\mathcal{F}_A^w] \right) = 0 \tag{5.66}$$

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 $\mathcal{F}^r, \mathcal{F}^w$.

5.5 THE EINSTEIN'S THEORY OF GRAVITATION

So far we have considered Gravitation as a force field, with specific properties, but nonetheless which can be addressed as it is done in Classic Physics, through its action on particles by forces and potential. The Principle of Equivalence - the equality between inertial mass and gravitational charge - is then simply expressed with the spinor. Einstein's Theory of Gravitation is based on different assumptions.

5.5.1 The assumptions of the theory

In Special Relativity the usage of orthonormal frames is restricted to inertial observers, whose velocity is constant. In GR the definition of inertial observers is the same and does not depend on a chart : their velocity, which is an intrinsic quantity, is constant (the derivative is with respect to the proper time). According to the first law of Mechanics, their momentum is constant and they do not feel any inertial forces. But if their velocity changes, then they feel inertial forces, which can be seen as the result of external forces. Because they are expressed with quantities which are equal - the inertial mass and the gravitational charge - the effects of gravitation can be seen, not as the action of a force field, but as the consequence of the "curved" geometry of the universe. In this picture the geometry, through the metric, is at the root of gravitation.

However in this picture one needs to explain why an observer, who is not submitted to any obvious external force (such as a direct contact or an EM field), has not a constant velocity. So it is assumed that a particle (in our usual meaning) always follows a trajectory which is a geodesic, defined as a curve of minimal length, using the metric.

This definition coincides with the definition by covariant derivative, but only with the Levy-Civita connection. As it can be computed from the metric itself, the theory can then be expressed with covariant derivative, the usual concepts of forces are recovered by the Christofell coefficients. Moreover the choice of the connection is justified on the ground that it is the only metric symmetric connection.

In this theory one needs a mechanism to explain the variation of the gravitational forces with the location and the presence of sources. Einstein postulated his equation $Ric_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta} = \frac{8\pi G}{\sqrt{c}}T_{\alpha\beta}$ with the momentum energy tensor $T_{\alpha\beta} = \frac{\partial T}{\partial g^{\alpha\beta}} - \frac{1}{2}g_{\alpha\beta}T$.

Einstein's assumptions were based on general considerations, but the results can be proven in the general framework that we have introduced and the implementation of the principle of least action, with 2 basic assumptions :

- the Levy-Civita connection, to define geodesics

- the choice of the scalar curvature to represent the self-interaction of the gravitational field in the lagrangian (it replaces our scalar product).

So starting from a very different point of view, the theory can be expressed with more conventional concepts, and implemented in our framework, where it appears as a special case. Its specificity is that all the quantities can be computed from the metric, which is the unique variable.

The assumptions can be enlarged, and many attempts have been done. However it is clear that the introduction of a non symmetric connection entails that of an additional variable (usually expressed as a torsion) besides the metric, that is of a gravitational field in the usual meaning. And similarly if one replaces the scalar curvature by another variable in the lagrangian.

It is useful to review the computations related to inertial observer and the scalar curvature.

5.5.2 The geodesic trajectories

According to the assumption above the world line $q(\tau)$ of a particle is such that $V = \frac{dq}{d\tau}$ follows :

$\widehat{\nabla}_V V = 0 : \sum_{\alpha=0}^{3} \left(\frac{dV^\alpha}{d\tau} + \sum_{\beta,\gamma=0}^{3} \widehat{\Gamma}^\alpha_{\beta\gamma} V^\beta V^\gamma \right) \partial\xi_\alpha = 0$

The inertial force $m_I \frac{dV^\alpha}{d\tau}$ is equal and opposed to the gravitational force $\left(m_G \sum_{\beta,\gamma=0}^{3} \widehat{\Gamma}_{\beta\gamma}^{\alpha} V^\beta V^\gamma\right)$. Because the inertial mass m_I is equal to the gravitational charge m_G all particles follow the same trajectory. The scalar product is preserved on a geodesic, so $\langle V, V\rangle = -c^2$. If an additional force, such as the EM field, is acknowledged, the trajectory is still a geodesic but the gravity is then deduced from the total forces.

In a strict interpretation of the Principle of Relativity, the observer has freedom of gauge. In the tetrad formalism he can choose at any point the vectors $\varepsilon_i = \sum_{\alpha=0}^{3} P_i^\alpha \partial\xi_\alpha$, submitted only to the constraint $[g_{\alpha\beta}] = \sum_{i,j=0}^{3} \eta_{ij} P_\alpha'^i P_j'^\beta$, so they are defined up to a rotation by a section of P_G. But in Einstein's theory, to be consistent we must assume that the observer himself follows a geodesic, the tetrad is transported by the connection and cannot any longer be arbitrary. The vectors P_i become vector fields, and

$i = 0..3 : \widehat{\nabla}_V \varepsilon_i = 0 \Leftrightarrow \sum_{\alpha=0}^{3} P_i^\alpha \partial_\alpha P_i^\alpha + \sum_{\beta,\gamma=0}^{3} \widehat{\Gamma}_{\beta\gamma}^{\alpha} P_i^\beta P_i^\gamma = 0 \Leftrightarrow \partial_\lambda P_i^\alpha + \sum_{\beta,\gamma=0}^{3} \widehat{\Gamma}_{\beta\lambda}^{\alpha} P_i^\beta = 0$

The orthonormal basis is transported by the connection, which is then metric. If one assumes that it is also symmetric it is then necessarily the Levy-Civita connection, computed from the metric :

$\widehat{\Gamma}_{\beta\gamma}^{\alpha} = \frac{1}{2} \sum_\eta g^{\alpha\eta} \left(\partial_\beta g_{\gamma\eta} + \partial_\gamma g_{\beta\eta} - \partial_\eta g_{\beta\gamma}\right)$

The dual tetrad P' is also transported by the connection :

$[\partial_\alpha P'] - \sum_{\beta,\gamma=0}^{3} \left[\widehat{\Gamma}_\alpha\right] [P'] = 0.$

As a consequence the connection associated to the principal connection is null :

$[\Gamma_{M\alpha}] = \left([P'] \left[\widehat{\Gamma}_\alpha\right] - [\partial_\alpha P']\right) [P]$

and, because $[\Gamma_{M\alpha}] = \sum_{a=1}^{6} G_\alpha^a [\kappa_a]$ then $G = 0$. In the theory there is no need for a gravitational field. However it is still necessary to explain the variation of the metric with the sources, and this is done with the scalar curvature as we will see below.

So the Theory is consistent. However if we relax the condition on the connection, we are lead to reintroduce, in one way or another, a potential, that is a gravitational field. In the most general context we come back to our formalism. The connection is still metric, but not necessarily symmetric. Then the assumption about the trajectories give very strong prescriptions for the motion. For any spinor S the condition $\nabla_V S = 0$ reads :

$\nabla_V S = [\gamma C(\sigma)] \left[\gamma C\left(\sigma^{-1} \cdot \frac{d\sigma}{dt} + \mathbf{Ad}_{\sigma^{-1}} \widehat{G}\right)\right] [S_0] = 0$

$\Leftrightarrow \sigma^{-1} \cdot \frac{d\sigma}{dt} + \mathbf{Ad}_{\sigma^{-1}} \widehat{G} = 0$

$\Leftrightarrow \widehat{G} = -\frac{d\sigma}{dt} \cdot \sigma^{-1} = -\upsilon(X_r, X_w)$

With :

$X_\alpha = [C(r)]^t \left(\left(\left[\frac{1}{a_r} + \frac{1}{2} j(r) + \frac{1}{4a_r} j(r) j(r)\right] \partial_\alpha r + \frac{1}{2} j(w) \partial_\alpha w + [A(w)] G_{r\alpha} + [B(w)] G_{w\alpha}\right)$

$Y_\alpha = [C(r)]^t \left(\frac{1}{4a_w} [4 - j(w) j(w)] \partial_\alpha w - [B(w)] G_{r\alpha} + [A(w)] G_{w\alpha}\right)$

the condition reads :

$\sum_{\alpha=0}^{3} V^\alpha \left\{\left[\frac{1}{a_r} + \frac{1}{2} j(r) + \frac{1}{4a_r} j(r) j(r)\right] \partial_\alpha r + \frac{1}{2} j(w) \partial_\alpha w + [A(w)] G_{r\alpha} + [B(w)] G_{w\alpha}\right\} = 0$

$\sum_{\alpha=0}^{3} V^\alpha \left\{\frac{1}{4a_w} [4 - j(w) j(w)] \partial_\alpha w - [B(w)] G_{r\alpha} + [A(w)] G_{w\alpha}\right\} = 0$

$\Leftrightarrow$

$\left[\frac{1}{a_r} + \frac{1}{2} j(r) + \frac{1}{4a_r} j(r) j(r)\right] \frac{dr}{d\tau} + \frac{1}{2} j(w) \frac{dw}{d\tau} + [A(w)] \widehat{G}_r + [B(w)] \widehat{G}_w = 0$

$\frac{1}{4a_w} [4 - j(w) j(w)] \frac{dw}{d\tau} - [B(w)] \widehat{G}_r + [A(w)] \widehat{G}_w = 0$

$\frac{dw}{d\tau} = \left[a_w I + \frac{1}{4a_w} j(w) j(w)\right] \left([B(w)] \widehat{G}_r - [A(w)] \widehat{G}_w\right)$

$= \left[a_w I + \frac{1}{4a_w} j(w) j(w)\right] \left([a_w j(w)] \widehat{G}_r - \left[1 - \frac{1}{2} j(w) j(w)\right] \widehat{G}_w\right)$

$= j(w) \widehat{G}_r + \left(a_w I - \frac{1}{4a_w} j(w) j(w)\right) \widehat{G}_w$

$\frac{dr}{dt}$

$= \left[\frac{1}{a_r} + \frac{1}{2}j(r) + \frac{1}{4a_r}j(r)j(r)\right]^{-1} \left(-\frac{1}{2}j(w)\frac{dw}{d\tau} - [A(w)]\widehat{G}_r - [B(w)]\widehat{G}_w\right)$

$= \left[a_r - \frac{1}{2}j(r)\right]\left(-\frac{1}{2}j(w)\left[j(w)\widehat{G}_r + \left(a_w I - \frac{1}{4a_w}j(w)j(w)\right)\widehat{G}_w\right] - [A(w)]\widehat{G}_r - [B(w)]\widehat{G}_w\right)$

$= \left[a_r - \frac{1}{2}j(r)\right]\left(-\widehat{G}_r + \left(\frac{1}{2a_w} - 2a_w\right)j(w)\widehat{G}_w\right)$

The trajectories are the same for any particle (they do not depend on S_0) and are fully given by the potential along the trajectory $\widehat{G}$. These conditions should apply also to the observers, who cannot stay spatially immobile : their motion must compensate the changes in the gravitational field. And this makes difficult to define physically their charts.

5.5.3 The Einstein equation

The only known model of the variation of a force field with the sources is the implementation of the Principle of Least Action with a lagrangian. And indeed the Einstein's equation can easily be proven with the lagrangian $L = T\left(g, z^i, z^i_\alpha\right) + \frac{\sqrt{c}}{8\pi G}\mathbf{R}$ where T corresponds to a stress tensor, whose specification depends on the problem : it is based on phenomenological laws and can include the EM field. The key variable is the scalar curvature $\mathbf{R}$, it is the only one involved in the vacuum where the equation reads : $Ric_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}\mathbf{R} = 0$

This assumption about the lagrangian is independent from the others, and the equation can be expressed with our usual variables. For any connection :

$Ric = \sum_{\alpha\beta=0}^{3}\sum_{a=1}^{6}\left([\mathcal{F}^a][P][\kappa_a][P']\right)^\alpha_\beta d\xi^\alpha \otimes d\xi^\beta$

So the equation reads :

$\sum_{a=1}^{6}\left([\mathcal{F}^a][P][\kappa_a][P'] - \frac{1}{2}[g]\mathbf{R}\right) = 0$

$[g] = [P']^t[\eta][P']$

$\Rightarrow$

$\sum_{a=1}^{6}[\mathcal{F}^a][P][\kappa_a] - \frac{1}{2}[P']^t[\eta]\mathbf{R} = 0$

$\Rightarrow$

$\sum_{a=1}^{6}[P]^t[\mathcal{F}^a][P][\kappa_a][\eta] = \frac{1}{2}\mathbf{R}I_4$

and (see Scalar Curvature above) :

$$\sum_{a=1}^{6}[P]^t[\mathcal{F}^a][P][\kappa_a][\eta] = \sum_{a=1}^{3}\begin{bmatrix} -\left[\widetilde{\mathcal{F}^w_w}\right]^a[\varepsilon_a] & \left[\widetilde{\mathcal{F}^w_r}\right]^a j(\varepsilon_a) \\ -j\left(\left[\widetilde{\mathcal{F}^r_w}\right]^a\right)[\varepsilon_a] & j\left(\left[\widetilde{\mathcal{F}^r_r}\right]^a\right)j(\varepsilon_a) \end{bmatrix}$$

with

$\left[\widetilde{\mathcal{F}}^r\right]^a = [\mathcal{F}^r]^a[Q']^t\det[Q] + [\mathcal{F}^w]^a[Q]j\left(\left[P^0\right]\right)$

$\left[\widetilde{\mathcal{F}}^w\right]^a = -[\mathcal{F}^r]^a j([P_0])[Q] + [\mathcal{F}^w]^a\left(P^0_0[Q] - [P_0]\left[P^0\right]\right)$

$\mathbf{R} = Tr\left\{-2[\mathcal{F}^r_r][Q']^t\det[Q] - 2[\mathcal{F}^w_r][Q]j\left(\left[P^0\right]\right) + [\mathcal{F}^r_w]j([P_0])[Q] - [\mathcal{F}^w_w]\left(P^0_0[Q] - [P_0]\left[P^0\right]\right)\right\}$

The Einstein equation sums up to :

$[\mathcal{F}^r_r][Q']^t\det[Q] + [\mathcal{F}^w_r][Q]j\left(\left[P^0\right]\right) = \frac{1}{2}Tr\left\{[\mathcal{F}^r_w]j([P_0])[Q] - [\mathcal{F}^w_w]\left(P^0_0[Q] - [P_0]\left[P^0\right]\right)\right\}I_3$

$\sum_{a=1}^{3}\left(-[\mathcal{F}^r_r]^a j([P_0])[Q] + [\mathcal{F}^w_r]^a\left(P^0_0[Q] - [P_0]\left[P^0\right]\right)\right)j(\varepsilon_a) = 0$

$\Rightarrow \mathbf{R} = -\frac{4}{3}Tr\left\{[\mathcal{F}^r_r][Q']^t\det[Q] + [\mathcal{F}^w_r][Q]j\left(\left[P^0\right]\right)\right\}$

that is a set of equations, linear in $\mathcal{F}$, with coefficients of the second order in $[P]$ which can be computed quite easily. As noticed above the tetrad is no longer arbitrary, but a section submitted to strong conditions.

In Einstein's theory the Ricci tensor is computed from the metric, and the tetrad itself depends on the metric, then the equations can be expressed with g and its derivatives.

With a more general connection the results above still hold and, because the observer gets back his freedom of gauge, the tetrad becomes a free variable, which must be accounted for.

5.5.4 Cosmological models

The Einstein's equation in the vacuum can be seen as a propagation equation for the gravitational field. This is also the starting point of cosmological models, representing the universe. It is generally assumed that the physical universe is spatially isotropic at large scale (there is no preferred spatial direction), then its dominant feature is the propagation of the gravitational field, and this leads to models with a singularity (the big bang). The main hypotheses are then about the metric. However these models lead also to static universe (as can be seen in the standard chart). To give more flexibility to the model, a fixed scalar Λ is 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. This issue is not related to the formalism used (the use of the Lévi-Civita connection or the scalar curvature) but to the implementation of the Principle of Least Action. It describes the conditions of an equilibrium, but at the cosmological scale the gravitational field is always expanding, it is never at equilibrium (at least in an infinite universe), so the Principle of Least Action does not hold in this framework.

5.5.5 Conclusion

The Einstein's theory of gravitation is a beautiful, consistent theory. The center role given to the metric has a clear meaning, and enables to proceed to explicit computations. However they are difficult and few solutions are known, variations around the Schwartzchild solution for a spherically symmetric system. Most models use linearized approximations.

The prediction of the Mercury's perihelion in a famous 1915 paper is considered as the first experimental verification of the theory. However other methods provide similar results, and it is not obvious that Einstein himself used strictly his theory (Engelhardt). The experiments involving the propagation of light (red shift, gravitational lenses,...) require additional assumptions, and the models are simplified to the extent that the results could be explained by different ways. The Theory encountered problems in Cosmology but, due to the speculative nature of the topic, they can be easily dismissed. The main issue is its failure to explain the motion of stars in the Galaxy. To keep the equation the usual patch has been proposed : invent a new mystery (dark matter).

Overall, any Theory of Gravitation faces very difficult experimental issues, due to the discrepancy (which, by itself, needs to be explained) between the gravitational and the other fields.

Many adjustments have been proposed to solve the problems. Besides those which involve explicitly Quantum Physics or another geometry (additional dimensions), the idea is to give more flexibility to the Theory by relaxing its constraints on the connection or the lagrangian. Einstein himself, with Cartan and Eisenhart, has considered connections with torsions (the so called "fernparallelism"). However these adjustments, when considered in the traditional framework based on the metric, lead to more complicated computations, in what is already a dreadful endeavour. Moreover they break with the genuine originality of the theory by reintroducing a gravitational field, with all its implications. Anyway the gauge field formalism is then more convenient to study these ideas.

5.6 THE PHENOMENON OF PROPAGATION

The two main features of force fields are their interaction with particles, which involves the connection through the potential, and their propagation, which involves the connection through its "derivative", the strength. The Principle of Least Action gives the conditions of an equilibrium, and it is the way to prove the Maxwell's equations for the EM field. However they do not tell all the story. Solutions of the Maxwell's equations can be found using the retarded potentials of Liénard-Wiechert, but usually only special solutions, plane waves, are considered, and for Optics "light rays", which both assume some preferred modes of propagation. The propagation of the EM field is a topic of great theoretical and practical interest and requires a study by itself. It has been at the origin of the Relativity, and the introduction of the photon seems in opposition to the usual wave behavior.

We do not know much about the gravitational field, but it is usually assumed that its propagation shares the same features as the EM field. We know even less about the other fields, which manifest essentially by discontinuities. However they all can be represented as gauge fields, and we will focus on the propagation of force fields in this framework.

5.6.1 Propagation of a signal

The main features of propagation

The interaction of a field with a particle entails a local change in the value of the field, that we will call a **signal**, which is assumed to occur at a point $O \in M$, that is a given, fixed , location in space and time. This signal then propagates in the vacuum, but it does not stay the same. Its transformation is the result of the interaction of the field with itself : there is no other interaction with a particle, creation or annihilation of bosons, and it is assumed that force fields do not interact with each other. However the metric, which is also present everywhere, can play a role as in the Einstein's Theory.

The implementation of the Principle of Least Action through a lagrangian, based on energy, provides the conditions at equilibrium, as a set of PDE (the Maxwell's equations for the EM field). On a given area, their solutions should be adjusted to the initial conditions. However, in a realist picture, the field is a physical entity : there is only one field at a given point, and there is no reason why the value of the field at a given point should keep the "memory" of its value at a point in the past. But this is a fact that it does. A far away star does not give us any favor in dispensing its EM field, however we can guess, from the value of the EM field at our location, the value of the field originating from the star. All our communications, either by light or radio, use this fact. Optics is just the study of the propagation of light rays between points. They do not involve discontinuous process, so we stay well inside the framework of continuous fields.

These phenomena can be modelled as the propagation in the vacuum of a variation $\delta\mathcal{F}$ of the value of the field from the source located at O to another point A. Its main features are, for the EM field at least :

i) At A is received at most one instance of the signal. The time elapsed between the emission and the reception of the signal depends only on the spatial distance between O and A. This is at the foundation of the concept of "speed of propagation".

ii) All the points located at the same spatial distance from O receive similar signals

iii) The intensity of the signal (which can be measured by different quantities) decrease with the spatial distance.

iv) The signal is measured along directions at the point of reception : propagation is modeled by the strength, which is a 2 form. However the potential, as measured by interaction with known particles, follow similar features.

If we represent the gravitation as a field, it shares similar properties : we do not know how to emit a signal, but we can tell that the field of a planet depends only on the spatial distance. And this is the basis for the concept of gravitational waves.

From these features one can view the field at a given point m as the sum of all the fields $\delta\mathcal{F}$ emanating from sources located in the past of m. And this is at the foundation of the principle of superposition, which is one of the tenets of the theory of fields.

The speed of propagation

The speed of propagation between 2 points O, A is measured with respect to the spatial distance between 2 points. In GR the distance between 2 points depends on a curve C with tangent V between O and A. So it is assumed that the propagation is along some curve, and the spatial distance is between O and the projection $OA\prime$ of C on the hypersurface $\Omega_3(t)$ of an observer, as in the diagram below :

$$\begin{bmatrix} & \Omega_3(t) & & \varepsilon_0 & & \Omega_3(t+\tau) \\ & A' = \varphi_o(ct, \xi_A) & \rightarrow & \rightarrow & \rightarrow & A = \varphi_o(c(t+\tau), \xi_A) \\ & \uparrow & & & \nearrow & \\ \overrightarrow{v} & \uparrow & & \nearrow & V & \\ & \uparrow & \nearrow & & & \\ & O = \varphi_0(ct, \xi_O) & \rightarrow & \rightarrow & \rightarrow & O' = \varphi_0(c(t+\tau), \xi_O) \end{bmatrix}$$

For a given observer C is given by a path $q(\tau) = \varphi_o(c(t+\tau), \xi(\tau))$

$V = c\varepsilon_0 + v$

and the spatial length of OA' is :

$\ell(OA') = \int_0^\tau \sqrt{g_3(q(s))(v(q(s)), v(q(s)))}ds$

$\langle V, V\rangle = g_3(q(s))(v(q(s)), v(q(s))) - c^2$

The speed of propagation does not depend on

- the point on the curve, so that : $g_3(q(s))(v(q(s)), v(q(s))) = w$
- the observer, and because $\langle V, V\rangle$ does not depend on the observer,w does not depend on the observer
- the curve : for any other point B we would have similar curves, with the same spatial speed.

For the EM field $w = c$ and it is generally assumed the gravitational field propagates at the same speed.

So the first assumption is :

Proposition 97 *Force fields propagate on curves of vectors* $V = v + c\varepsilon_0$ *such that* $\langle V, V\rangle = w^2 - c^2$ *where* $w = c$ *for the gravitational and the EM field.*

Equivalently, in the standard chart $\varphi_o(ct, \xi), x = \varphi_\Omega(\xi)$ of the observer there is, for a given point O, a function $f : \Omega_3(t) \to \mathbb{R}_+$ such that the signal is received at the time $t + f(x)$ at the point $\varphi_o(c(t + f(\varphi_\Omega(\xi))), \xi)$. The function f is continuous, $f(x) \geq 0$ and null at O. The set of points A' which receive the signal at $t + \tau$ is a 2 dimensional surface $S_3(O, \tau)$. If f is differentiable the gradiant *gradf* computed with the Riemannian metric on $\Omega_3(t)$ defines a vector field on the tangent space to $\Omega_3(t)$ which is normal to the surfaces $S_3(O, \tau)$, which are themselves diffeomorphic by the flow of v. The set $S_4(O, \tau)$ of points where the signal at O is received after a delay τ is a 4 dimensional cone, with apex O and sections the surfaces $S_3(O, \tau)$. The point O is singular : there are infinitely many curves passing through O.

In SR the propagation curves are straight lines : $V = Ct$ with $\langle V, V\rangle = w^2 - c^2$.

Evolution of the signal

The right variable to study the propagation of a field is its strength $\mathcal{F}$, so we assume that the signal $\delta\mathcal{F}$ is represented by a variable with the same property. In the assumption above the propagation follows the same rules for all the components $\delta\mathcal{F}^a$, but w and the curves are not necessarily the same for different forces fields.

i) A signal is identified by the measure of $\delta\mathcal{F}$ along vectors at the point of reception. To compare the signal at O and A we need to use the same set of coordinates on TM. The mechanism suggests a natural way : pull back the signal $\delta\mathcal{F}(\tau)$ from A to O using the curve with tangent V :

V is a vector field supporting the tangent to the curve

Φ_V is its flow, by construct the parameter on its integral curves is τ so that $q(\tau) = \Phi_V(\tau, O)$.

$\delta\mathcal{F}(\tau)$ is pulled back by $\Phi_V(-\tau, A)$. We take for example a field such that $\mathcal{F} \in L^2(\Omega, T_1U, \varpi_4)$:

$\delta\mathcal{F}(\tau) \in \Lambda_2 T_{q(\tau)}M^* \otimes T_1U \to \widehat{\delta\mathcal{F}}(\tau) \in \Lambda_2 T_O M^* \otimes T_1U$

and $\widehat{\delta\mathcal{F}}(\tau)$ is measured by its action on vectors transported by Φ'_{Vm} :

$\forall u, v \in T_O M : \widehat{\delta\mathcal{F}}(\tau)(u, v) = \delta\mathcal{F}(\tau)(\widetilde{u}, \widetilde{v})$

$\widetilde{u} = \Phi'_{Vm}(\tau, O)(u) \in \Lambda_2 T_{q(\tau)}M$

$\widehat{\delta\mathcal{F}}(\tau)$ depends on $\delta\mathcal{F}(\tau)$ and the curve. Physically it means that the value of $\delta\mathcal{F}(\tau)$ is compared to the value of $\delta\mathcal{F}(O)$, using the same gauge, that is vectors at O.

ii) The relation between $\delta\mathcal{F}(O)$ and the signal received $\delta\mathcal{F}(O)$ can then be expressed by :

$\widehat{\delta\mathcal{F}}(\tau) = \widehat{\mathcal{E}}(\tau, V)(\delta\mathcal{F}(O))$

with a map $\widehat{\mathcal{E}}(\tau, V)$ depending on both the distance between the source and the reception, represented by τ, and the curve, represented by V.

From the experimental facts :

- along the same curve, the signal is similar, but its intensity, which can be represented by the density of energy $\langle \delta\mathcal{F}(q(\tau)), \delta\mathcal{F}(q(\tau)) \rangle$ decreases with τ.

- along different propagation curves but at the same distance the intensity of the signal is similar. That is : $\widehat{\mathcal{E}}(\tau, V) = \widehat{\mathcal{E}}(\tau) \circ \widehat{\mathcal{E}}(V)$

We can express these results as follows :

- let $\widetilde{\delta\mathcal{F}}(\tau)$ be the push-forward of $\delta\mathcal{F}(O)$ along the curve with tangent V :

$\widetilde{\delta\mathcal{F}}(\tau) = \Phi'_{Vm}(\tau, O)_* \delta\mathcal{F}(O)$

for each component $\widetilde{\delta\mathcal{F}}^a$

$\left[\widetilde{\delta\mathcal{F}}^a(\tau)\right] = [K(\tau)]^t [\delta\mathcal{F}^a(O)] [K(\tau)]$ where K depends on τ, V

- along the same curve : $\delta\mathcal{F}(\tau) = \mathcal{E}(\tau) \widetilde{\delta\mathcal{F}}(\tau)$ where $\mathcal{E}(\tau)$ is a linear map, that we express in the simplest possible way by :

$[\delta\mathcal{F}^a(\tau)] = \theta(\tau) \left[\widetilde{\delta\mathcal{F}}^a(\tau)\right]$

with some scalar function $\theta(\tau)$, depending only on τ, and not on the curve.

- the density of energy of the signal is $\langle \delta\mathcal{F}_A(q(\tau)), \delta\mathcal{F}_A(q(\tau)) \rangle = \delta E(q(\tau))$

$\langle \delta\mathcal{F}_A(q(\tau)), \delta\mathcal{F}_A(q(\tau)) \rangle = (\theta(\tau))^2 \left\langle \widetilde{\delta\mathcal{F}}_A(\tau), \widetilde{\delta\mathcal{F}}_A(\tau) \right\rangle$

And $\delta E(q(\tau))$ depends only on $\langle \delta\mathcal{F}_A(O), \delta\mathcal{F}_A(O) \rangle$ and τ, and not on the curve. Meaning that

$\left\langle \widetilde{\delta\mathcal{F}}_A(\tau), \widetilde{\delta\mathcal{F}}_A(\tau) \right\rangle = \langle \delta\mathcal{F}_A(O), \delta\mathcal{F}_A(O) \rangle$

or equivalently that the curves C over which the fields propagate are such that they preserve the scalar product.

$\left\langle \widetilde{\delta\mathcal{F}}_A(\tau), \widetilde{\delta\mathcal{F}}_A(\tau) \right\rangle = \sum_{a=1}^m G_2\left(\widetilde{\delta\mathcal{F}^a}_A(\tau), \widetilde{\delta\mathcal{F}^a}_A(\tau)\right)$

Each component $\widetilde{\delta\mathcal{F}^a}_A$ is transported according to the same rule, so the curve is such that it preserves the scalar product G_2 of 2 forms. The curves with this property are Killing curves.

And we state :

Proposition 98 *Force fields propagate along Killing curves.*

We need to tell more about Killing vector fields and isometries.

5.6.2 Killing vector fields and isometries

Killing vector fields

A Killing curve is a curve whose tangent V is such that the metric g is transported by the flow of V. Killing vector fields are fields of vectors such that their integral curves are Killing curves.

For any vector field V the flow Φ_V is a map : $\Phi_V : \mathbb{R} \times M \to M$. The range of τ is $\mathbb{R}$ if Φ_V is defined on a relatively compact area Ω. It has the properties :

$\Phi_V(-\tau, m) = \Phi_V(\tau, m)^{-1}$

$\Phi_V(\tau + h, m) = \Phi_V(\tau, \Phi_V(h, m))$

By definition the derivative of $\Phi_V(\tau, m)$ with respect to τ gives back the vector field :

$\frac{\partial}{\partial \tau}(\Phi_V(\tau, m))|_{\tau=\tau_0} = V(\Phi_V(\tau_0, m)) = \frac{dq}{d\tau}|_{\tau=\tau_0}$

The derivative with respect to m of the map Φ_V is a linear map :

$\Phi'_{Vm}(\tau, m) :: T_m M \to T_{\Phi_V(\tau,m)} M$

which has for matrix $[J]$, in the holonomic basis in $m = q(0)$ and $q(\tau) = \Phi_V(\tau, m)$:

$\Phi'_{Vm}(\tau, m)(\partial\xi_\alpha(q(0))) = \sum_{\lambda=0}^{3} [J(\tau)]_\alpha^\lambda \partial\xi_\lambda(q(\tau))$

A vector $U(0) \in T_m M \to U(\tau) \in T_{\Phi_V(\tau,m)} M$ is transported by :

$\Phi'_{Vm}(\tau, m) U(0) = \sum_{\alpha,\lambda=0}^{3} [J(\tau)]_\lambda^\alpha U^\lambda(0) \partial\xi_\alpha(q(\tau))$

This is extended to covectors :

$\mu(0) \in T_m M^* \to \mu(\tau) \in T_{\Phi_V(\tau,m)} M$

$\Phi'_{Vm}(\tau, m)(\mu(0)) = \sum_{\beta\beta=0}^{3} [K(\tau)]_\alpha^\beta \mu_\beta(0) d\xi^\alpha(q(\tau))$ with $[K(\tau)] = [J(\tau)]^{-1}$

and more generally to any tensor S, and the transport along V is equivalent to $\pounds_V S = 0$ or equivalently that S is push forward by $\Phi'_{Vm}(\tau, m)$ from m to $\Phi_V(\tau, m)$: $S(\Phi_V(\tau, m)) = \Phi'_{Vm}(\tau, m)_* S(m)$ or that S is pulled back from $\Phi_V(\tau, m)$ to m.

Usually the matrix $[J(\tau)]$ is *not* expressed as an exponential : $[J(\tau)] = \exp \tau [J]$ with a fixed matrix (Maths.1456).

Killing curves are such that the metric g is transported along the curve :

$$\pounds_V g = 0 \Leftrightarrow [g(q(\tau))] = [K(\tau)]^t [g(q(0))] [K(\tau)] \tag{5.67}$$

where $[K(\tau)]$ depends on both the curve and τ.

Isometries

Any differentiable map : $F : M \to M$ reads in any chart : $F(\varphi(\xi)) = \varphi(f(\xi))$ and its derivative $T_m F \in \mathcal{L}(T_m M; T_{F(m)} M)$ has for matrix in the holonomic basis the jacobian : $[J] = [\partial_\beta f^\alpha]$. An isometry is a map such that it preserves the metric : $\forall u, v \in T_m M : \langle F'(m) u, F'(m) v \rangle = \langle u, v \rangle \Longleftrightarrow \left([J]^{-1}\right)^t [g(m)] [J]^{-1} = [g(F(m))]$. The image of an orthonormal basis is an orthonormal basis. The coordinates of u in an orthonormal basis $\varepsilon_i(m)$ are the same as the coordinates of $F'(m) u$ in the orthonormal basis $F'(m) \varepsilon_i(m)$. As any diffeomorphism they transport tensors and in this operation the coordinates of tensors in an orthonormal basis are conserved. As a consequence an isometry preserves the scalar product of forms : $G_r(\lambda, \mu)$ does not depend on the basis, expressed in an orthonormal basis it is a simple expression which is preserved by $F'(m)$. Thus the Hodge dual of $F^*\lambda$ is the image of the Hodge dual : $*(F^*\lambda) = F^*(*\lambda)$. And the density of energy is preserved by an isometry.

Isometries on a manifold constitute a Lie group. The set $\mathfrak{X}(TM)$ of vector fields on a manifold has the structure of Lie algebra with the commutator. Killing vector fields have a structure of Lie subalgebra. If $V_i, i = 1...p$ are Killing vector fields,$V = \sum_{i=1}^{n} a_i V_i$ with fixed scalars is a Killing vector field. If their flow is complete (which is the case if Ω is relatively compact) the Lie algebra of isometries is isomorphic to the Lie algebra of Killing vector fields : Killing vector fields are the infinitesimal generators of isometries. Indeed, for any Killing vector field V and fixed scalar τ one can define the map : $F(m) = \Phi_V(\tau, m)$ which is often denoted : $F(m) = \exp \tau V(m)$ which is an isometry if V is a Killing vector field. Notice that $\exp \tau_1 V_1 \circ \exp \tau_2 V_2 (m) = \exp(\tau_1 + \tau_2) V(m)$ only if $V_1 = V_2$.

As a consequence the transport of a r form along a Killing vector field preserves the scalar product $G_r(\lambda, \mu)$ and the density of energy.

$V = 0$ corresponds to the isometry $F(m) = m$. As a consequence if a Killing vector field is null at a point, it is null everywhere, and if the vector fields $(V_i)_{i=1...N}$ are linearly independent at a point, they are linearly independent everywhere, and conversely, if they are linearly dependent at some point, they are linearly dependent everywhere.

The condition $\pounds_V g = 0$ for a Killing vector field is expressed by (see Annex) :

$$\pounds_V g = 0 \Leftrightarrow \alpha, \beta = 0...3 : \sum_{\gamma=0}^{3} V^\gamma [\partial_\gamma g]^\alpha_\beta + [g]^\beta_\gamma [\partial_\alpha V]^\gamma + [g]^\alpha_\gamma [\partial_\beta V]^\gamma = 0 \tag{5.68}$$

So, for a given metric, Killing vector fields are defined by a set of 10 linear PDE. The initial conditions on $V(a), \partial_\alpha V^\beta(a)$, which can be chosen at any point a, give 20 parameters, of which 10 are related, so the set of Killing vector fields on M is a Lie algebra with dimension at most 10, and similarly for the group of isometries. They represent the possible symmetries of the physical universe, which is characterized by the metric.

The isometries which preserve the vector ε_0 are such that their jacobian : $[J] = \begin{bmatrix} 1 & 0 \\ 0 & [j]_{3\times 3} \end{bmatrix}$ and they are generated by Killing vector fields : $V = c\varepsilon_0 + v(m)$. In the standard gauge, the conditions sum up to :

$\alpha = \beta = 0 : \sum_{\gamma=0}^{3} V^\gamma [\partial_\gamma g_{00}] + [g_{0\gamma}] [\partial_0 V]^\gamma + [g_{0\gamma}] [\partial_0 V]^\gamma = 0$ which is always met

$\alpha = 0, \beta = 1,2,3 : \sum_{\gamma=1}^{3} [g_{\beta\gamma}] \partial_0 v^\gamma = 0 \Leftrightarrow [g_3][\partial_0 v] = 0 \Leftrightarrow [\partial_0 v] = 0$

$\alpha, \beta = 1...3 : c [\partial_0 g_{\alpha\beta}] + \sum_{\gamma=1}^{3} v^\gamma [\partial_\gamma g_{\alpha\beta}] + [g_{\beta\gamma}] [\partial_\alpha v]^\gamma + [g_{\alpha\gamma}] [\partial_\beta v]^\gamma = 0$

If $V_i, i = 1...n$ are such vector fields, then $\sum_{i=1}^{n} A_i V_i$ still belongs to the family iff the scalar constants $\sum_{i=1}^{n} A_i = 1$.

Propagation curves belong to special Killing vector fields : they are related to an observer (they preserve ε_0) and $\langle V, V \rangle = w^2 - c^2$. So we have :

$$\begin{array}{c} [v]^t [g_3] [v] = w^2 \\ [v]^t [\partial_0 g_3] [v] = 0 \\ [\partial_0 v]^t [g_3] [v] = 0 \end{array} \tag{5.69}$$

So propagation curves define a set of at most 6 linearly independent Killing vector fields.

Propagation curves and charts

Charts of the manifold M have for only purpose to locate a point. They are arbitrary. However physical charts are built using physical processes. We have seen in the 3d Chapter how an observer O can build a chart : practically this is done by building a spatial grid by the transmission of an EM signal. To any spatial vector there is an associated propagation curve, so that the spatial chart of $\Omega_3(t)$ is actually a grid made of propagation curves. A spherical chart is an example of such a grid. So we can safely assume that :

Proposition 99 *The spatial axis $\partial\xi_\alpha, \alpha = 1,2,3$ of a standard chart correspond to propagation curves of the EM field.*

Of course there is no propagation curve along the time axis (it would imply $[v]^t [g_3] [v] = 0$). However, for an observer who uses his standard chart the vector ε_0 is transported along the propagation curve.

So we have the 3 Killing vector fields :

$\gamma = 1,2,3 : U_\gamma (m) = u_\gamma (m) \partial\xi_\gamma + c\partial t$

The identities above imply :

$[u_\gamma] [g_{\gamma\gamma}] [u_\gamma] = c^2 \Rightarrow g_{\gamma\gamma} > 0, u_\gamma (m) = c (g_{\gamma\gamma} (m))^{-1/2}$

$[u_\gamma] [\partial_0 g_{\gamma\gamma}] [u_\gamma] = 0 \Rightarrow \partial_0 g_{\gamma\gamma} (m) = 0$

The PDE $\sum_{\lambda=0}^{3} V^\lambda [\partial_\lambda g]^\alpha_\beta + [g]^\beta_\lambda [\partial_\alpha V]^\lambda + [g]^\alpha_\lambda [\partial_\beta V]^\lambda = 0$ gives, with each U_γ differential equations for g which hold at any point :

$\alpha, \beta = 0...3, \gamma = 1,2,3 : 2 [\partial_0 g_{\alpha\beta}] (g_{\gamma\gamma})^{3/2} + 2 [\partial_\gamma g_{\alpha\beta}] (g_{\gamma\gamma}) = [g_{\beta\gamma}] [\partial_\alpha g_{\gamma\gamma}] + [g_{\alpha\gamma}] [\partial_\beta g_{\gamma\gamma}]$

$\alpha, \beta = 0 :$

$2 [\partial_0 g_{00}] (g_{\gamma\gamma})^{3/2} + 2 [\partial_\gamma g_{00}] (g_{\gamma\gamma}) = [g_{0\gamma}] [\partial_0 g_{\gamma\gamma}] + [g_{0\gamma}] [\partial_0 g_{\gamma\gamma}]$

$\alpha = 0, \beta = 1,2,3 :$

$2 [\partial_0 g_{0\beta}] (g_{\gamma\gamma})^{3/2} + 2 [\partial_\gamma g_{0\beta}] (g_{\gamma\gamma}) = [g_{\beta\gamma}] [\partial_0 g_{\gamma\gamma}] + [g_{0\gamma}] [\partial_\beta g_{\gamma\gamma}]$

These equations are always met in the standard gauge.

For $\alpha = \beta = 1,2,3 : [\partial_\gamma g_{\beta\beta}] = \frac{[g_{\beta\gamma}]}{[g_{\gamma\gamma}]} [\partial_\beta g_{\gamma\gamma}]$

By inverting $\beta \leftrightarrow \gamma : \partial_\beta g_{\gamma\gamma} = \frac{g_{\beta\gamma}}{g_{\beta\beta}} \partial_\gamma g_{\beta\beta} = \frac{g_{\beta\gamma}}{g_{\beta\beta}} \frac{g_{\beta\gamma}}{g_{\gamma\gamma}} \partial_\beta g_{\gamma\gamma} = \frac{(g_{\beta\gamma})^2}{g_{\beta\beta} g_{\gamma\gamma}} \partial_\beta g_{\gamma\gamma}$

Because $g_{\beta\beta} > 0$ one can write : $g_{\beta\beta} = \lambda_\beta^2$ and

$g_{32} = a_1 \lambda_2 \lambda_3$

$g_{13} = a_2 \lambda_1 \lambda_3$

$g_{21} = a_3 \lambda_1 \lambda_2$

and the equations read :

(1)

$\partial_2 \lambda_1 = a_3^2 \partial_2 \lambda_1 = a_3 \partial_1 \lambda_2$

$\partial_1 \lambda_2 = a_3^2 \partial_1 \lambda_2 = a_3 \partial_2 \lambda_1$

which have for solutions :

either $(1.a) : \partial_2 \lambda_1 = \partial_1 \lambda_2 = 0$

or $(1.b) : \partial_2 \lambda_1 = \epsilon_3 \partial_1 \lambda_2 \neq 0, a_3 = \epsilon_3, g_{21} = \epsilon_3 \lambda_1 \lambda_2; [\partial_0 g_{21}] = 0$

$\partial_3 \lambda_1 = a_2^2 \partial_3 \lambda_1 = a_2 \partial_1 \lambda_3$

$\partial_1 \lambda_3 = a_2^2 \partial_1 \lambda_3 = a_2 \partial_3 \lambda_1$

which have for solutions :

either $(2.a) : \partial_3 \lambda_1 = \partial_1 \lambda_3 = 0$

or $(2.b) : \partial_3 \lambda_1 = \epsilon_2 \partial_1 \lambda_3 \neq 0; a_2 = \epsilon_2, g_{13} = \epsilon_2 \lambda_1 \lambda_3; \partial_0 g_{13} = 0$

$\partial_3 \lambda_2 = a_1^2 \partial_3 \lambda_2 = a_1 \partial_2 \lambda_3$

$\partial_2 \lambda_3 = a_1^2 \partial_2 \lambda_3 = a_1 \partial_3 \lambda_2$

which have for solutions :

either $(3.a) : \partial_3 \lambda_2 = \partial_2 \lambda_3 = 0$

or $(3.b) : \partial_3 \lambda_2 = \epsilon_1 \partial_2 \lambda_3 \neq 0; a_1 = \epsilon_1, g_{32} = \epsilon_1 \lambda_3 \lambda_2; \partial_0 g_{32} = 0$

For $\alpha \neq \beta = 1,2,3 :$

$2 [\partial_0 g_{\alpha\beta}] (g_{\gamma\gamma})^{3/2} + 2 [\partial_\gamma g_{\alpha\beta}] (g_{\gamma\gamma}) = [g_{\beta\gamma}] [\partial_\alpha g_{\gamma\gamma}] + [g_{\alpha\gamma}] [\partial_\beta g_{\gamma\gamma}]$

$2 [g_{\gamma\gamma}] \left([\partial_0 g_{\alpha\beta}] [g_{\gamma\gamma}]^{1/2} + [\partial_\gamma g_{\alpha\beta}] \right) = [g_{\alpha\gamma}] [g_{\beta\gamma}] \left(\frac{1}{[g_{\alpha\alpha}]} \partial_\gamma g_{\alpha\alpha} + \frac{1}{[g_{\beta\beta}]} \partial_\gamma g_{\beta\beta} \right)$

$\lambda_\gamma^2 ([\partial_0 g_{\alpha\beta}] \lambda_\gamma + [\partial_\gamma g_{\alpha\beta}]) = [g_{\alpha\gamma}] [g_{\beta\gamma}] \left(\lambda_\alpha^{-1} \partial_\gamma \lambda_\alpha + \lambda_\beta^{-1} \partial_\gamma \lambda_\beta \right)$

we have the set (4) of 9 PDE :

$\partial_2 a_1 = -(\partial_0 a_1) \lambda_2$

$\partial_3 a_1 = -(\partial_0 a_1) \lambda_3$

$\partial_1 a_2 = -(\partial_0 a_2) \lambda_1$

$\partial_3 a_2 = -(\partial_0 a_2)\lambda_3$
$\partial_1 a_3 = -(\partial_0 a_3)\lambda_1$
$\partial_2 a_3 = -(\partial_0 a_3)\lambda_2$
$\partial_1 a_1 = -(\partial_0 a_1)\lambda_1 + (a_2 a_3 - a_1)\left(\lambda_2^{-1}\partial_1\lambda_2 + \lambda_3^{-1}\partial_1\lambda_3\right)$
$\partial_2 a_2 = -(\partial_0 a_2)\lambda_2 + (a_1 a_3 - a_2)\left(\lambda_1^{-1}\partial_2\lambda_1 + \lambda_3^{-1}\partial_2\lambda_3\right)$
$\partial_3 a_3 = -(\partial_0 a_3)\lambda_3 + (a_1 a_2 - a_3)\left(\lambda_1^{-1}\partial_3\lambda_1 + \lambda_2^{-1}\partial_3\lambda_2\right)$
Moreover $\det[g_3] = \lambda_1^2\lambda_2^2\lambda_3^2\left(2a_1a_2a_3 - a_1^2 - a_2^2 - a_3^2 + 1\right)$

If all the partial derivatives $\gamma \neq \beta : \partial_\gamma g_{\beta\beta} \neq 0$, that is (1.b),(2.b),(3.b) then $a_\gamma = \epsilon_\gamma$,$\det[g_3] = 2\lambda_1^2\lambda_2^2\lambda_3^2(\epsilon_1\epsilon_2\epsilon_3 - 1)$ and the equations (4) read :
$\lambda_2^{-1}\partial_1\lambda_2 + \lambda_3^{-1}\partial_1\lambda_3 = 0$
$\lambda_1^{-1}\partial_2\lambda_1 + \lambda_3^{-1}\partial_2\lambda_3 = 0$
$\lambda_1^{-1}\partial_3\lambda_1 + \lambda_2^{-1}\partial_3\lambda_2 = 0$
with $\partial_2\lambda_1 = \epsilon_3\partial_1\lambda_2 \neq 0; \partial_3\lambda_1 = \epsilon_2\partial_1\lambda_3 \neq 0; \partial_3\lambda_2 = \epsilon_1\partial_2\lambda_3 \neq 0;$
and one can check that there is no solution.
So at least some of these partial derivatives are null.
If (1.b),(2.b),(3.a) :
$(1.b) : \partial_2\lambda_1 = \epsilon_3\partial_1\lambda_2 \neq 0, a_3 = \epsilon_3, g_{21} = \epsilon_3\lambda_1\lambda_2; [\partial_0 g_{21}] = 0$
$(2.b) : \partial_3\lambda_1 = \epsilon_2\partial_1\lambda_3 \neq 0; a_2 = \epsilon_2, g_{13} = \epsilon_2\lambda_1\lambda_3; [\partial_0 g_{13}] = 0$
$(3.a) : \partial_3\lambda_2 = \partial_2\lambda_3 = 0$
the equations (4) sum up to :
$\partial_2 a_1 = -(\partial_0 a_1)\lambda_2$
$\partial_3 a_1 = -(\partial_0 a_1)\lambda_3$
$\partial_1 a_1 = -(\partial_0 a_1)\lambda_1 + (\epsilon_2\epsilon_3 - a_1)\left(\lambda_2^{-1}\partial_1\lambda_2 + \lambda_3^{-1}\partial_1\lambda_3\right)$
$(a_1 - \epsilon_2\epsilon_3)(\partial_2\lambda_1) = 0$
$(a_1 - \epsilon_3\varepsilon_2)(\partial_3\lambda_1) = 0$
$\Rightarrow a_1 = \epsilon_2\epsilon_3 \Rightarrow \det[g_3] = 0$
So the only solutions are either
$[(1.a),(2.a),(3.a)]$,$[(1.b),(2.a),(3.a)]$, $[(1.a),(2.b),(3.a)]$, $[(1.a),(2.a),(3.b)]$
If no direction is privileged, the only solution is the first. Then the components $g_{\gamma\gamma}$ depend only on the coordinate $\xi_\gamma : g_{11}(\xi_1), g_{22}(\xi_2), g_{33}(\xi_3)$ and the solution is :

$$\begin{gathered} g_{32} = a_1(t,\xi_1,\xi_2,\xi_3)\sqrt{g_{22}(\xi_2)\, g_{33}(\xi_3)} \\ g_{13} = a_2(t,\xi_1,\xi_2,\xi_3)\sqrt{g_{11}(\xi_1)\, g_{33}(\xi_3)} \\ g_{21} = a_3(t,\xi_1,\xi_2,\xi_3)\sqrt{g_{22}(\xi_2)\, g_{11}(\xi_1)} \\ \partial_\gamma a_p = -(g_{\gamma\gamma})^{1/2}\partial_0 a_p, \gamma, p = 1,2,3 \end{gathered} \tag{5.70}$$

which can be written :
$[g_3] = \begin{bmatrix} \lambda_1^2 & a_3\lambda_1\lambda_2 & a_2\lambda_1\lambda_3 \\ a_3\lambda_1\lambda_2 & \lambda_2^2 & a_1\lambda_2\lambda_3 \\ a_2\lambda_1\lambda_3 & a_1\lambda_2\lambda_3 & \lambda_3^2 \end{bmatrix}$,

For any fixed scalars A_1, A_2, A_3 the vector field $V = \sum_{\gamma=1}^3 A_\gamma U_\gamma$ is a Killing vector field. Its integral curves are given by :
$q(\tau) = \varphi_o(\xi_0(\tau), \xi_1(\tau), \xi_2(\tau), \xi_3(\tau))$
$\xi_0(\tau) = \sum_{\beta=1}^3 A_\beta c(t+\tau)$
$V = c\sum_{\gamma=1}^3 A_\gamma\left(\varepsilon_0 + g_{\gamma\gamma}(q(\tau))^{-1/2}\partial\xi_\gamma\right)$
They are propagation curves for the EM field iff
$A = \sum_{\beta=1}^3 A_\beta = 1,$
$\langle V, V\rangle = 2c^2(A_1A_2(a_3 - 1) + A_1A_3(a_2 - 1) + A_2A_3(a_1 - 1)) = 0$

so, practically, if $V = U_\gamma$ for some $\gamma = 1, 2, 3$.

Let us denote $f_\gamma : \mathbb{R}^2 \to \mathbb{R} :: f_\gamma(\tau, \xi_\gamma)$ the solution of the differential equation :

$\frac{\partial f_\gamma}{\partial \tau}(\tau, \xi_\gamma) = A_\gamma c g_{\gamma\gamma}(f_\gamma(\tau, \xi_\gamma))^{-1/2}$

$f_\gamma(0, \xi_\gamma) = \xi_\gamma$

then $\Phi_V(\tau, \varphi_o(ct, \xi_1, \xi_2, \xi_3)) = \varphi_o(c(t+\tau)A, f_1(\tau, \xi_1), f_2(\tau, \xi_2), f_3(\tau, \xi_3))$ is an isometry for τ fixed :

$F(m) = \exp \tau V(m) = \Phi_V(\tau, m)$

Its jacobian is : $[J] = \begin{bmatrix} A & 0 & 0 & 0 \\ 0 & j_1 & 0 & 0 \\ 0 & 0 & j_2 & 0 \\ 0 & 0 & 0 & j_3 \end{bmatrix}$ with $j_\gamma = \frac{\partial}{\partial \xi_\gamma} f_\gamma(\tau, \xi_\gamma)$

We want that the isometry preserves the vector ε_0 which implies $A = 1$.

Then from the relation :

$[J]^t [g(F(m))] [J] = [g(m)]$

we get, in the standard gauge of the observer, the relations :

$a_\gamma(F(m)) = a_\gamma(m)$

The functions a_p are symmetric with respect to F.

The functions f_γ are such that :

$\lambda_\gamma(f_\gamma(\tau, \xi_\gamma)) \frac{\partial}{\partial \tau} f_\gamma(\tau, \xi_\gamma) = A_\gamma c$

$\lambda_\gamma(f_\gamma(\tau, \xi_\gamma)) \frac{\partial}{\partial \xi_\gamma} f_\gamma(\tau, \xi_\gamma) = \lambda_\gamma(\xi_\gamma)$

Isometries and the tetrad

It is useful to review the distinctive properties of the metric, the tetrad, the principal bundle P_G, and isometries.

For any isometry F the tetrad of any observer is transported as an orthonormal basis,

$F'(m)(\varepsilon_i(m)) = \sum_{\alpha,\beta=0}^3 [F'(m)]^\alpha_\beta [P(m)]^\beta_i \partial\xi_\alpha(F(m))$

so there is some matrix $[\mathfrak{T}(m)] \in SO(3,1)$:

$F'(m)(\varepsilon_i(m)) = \sum_{j=0}^3 [\mathfrak{T}(m)]^j_i \varepsilon_j(F(m))$

$[F'(m)] = [P(F(m))] [\mathfrak{T}(m)] [P'(m)]$

However in a change of gauge : $\chi(m)^{-1} \in \mathfrak{X}(P_G)$

$[P(m)] \to \left[\widetilde{P}(m)\right] = [P(m)] [\chi(m)]^{-1}$

$[F'(m)] = [P(F(m))] [\chi(F(m))]^{-1} \left[\widetilde{\mathfrak{T}}(m)\right] [\chi(m)] [P'(m)]$

and $[\mathfrak{T}(\tau, m)]$ transforms as :

$\left[\widetilde{\mathfrak{T}}(m)\right] = [\chi(F(m))] [\mathfrak{T}(m)] [\chi(m)]^{-1}$

so one cannot associate a section $\sigma \in \mathfrak{X}(P_G)$ to an isometry.

Similarly, in the product of isometries :

$F_1 \circ F_2(m) = F_1(F_2(m))$

$(F_1 \circ F_2)'(m) = F_1'(F_2(m)) \circ F_2'(m) = [P(F_1 \circ F_2(m))] [\mathfrak{T}_1(F_2(m))] [\mathfrak{T}_2(m)] [P'(m)]$

$\Leftrightarrow [\mathfrak{T}_1 \mathfrak{T}_2(m)] = [\mathfrak{T}_1(F_2(m))] [\mathfrak{T}_2(m)]$

thus this is not a group isomorphism.

An isometry is defined with respect to a given metric. The metric of the Universe is the main physical property of its Geometry. There are *at most* 10 isometries, they can exist for some areas, and their existence is a physical fact which can be checked. The principal bundle P_G is more general : it gives only, point wise, the rules in a change of orthonormal basis, whatever the metric. The link between the physical metric and P_G is given by the associated vector bundle $P_G\left[\mathbb{R}^4, \mathbf{Ad}\right]$: holonomic bases are defined formally by : $\varepsilon_i(m) = (\mathbf{p}(m), \varepsilon_i)$ and physically by $\varepsilon_i(m) = \sum_{\alpha=0}^3 P_i^\alpha(m) \partial\xi_\alpha$. Then any other basis deduced by rotation with $\sigma \in P_G$ is geometrically equivalent. But it does not tell that $\varepsilon_i(m)$ is orthonormal, one can only say that, if $\varepsilon_i(m)$ is orthonormal, then the basis deduced by rotation with $\sigma \in P_G$ is still orthonormal (because **Ad** preserves the scalar product). The Principle of Relativity is expressed in the choice of the tetrad $P_i^\alpha(m)$, with the constraint that it is orthonormal, which can be physically checked at any point.

One could consider to take the group of isometries as base for the formalism in the Geometry, and this way has been largely explored by physicists and mathematicians. But there are two issues. First we cannot assert the existence of an isometry without the possibility to check it. Second, an isomorphism, by definition, is not local, it relates 2 points and does not give convenient tools to study the motion. The simplest solution is to look for Killing vector fields, which are generators of isometries, and the assumption above gives a natural and simple choice, which has a physical meaning. However, when a special point O is singled out in a problem, the isometries from this point, using the vectors U_γ, provide both a natural system of coordinates and of tetrads (which are then image of the tetrad at O).

The assumption about the propagation of fields along Killing curves, and its consequences on spatial coordinates, answer also to an intriguing question. In the Geometry chapter we have seen that, because of the fundamental symmetry breakdown, the measures of lengths and times cannot be done with the same processes, so a universal constant is needed to relate the measures. However this constant, whatever the units used, happens also to be the "speed of light". We have here an explanation : actually measures of lengths, done along the spatial coordinates, use the propagation of an EM signal.

Killing curves and geodesics

It is usually assumed that fields propagate along geodesics. In the SR (or Newtonian) context the metric is constant and the connection null, fields propagate along straight lines, and there is no difference. However there is one in the GR context. It is useful to review the differences between both assumptions.

The Killing curves are similar to geodesics, they are defined by the properties of their tangent V

geodesic : $\nabla_V V = 0$

Killing curve : $\pounds_V g = 0$

Geodesics are related to connections, Killing curves are related to the metric. The definition of geodesic requires a connection, the definition of Killing curves requires a metric.

Tensors can be transported along any curve, so along geodesics as well as along Killing curves, either by covariant derivative $\nabla_V S = 0$ if there is a connection, or by Lie derivative : $\pounds_V S = 0$ (without any special condition), in both cases.

By definition the transport by Lie derivative along a Killing curve preserves the scalar product. The transport by covariant derivative, along any curve, preserves the scalar product only if the connection is metric.

Historically fiber bundles have been defined through the bundle of frames (local bases) and affine connections. Affine connections define geodesics, which have no other properties. In a picture, such as Einstein's theory, based on the metric, actually Killing curves would be the natural choice of "special curves", geodesics involve a detour by two additional assumptions related to the connection (which must be metric and symmetric), however, as we have seen, there is no gravitational field to speak of, and the propagation of the EM field is an independent issue. The role given to geodesics comes from the extrapolation of the Galilean or SR context, where geodesics are defined as the shortest lines between two points. This properly holds in a Geometry with a varying metric only if the connection is the Levi-Civita connection, and we have seen that this assumption imposes heavy constraints to the theory. It cannot be deduced from a lagrangian based on the scalar curvature. And the issues of gravitational waves and graviton are awkward in this picture. When one goes further than the rigid framework of Einstein's theory, one needs to consider the strength $\mathcal{F}$ as key variable for the propagation. As it comes from the Lie derivative, the quantization of the field leads naturally to Killing vector fields, and the gravitational field has no special role.

With the structure of principal bundle P_G any principal connection provides a linear connection which is necessarily metric, thus the transport by covariant derivative preserves the scalar product along any curve. But there is a connection attached to each force field, and even if the associated geodesics have little meaning here, to give a special role to the geodesics computed from $\mathbf{G}$ is not obvious, notably in the prospect of an unification of force fields. In our picture the metric is a specific feature of the physical universe, Killing curves define physical symmetries which are of geometrical nature. They can be experimentally related to the definition of charts, through the propagation of fields, so their special role seems natural.

However, because geodesics and Killing curves share many mathematical properties, it would be difficult to check experimentally one or the other assumption, and it is more a matter of general consistency of the Theory.

Notice that, whatever the assumption, in any physical chart the direction of the spatial axis are propagation curves, so the spatial speed of propagation implies that the vector fields

$\gamma = 1,2,3 : U_\gamma(m) = w\left(g_{\gamma\gamma}(m)\right)^{-1/2} \partial\xi_\gamma + c\partial t$

are such that their integral curves are propagation curves. If they were geodesics there would be constraints on the Christofell coefficients and the Levy-Civita connection.

Example : single particle

Take a system with a single particle, and only the field originating from the particle itself. The system is spatially symmetric, and we can take a spherical chart centered on the particle : the coordinates of any point are $m = \varphi_o(t,\rho,\theta,\varphi)$ with holonomic basis $(\partial t, \partial\rho, \partial\theta, \partial\phi)$. Then all the variables depend only on t, ρ.

The only propagation curves have for origin O and are along a spatial radial :

$q(\tau) = \varphi_o(t+\tau, \rho(\tau), \theta, \phi)$ with fixed θ, ϕ. The vector $V(q(\tau)) = \frac{dq}{d\tau} = \frac{d\rho}{d\tau}\partial\rho + c\partial t$

The metric in the standard chart is :

$g(m) = \begin{bmatrix} -1 & 0 \\ 0 & [g_3(t,\rho)] \end{bmatrix}$

The spatial length between $O(t+\tau), q(\tau)$ is :

$\ell(O(t+\tau), q(\tau)) = \int_0^\tau \sqrt{g_{\rho\rho}(t+s,\rho(s))\left(\frac{d\rho}{d\tau}(s)\right)^2} ds = \int_0^\tau \frac{d\rho}{d\tau}(s)\sqrt{g_{\rho\rho}(t+s,\rho(s))}ds$

The average speed of spatial propagation is :

$\frac{1}{\tau}\int_0^\tau \frac{d\rho}{d\tau}(s)\sqrt{g_{\rho\rho}(t+s,\rho(s))}ds = c$

By derivation :

$\frac{d\rho}{d\tau}(\tau)\sqrt{g_{\rho\rho}(t+\tau,\rho(\tau))} = c$

and the vector $V(\varphi_o(t+\tau,\rho(\tau),\theta,\phi)) = \frac{c}{\sqrt{g_{\rho\rho}(t+\tau,\rho(\tau))}}\partial\rho + c\partial t$

The vector fields of propagation are then : $V(\varphi_o(t,\rho,\theta,\phi)) = \frac{c}{\sqrt{g_{\rho\rho}(t,\rho)}}\partial\rho + c\partial t$

$V(\Phi_V(\tau, q(0))) = \sum_{\alpha,\beta=0}^3 [J(\tau)]^\alpha_\beta V^\beta(q(0))\partial\xi_\alpha(q(\tau))$

$[J(\tau)] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{\frac{g_{\rho\rho}(t,0)}{g_{\rho\rho}(t+\tau,\rho(\tau))}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Using the value of $V(m) = \frac{w}{\sqrt{g_{\rho\rho}(t,\rho)}}\partial\rho + c\partial t$ in the PDE

$\alpha,\beta = 0...3 : \sum_{\lambda=0}^3 V^\lambda \partial_\lambda [g]^\alpha_\beta + \sum_{\gamma=0}^3 [g]^\beta_\gamma [\partial_\alpha V]^\gamma + [g]^\alpha_\gamma [\partial_\beta V]^\gamma = 0$

provides PDE on g :

$\partial_0 g_{\rho\rho} = 0$

$\partial_\rho g_{\theta\theta} = -\sqrt{g_{\rho\rho}}\partial_0 g_{\theta\theta}$

$\partial_\rho g_{\phi\phi} = -\sqrt{g_{\rho\rho}}\partial_0 g_{\phi\phi}$

$\partial_\rho g_{\theta\phi} = -\sqrt{g_{\rho\rho}}\partial_0 g_{\theta\phi}$
$\partial_0 g_{\rho\theta} + \frac{1}{\sqrt{g_{\rho\rho}}}\partial_\rho g_{\rho\theta} = \frac{1}{2} g_{\theta\rho} (g_{\rho\rho})^{-1/2} \partial_\rho g_{\rho\rho}$
$\partial_0 g_{\rho\phi} + \frac{1}{\sqrt{g_{\rho\rho}}}\partial_\rho g_{\rho\phi} = \frac{1}{2} g_{\rho\phi} (g_{\rho\rho})^{-1/2} \partial_\rho g_{\rho\rho}$

5.6.3 Evolution of the field on the propagation curves

On a propagation curve the signal follows the law :

$$\begin{array}{c} [\delta\mathcal{F}(\tau)] = \theta(\tau)[K(\tau)]^t[\delta\mathcal{F}(O)][\theta(\tau)][K(\tau)] \\ \left[\begin{array}{cc} [\delta\mathcal{F}^r(\tau)] & [\delta\mathcal{F}^w(\tau)] \end{array}\right] = \theta(\tau)\left[\begin{array}{cc} [\delta\mathcal{F}^r(O)] & [\delta\mathcal{F}^w(O)] \end{array}\right]\left[L_{K(\tau)}\right] \end{array} \tag{5.71}$$

where $[\delta\mathcal{F}(\tau)] = [\delta\mathcal{F}_G(\tau)]_{3\times 4}$ (in complex format) or $[\delta\mathcal{F}_A(\tau)]_{m\times 4}$, and $[K(\tau)]$ is defined by the propagation curve. It is such that :

$d\xi^\alpha(q(\tau)) = \sum_{\beta=0}^{3}[K(\tau)]^\alpha_\beta d\xi^\alpha(O)$

In the standard chart the axis $d\xi^0$ is preserved, so :

$[K(\tau)]_{4\times 4} = \begin{bmatrix} 1 & 0 \\ 0 & [k(\tau)]_{3\times 3} \end{bmatrix}$

It is the same for all components and depends on the speed of propagation, so it should be the same for the gravitational and EM fields.

Energy

The function $[\theta(\tau)]$ depends on τ only, the propagation is assumed to the same for all the components $\mathcal{F}^a$ but can possibly be different from one type of field to another (as well as w). Notably the weak and strong interactions have a short range, represented usually by an exponential. As a consequence, for these fields the phenomenon of propagation manifests itself essentially in discontinuous processes, involving elementary particles, and the concept of “propagation in the vacuum” has little physical meaning.

For the EM and gravitational fields, which have an infinite range, the phenomenon of propagation of a signal should follow the Principle of Conservation of Energy. For the observer it is measured at each time over $\Omega_3(t+\tau)$ and the process happens on the spheres $S_3(t+\tau)$. So for $\tau \geq 0$:

$\int_{S_3(t+\tau)} \delta E(\tau)\varpi_2(q(\tau)) = Ct$

and on $S_3(\tau+\tau) : \delta E(\tau) = \theta^2(\tau)\langle\delta\mathcal{F}_A(O), \delta\mathcal{F}_A(O)\rangle = \theta^2(\tau)\delta E(0)$

The sphere $S_3(t+\tau)$ centered at O has a radius $\rho = c\tau$ thus :

$\int_{S_3(t+\tau)} \delta E(\tau)\varpi_2(q(\tau)) = \theta^2(\tau)\delta E(0)\int_{S_3(t+\tau)} \varpi_2(q(\tau)) \simeq 4\pi c^2\tau^2\theta^2(\tau)\delta E(0) = Ct$

and $\theta(\tau) \simeq Ct \times \frac{1}{\tau}, \delta E(\tau) \simeq \frac{1}{\tau^2}\delta E(0) Ct$

Which is consistent with what we know about the EM field. In Quantum Chemistry it is usual to use models with specifications laws in τ^{-n} such as the Lennard-Jonnes potential, which accounts for the local interactions with charged particles.

Derivative of $\delta\mathcal{F}$

$\left[\widetilde{\delta\mathcal{F}}^a(\tau)\right] = [K(\tau)]^t[\delta\mathcal{F}^a(O)][K(\tau)] \Leftrightarrow \pounds_V\widetilde{\delta\mathcal{F}}^a = 0$

The value of the Lie derivative of a 2 form is (see Annex) :

$\pounds_V\delta\mathcal{F}(q(\tau)) = \sum_{a=1}^{m}\left(\sum_{\{\alpha\beta\}}\frac{d}{d\tau}\left(\delta\mathcal{F}^a_{\alpha\beta}\right)d\xi^\alpha\wedge d\xi^\beta + \delta\mathcal{F}^a_{\alpha\beta}\left(\partial_\gamma V^\alpha d\xi^\gamma\wedge d\xi^\beta + \partial_\gamma V^\beta d\xi^\alpha\wedge d\xi^\gamma\right)\right)\otimes \overrightarrow{\theta}_a$

In the standard chart, along a propagation curve : $V = v + c\varepsilon_0, \partial_0 v = 0$

which reads :

$\left[\pounds_V\widetilde{\delta\mathcal{F}}^r(q(\tau))\right] = \frac{d}{d\tau}\left[\widetilde{\delta\mathcal{F}}^r\right] + \left[\widetilde{\delta\mathcal{F}}^w\right]j\left(\left[\partial V^0\right]\right) + \left[\widetilde{\delta\mathcal{F}}^r\right]\left(-[\partial v]^t + (div(V))I_3\right) = 0$

$\left[\pounds_V \widetilde{\delta\mathcal{F}}^w (q(\tau))\right] = \frac{d}{d\tau}\left[\widetilde{\delta\mathcal{F}}^w\right] + \left[\widetilde{\delta\mathcal{F}}^w\right]\left(\partial_0 V^0 + [\partial v]\right) - \left[\widetilde{\delta\mathcal{F}}^r\right] j(\partial_0 V) = 0$

it reads :

$\frac{d}{d\tau}\left[\left[\widetilde{\delta\mathcal{F}}^r (q(\tau))\right] \quad \left[\widetilde{\delta\mathcal{F}}^w (q(\tau))\right] \right] = \left[\left[\widetilde{\delta\mathcal{F}}^r (q(\tau))\right] \quad \left[\widetilde{\delta\mathcal{F}}^w (q(\tau))\right] \right][D(\tau)]$

with

$[D(\tau)] = \begin{bmatrix} \left([\partial v]^t - (Tr[\partial v]) I_3\right) & 0 \\ 0 & -[\partial v] \end{bmatrix}$

$[\partial v] = \begin{bmatrix} \partial_1 v^1 & \partial_2 v^1 & \partial_3 v^1 \\ \partial_1 v^2 & \partial_2 v^2 & \partial_3 v^2 \\ \partial_1 v^3 & \partial_2 v^3 & \partial_3 v^3 \end{bmatrix}$

We have the 2 linear differential equations :

$\frac{d}{d\tau}\left[[\delta\mathcal{F}^r(\tau)] \quad [\delta\mathcal{F}^w(\tau)] \right] = \left[[\delta\mathcal{F}^r(O)] \quad [\delta\mathcal{F}^w(O)] \right]\left[L_{K(\tau)}\right]\left(\frac{1}{\theta}\frac{d\theta}{d\tau} I_3 + [D(\tau)]\right)$

$\frac{d}{d\tau}\left[[\delta\mathcal{F}^r(\tau)] \quad [\delta\mathcal{F}^w(\tau)] \right] = \left[[\delta\mathcal{F}^r(\tau)] \quad [\delta\mathcal{F}^w(\tau)] \right]\frac{1}{\theta}\left(\frac{1}{\theta}\frac{d\theta}{d\tau} I_3 + [D(\tau)]\right)$

Potentials

Meanwhile the potential and the strength do not transform in the same way in a change of gauge in the fiber bundles, they are one and 2 forms on TM and transform as such in a change of chart on M. The strength is defined from the connection form $\widehat{\grave{A}}$ by the relation :

$\mathcal{F}_A(m) = -\mathbf{p}^*(m)\,\pounds\widehat{\grave{A}} = -\mathbf{p}^*(m)\,\chi^* d\grave{A} \in \Lambda_2(M; T_1U)$

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 relation is purely geometric, the definition of horizontal vector fields does not depend on a chart on M, so the relation :

$\mathcal{F}^a_{A\alpha\beta} = \partial_\alpha \grave{A}^a_\beta - \partial_\beta \grave{A}^a_\alpha + 2\left[\grave{A}_\alpha, \grave{A}_\beta\right]^a$

holds in a change of chart on M.

The Lie bracket is a bilinear, antisymmetric, map, and $\left[\grave{A}_\alpha, \grave{A}_\beta\right]^a$ can be written :

$\left[\grave{A}_\alpha, \grave{A}_\beta\right]^a = \sum_{b,c=1}^m C^a_{bc} \grave{A}^b_\alpha \grave{A}^c_\beta$

where C^a_{bc} are fixed scalars, called the structure constants, such that $C^a_{bc} = -C^a_{cb}$.

So, in matrix form : $[\mathcal{F}^a]^\alpha_\beta = \left[d\grave{A}^a\right]^\alpha_\beta + 2\sum_{b,c=1}^m C^a_{bc} \grave{A}^b_\alpha \grave{A}^c_\beta$

The matrix for the last quantity reads :

$\sum_{b,c} C^a_{bc} \begin{bmatrix} 0 & \grave{A}^b_0 \grave{A}^c_1 & \grave{A}^b_0 \grave{A}^c_2 & \grave{A}^b_0 \grave{A}^c_3 \\ -\grave{A}^b_0 \grave{A}^c_1 & 0 & -\grave{A}^b_2 \grave{A}^c_1 & \grave{A}^b_1 \grave{A}^c_3 \\ -\grave{A}^b_0 \grave{A}^c_2 & \grave{A}^b_2 \grave{A}^c_1 & 0 & -\grave{A}^b_3 \grave{A}^c_2 \\ -\grave{A}^b_0 \grave{A}^c_3 & -\grave{A}^b_1 \grave{A}^c_3 & \grave{A}^b_3 \grave{A}^c_2 & 0 \end{bmatrix} = \sum_{b,c} C^a_{bc} \begin{bmatrix} 0 & A^{bc}_w \\ -A^{bc}_w & j\left(A^{bc}_r\right) \end{bmatrix}$

with $\left[A^{bc}_w\right] = \left[\grave{A}^b_0 \grave{A}^c_1 \quad \grave{A}^b_0 \grave{A}^c_2 \quad \grave{A}^b_0 \grave{A}^c_3 \right], \left[A^{bc}_r\right] = \left[\grave{A}^b_3 \grave{A}^c_2 \quad \grave{A}^b_1 \grave{A}^c_3 \quad \grave{A}^b_2 \grave{A}^c_1 \right]$

So we have :

$[\mathcal{F}^a] = \left[d\grave{A}^a\right] + 2\sum_{b,c=1}^m C^a_{bc}\left[F^{bc}\right]$

In a change of chart :

$\grave{A}^a_\alpha \to \widetilde{\grave{A}}^a_\alpha = \sum_{\lambda=0}^3 [M]^\lambda_\alpha \grave{A}^a_\lambda \Leftrightarrow \left[\widetilde{\grave{A}}^a\right] = [A^a][M]$ with the matrix $\left[\widetilde{\grave{A}}^a\right]_{1\times 4}$

$\partial_\alpha \grave{A}^a_\beta - \partial_\beta \grave{A}^a_\alpha$ are the components of the 2 form $d\grave{A}$ thus :

$\left[d\grave{A}^a\right]^\alpha_\beta \to \widetilde{\left[d\grave{A}^a\right]}^\alpha_\beta = \left([M]^t\left[d\grave{A}^a\right][M]\right)^\alpha_\beta$

$\mathcal{F}^a_{A\alpha\beta} \to \widetilde{\mathcal{F}}^a_{A\alpha\beta} = \left([M]^t[\mathcal{F}^a][M]\right)^\alpha_\beta$

with the inverse $[M]$ of the jacobian.

$\left[\widetilde{\mathcal{F}^a}\right] = \left[\widetilde{d\grave{A}}^a\right] + 2\sum_{b,c=1}^m C_{bc}^a \left[\widetilde{F^{bc}}\right]$

$[M]^t [\mathcal{F}^a] [M] = [M]^t \left[d\grave{A}^a\right] [M] + 2\sum_{b,c=1}^m C_{bc}^a \left[\widetilde{F^{bc}}\right]$

$\Rightarrow [M]^t \left[F^{bc}\right] [M] = \left[\widetilde{F^{bc}}\right]$

In the present case we have :

$\left[\widetilde{\delta\mathcal{F}}^a(\tau)\right] = [K(\tau)]^t [\delta\mathcal{F}^a(O)] [K(\tau)]$

so we can write for the potential along a propagation curve :

$\left[d\grave{A}^a(O)\right] \to \left[d\grave{A}^a(\tau)\right] = [K(\tau)]^t \left[d\grave{A}^a(O)\right] [K(\tau)]$

$\sum_{b,c=1}^m C_{bc}^a \grave{A}_\alpha^b(O) \grave{A}_\beta^c(O) \to [K(\tau)]^t \sum_{b,c=1}^m C_{bc}^a \grave{A}_\alpha^b(O) \grave{A}_\beta^c(O) [K(\tau)]$

and :

$$\begin{array}{c} [\delta G(\tau)] = \theta(\tau)[\delta G(O)][K(\tau)] \\ \left[\delta\grave{A}(\tau)\right] = \theta(\tau)\left[\delta\grave{A}(O)\right][K(\tau)] \end{array} \quad (5.72)$$

as $[K(\tau)]_{4\times 4} = \begin{bmatrix} 1 & 0 \\ 0 & [k(\tau)]_{3\times 3} \end{bmatrix}$ we have :

$$\begin{array}{c} \delta G_0^a(\tau) = \theta(\tau)\delta G_0^a(O) \\ \delta\grave{A}_0^a(\tau) = \theta(\tau)\delta\grave{A}_0^a(O) \end{array} \quad (5.73)$$

The law $\left[\delta\grave{A}(\tau)\right] = \theta(\tau)\left[\delta\grave{A}(O)\right][K(\tau)]$ can be seen as an extension of the retarded Liénard-Wiechert potentials.

With $[\theta(\tau)] \simeq Ct\frac{1}{\tau}$ the potential decreases as $\frac{1}{\tau}$.

Propagation in SR geometry

In SR geometry, because the metric is fixed, the Killing curves are straight lines starting at O, the hypersurfaces $\Omega_3(t)$ are hyperplans orthogonal to ε_0.and the surfaces $S_3(O,\tau)$ are 2 dimensional spheres with radius $\rho = w\tau$ which can be expressed in a simple way from the coordinates in a spherical chart centered at O. The coordinates are then :

$m = ct\varepsilon_0 + \rho\overrightarrow{u}$ where $\overrightarrow{u}$ is a unitary vector $u \in \mathbb{R}^3$ normal to $S_3(O,\rho)$

The basis of the chart is constant, the strength is expressed as :

$\mathcal{F}(t,\rho,u) = \sum_{\alpha;\beta=0}^3 [\mathcal{F}(t,\rho,u)]_\beta^\alpha \varepsilon_\alpha \wedge \varepsilon_\beta$

The propagation of the signal is then :

$[\delta\mathcal{F}(t+\tau, w\tau, u)] = \theta(\tau)[\delta\mathcal{F}(t,0,u)]$

We have the model of the propagation of a wave with wave vector $\overrightarrow{u}$.

And similarly for the potential :

$\left[\delta\grave{A}(t+\tau, w\tau, u)\right] = \theta(\tau)\left[\delta\grave{A}(t,0,u)\right]$

which can be seen as a boson with spatial velocity in the direction u, and the Yukawa potential is then $[\theta(\tau)] = \frac{1}{w\tau}\exp(-mw\tau)$.

Discontinuities

The propagation curves can be defined through the gradiant of the function f such that the signal is received at the time $t + f(x)$ at the point $\varphi_o(t + f(x), x)$. But the point O is singular : all the propagation curves originating from O cannot be defined by a single vector field. However all the propagation curves are part of integral curves of Killing vector fields, for which the point O is regular.

It is clear that, with laws such as $[\theta(\tau)] = Ct\frac{1}{\tau}$ there is a discontinuity at O in the value of the field. However such laws come from the measure of the energy, which is never punctual. Actually any measure of the field is done through the interaction of the field with known particles, over some area $\omega \subset \Omega$. The quantities $\mathcal{F}, \grave{A}$ are then operators, acting on vectors, and the results must be understood "in the meaning of distributions". This is a well known problem in the solutions of Maxwell's equations (Maths.2569). However this interpretation depends also on the scale (in space and time) of the area ω used for sampling the data, with respect to the value of the interaction $\delta\mathcal{F}$. To assert that force fields are only "continuous" (as in the classic representation) or "discontinuous" (by bosons) is futile. Our models and representations shall be adjusted to the problem at hand, with the goal of efficiency. In the most usual cases, where the focus is on the propagation of a given signal, the continuous representation (usually by plane waves) suffices, and discontinuous models are required when the focus is on the interaction itself, as we will see in the last chapter.

Observables of propagation

Even if models could enable us to forecast the value of the field at a given point, this endeavor requires the knowledge of the field in a past region, that is an infinite number of data which is impossible to collect. Only in Cosmology the purpose is to model a field (the gravitational field) in its total extension. However the Physicist has usually more limited ambitions :

i) the measure of the field at a given point : it is essentially done through the interaction of the field with a known particle

ii) the computation of the field which results from a given system, that is a limited set of particles

iii) the measure of the field originating from a given source, that is of a signal

For almost all other applications the value of the field at any point it is of little interest. So, rather than trying to account for all the sources, and all the fields propagating from the past, usually the Physicist is focused on a delimited system involving only a given set of particles and the fields with which they interact locally. The goal is to build observables of the field, that is quantities which can be predicted, computed and measured. And, for this, the knowledge of the propagation laws of the field emanating from a given source is essential. Similarly the measure of the field itself is never done precisely at a point, but along a propagation curve (be it of the world line of a spatially immobile observer). So what matters, both practically and theoretically, is the knowledge of the propagation, that is of the curves, the matrices K and θ. With the previous results we have replaced the general model by something which is more manageable, and thus more efficient.

The propagation can then be represented by a map :

$\Theta(\tau)([\delta\mathcal{F}^a(O)]) = \theta(\tau)[K(\tau)]^t[\delta\mathcal{F}^a(O)][K(\tau)]$

on the set H of 4×4 matrices, which is a Hilbert space. The maps $\Theta(\tau)$ can be seen as observables and usually the physicist will choose a specification of $[K]$ such as plane waves, and of $\theta(\tau)$ simples polynomial laws. For this purpose the framework provided by H has many useful properties. For instance the differential equations for the field are linear. A linear differential operator D is a linear map : $J^rH \to H$, the jet extension of H can be assimilated to a product $H \times H.. \times H$ so that D can the expressed by a spectral integral, that is a pseudo differential operator using the properties of the Fourrier transform.

A special important case is the specification of the field by periodic maps.

Periodic fields

A field has a fixed value $\mathcal{F}(m)$ at a given point $m \in M$, so a "periodic field" should be a field such that, along a propagation curve : $\delta\mathcal{F}(q(\tau+T)) = \delta\mathcal{F}(q(\tau))$. However $\delta\mathcal{F}(q(\tau))$ decreases with τ, so it cannot be truly periodic. A periodic field is then necessarily such that the initial value $\delta\mathcal{F}(\varphi_0(ct_0,\xi))$ is periodic : $\delta\mathcal{F}(\varphi_0(c(t_0+T),\xi)) = \delta\mathcal{F}(\varphi_0(ct_0,\xi))$ then the propagated field, at a given point $\delta\mathcal{F}(\varphi_0(ct_0,\eta))$ located a fixed spatial distance from the origin, is periodic. And if the source and the observation points are moving relatively to each other we have the usual Doppler effect.

The set of square $n \times n$ matrices is a Hilbert space, and if the field is considered on a relatively compact area the flow is defined on $\mathbb{R}$, and the maps $[\Theta]$, at a given point m, have a spectral decomposition. The Fourier transform of $[\Theta]$ is :

$$\left[\widehat{\Theta}(\omega)\right] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} [\Theta(\tau)] \exp(-i\omega\tau)\, d\tau$$

and the inverse :

$$[\Theta(\tau)] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left[\widehat{\Theta}(\omega)\right] \exp(i\omega\tau)\, d\omega$$

$\left[\widehat{\Theta}(\omega)\right]$ gives the decomposition of $[\Theta]$ according to ω, so that the field can be seen as the superposition of periodic fields $\left[\widehat{\Theta}(\omega)\right] \exp(i\omega\tau)$. There are some mathematical restrictions to the validity of the usual formulas, but this formalism gives a clear meaning to the common physical definition of "monochromatic" fields and "plane waves", for any field.

5.6.4 Conclusion

That propagation curves are Killing curves is important from a theoretical point of view.

They define the physical symmetries of the universe. Any chart which can be built practically relies on a set of propagation curves, usually with the EM field, and it accounts for the symmetries of a system, as we have done above for a single particle with a spherical chart.

Killing curves are related to the metric. As we will see the metric is locally the result of the balance of energy between all the interactions fields / particles, fields / fields. That this equilibrium, on the most general level, leads to a finite dimensional structure is a striking example of (genuine) quantization : order is borne from chaos.

The phenomenon of propagation is linked to cosmology. The fundamental symmetry breakdown can be seen as a manifestation of an over extending domain of propagation, in a universe which, as a whole, is not fully at equilibrium (and its expansion can be seen as its entropy). The cosmological models of a warped universe, conversely, imply that all the propagation curves belong to some 3 dimensional hypersurface (which is a symmetry) and in this picture its expansion requires a special phenomenon.

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, it 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.

If the system is represented by variables $(z_i)_{i=1}^n$ defined on a fiber bundle $E(M,V,\pi)$, and their r jet extension $j^rZ = \left(z^i_{\alpha_1...\alpha_s}, i = 1...n, s = 0,...,r\right)$. The action is a functional, that is a map : $\ell : \mathfrak{X}(J^rE) \to \mathbb{R}$ usually defined by an integral over a compact area Ω of M with a volume form ϖ_4 :

$\ell(j^rZ) = \int_\Omega \mathcal{L}(j^rZ(m))\, d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$\mathcal{L}(j^rZ(m)) = L(j^rZ(m))\,\varpi_4$

$L(j^rZ(m))$ is the **scalar lagrangian**.

A state of the system represented by the value $j^r\zeta \in \mathfrak{X}(J^rE)$ is deemed to be an equilibrium if the functional is stationary, understood as a local extremum. It can be a maximum or a minimum. It leads to relations between the variables $\zeta^i_{\alpha_1...\alpha_s}$.

Important remark :

A lagrangian is a scalar function whose arguments are coordinates $\left(z^i_{\alpha_1...\alpha_s}, i = 1...n, s = 0,...,r\right)$ in the r-jet extension of some fiber bundle E, which represent a section $j^r\zeta$ of J^rE. So, at a given point m of the basis M of E they have definite values, but the variables $z^i_{\alpha_1...\alpha_s}$ are assumed to be independent : they are not necessarily the partial derivatives of a section z^i of E. This is consistent with the idea that all variations are considered, even if they are not “physically possible”, for instance the components of the vector velocity are not necessarily the derivatives of the coordinates which represent the location [1]. The evolution of the system is not assumed to be continuous in the specification of the lagrangian.

The first step in the search for a solution provides equations between the coordinates in the r-jet, that is between the quantities $z^i_{\alpha_1...\alpha_s}, i = 1...n, s = 0,...,r$. Then *in a second step* one states that, in a continuous process, the quantities $z^i_{\alpha_1...\alpha_s}$ come from a section Z on E : $j^r\zeta = J^rZ$:

[1] This is at the origin of confusion in the consideration of “virtual particles” in QTF. See Chap.8.

$z^i_{\alpha_1...\alpha_s} = \partial_{\alpha_1}...\partial_{\alpha_s} Z^i$. This is exactly the mathematical definition of a partial derivative equation : a relation between the $z^i_{\alpha_1...\alpha_s}$ such that its solutions are the partial derivatives of a common section.

This holds when the variables are maps $z(t)$ defined over an interval $[0, T] \subset \mathbb{R}$ the action takes the form :

$\int_0^T L\left(z^i(t), \delta z^i(t)\right) dt$

where δz^i are independent variables, which, *in a continuous process*, are equal to the derivatives : $\delta z^i(t) = \frac{dz^i}{dt}$.

One crucial step is the specification of the lagrangian, because L sums up much of the Physics of the model. This 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 ?

The r jet section $\left(z^i_{\alpha_1...\alpha_s}, i = 1...n, s = 0, ..., r\right)$ is composed of different variables z, and each of them gives its own r jet.

We have to decide which are the variables that enter the lagrangian and the order of their derivatives. We will limit ourselves to the variables which have been introduced previously, as they give a comprehensive picture of the problems.

For particles the key variable is the state ψ, which sums up all the properties including the motion, or, when only the EM and gravitational field are present, the spinor S. A collection of identical particles whose trajectories do not cross can be represented by a matter field with a density μ, then the measure with respect to which the integral is computed is $\mu\varpi_4$. It can be extended similarly to deformable solids.

The fields are represented by their potential, G for the gravitational field, $\grave{A}$ for the other fields, and their strength $\mathcal{F}_G$,$\mathcal{F}_A$ which accounts for the partial derivatives.

The tetrad P is, in the fiber bundle model, a variable as the others and defines the metric g.

All these variables are maps defined on a bounded area Ω of M, or a bounded interval $[0,T] \subset \mathbb{R}$, 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.

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 a first jet prolongation : $\delta_\alpha\psi, \delta_\alpha G_\beta, \delta_\alpha \grave{A}_\beta, \delta_\alpha P$.

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. Whenever particles are represented as matter fields these relations can be fully expressed with an element of $Spin\left(3,1\right)$.

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 is at the core of the theory of particles.

For elementary particles it is given by the type of the particle : this is the fundamental state ψ_0. Composite particles can be represented by tensor products. When only the EM and gravitational fields are present, the fundamental state is given by an inertial spinor S_0 with the charge.

In both cases each particle are represented by a map $\psi\left(t\right)$ or $S(t)$. Particles with the same fundamental state and whose trajectories do not cross can be represented by a single matter or spinor field with a density.

Deformable solids can be represented by a unique spinor $\gamma C\left(\sigma_B\left(t\right)\right)S_B\left(t\right)$.

Then the definition of the momenta relates the state or the spinor to the motion represented by a map $\sigma\left(t\right)$ or $\sigma\left(m\right)$, which is itself defined by maps $\left(r\left(t\right), w\left(t\right)\right)$ or $\left(r\left(m\right), w\left(m\right)\right)$.

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 in the vacuum.

Partial derivatives and covariant derivatives

In a lagrangian $L\left(z^i,z^i_\alpha\right)$ the variable belongs to a J^1 bundle. To implement the rules of Variational Calculus the partial derivatives $\partial_\alpha\psi,\partial_\alpha\grave{A},\partial_\alpha G,...$ and their 1 jet equivalent are required. However the lagrangian, can be expressed by 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.

In the lagrangian the action of the fields on particles depend on their trajectory through the covariant derivative :

$[\nabla_V\mathcal{M}]=\sum_{\alpha=0}^3 V^\alpha\vartheta\left(\sigma,\varkappa\right)\left(\left[\gamma C\left(\mathbf{Ad}_{\sigma^{-1}}\left(\upsilon\left(X_{r\alpha},X_{w\alpha}\right)+G_\alpha\right)\right)\right]\left[\psi_0\right]+\left[\psi_0\right]\left[Ad_\varkappa\grave{A}_\alpha\right]\right)$

With the representation of particles by spinors the velocity is deduced from σ, so V is not an independent variable in a continuous process. Moreover $V=\frac{dq}{dt}$ represent a trajectory and not a world line.

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 V : 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.32.1.8), including GR, and in our formalism its meaning is more obvious.

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 $\mathbf{e}_p\left(m\right)\otimes\mathbf{f}_q\left(m\right)$ by :

$v=\sum_{\alpha j=0}^3 v^\alpha P'^j_\alpha\varepsilon_j\left(m\right)$ in the orthogonal frame

ε_j acts on $\mathbf{e}_p\left(m\right)\otimes\mathbf{f}_q\left(m\right)$ by γC :

$\left(\mathbf{q}\left(m\right),\mathbf{e}_p\left(m\right)\otimes\mathbf{f}_q\left(m\right)\right)$

$\to\left(\mathbf{q}\left(m\right),\sum_{\alpha j=0}^3 v^\alpha P'^j_\alpha\left(\left[\gamma C\left(\varepsilon_j\right)\right]\mathbf{e}_p\left(m\right)\right)\otimes\mathbf{f}_q\left(m\right)\right)$

iii) thus there is an action of TM^* on $\mathbf{e}_p\left(m\right)\otimes\mathbf{f}_q\left(m\right)$ with $v=\omega^*$

$\left(\mathbf{q}\left(m\right),\gamma C\left(\omega^*\right)\psi\left(m\right)\right)$

$=\left(\mathbf{p}\left(m\right),\sum_{\alpha\beta j=0}^3 g^{\alpha\beta}\omega_\alpha P'^j_\beta\left(\left[\gamma C\left(\varepsilon_j\right)\right]\mathbf{e}_p\left(m\right)\right)\otimes\mathbf{f}_q\left(m\right)\right)$

and as the tetrad defines the metric g :

$\sum_\beta g^{\alpha\beta}P'^j_\beta=\sum_{\beta kl}\eta^{kl}P^\alpha_k P^\beta_l P'^j_\beta=\sum_k\eta^{kj}P^\alpha_k$

$\sum_{\alpha\beta j=0}^3 g^{\alpha\beta}\omega_\alpha P'^j_\beta\left[\gamma C\left(\varepsilon_j\right)\right]\mathbf{e}_p\left(m\right)\otimes\mathbf{f}_q\left(m\right)$

$=\sum_{\alpha\beta=0}^3 g^{\alpha\beta}\omega_\alpha\left[\gamma C\left(\partial\xi_\beta\right)\right]\mathbf{e}_p\left(m\right)\otimes\mathbf{f}_q\left(m\right)$

$=\sum_{\alpha=0}^3\omega_\alpha\left[\gamma C\left(d\xi^\alpha\right)\right]\mathbf{e}_p\left(m\right)\otimes\mathbf{f}_q\left(m\right)$

iv) the covariant derivative is a one form on M so one can take $\varpi=\nabla_\alpha$ and the Dirac operator is :

$$D:\mathfrak{X}\left(J^1Q\left[E\otimes F,\vartheta\right]\right)\to\mathfrak{X}\left(J^1Q\left[E\otimes F,\vartheta\right]\right)::D\psi=\sum_{\alpha=0}^3\left[\gamma C\left(d\xi^\alpha\right)\right]\left[\nabla_\alpha\psi\right]\qquad(6.1)$$

$D\psi=\sum_{\alpha=0}^3\left[P\right]^\alpha_i\left[\gamma C\left(\varepsilon^i\right)\right]\left[\nabla_\alpha\psi\right]$

$\varepsilon^i\left(\varepsilon_j\right)=\delta^i_j\Rightarrow\gamma C\left(\varepsilon^i\right)=\gamma C\left(\varepsilon_i\right)^{-1}$

$D\psi = \sum_{\alpha=0}^{3} [P]_i^\alpha [\gamma C(\varepsilon_i)] [\nabla_\alpha \psi]$

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 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).

Moreover, as we have seen, a lagrangian, which is the central piece in the implementation of the Principle of Least Action, requires that the variables depend on α for a matter field.

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 :

$\ell(Z) = \int_0^T L(t, q^i, y^i)\, dt$

where y^i stands for $\frac{dq^i}{dt}$ in the 1-jet formalism. Actually t is involved explicitly only if there are external (and known) processes.

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 led 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 banal 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 foliation : 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 foliation, 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 [2]. 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, there can be a known external field which adds up to

[2] 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.

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 subtraction of the external field from the value given by the model.

Similarly if the observer is subjected to a specific motion, such as the rotation of his basis with respect to a chart, this motion must be accounted for in the tetrad, but it would be easier to look for a solution for a spatially still observer and then to proceed to a change of observer.

6.1.2 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. 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.

Rules for a general lagrangian

The conditions for the covariance and equivariance of the lagrangian are expressed as relations between the partial derivatives of the lagrangian with respect to the variables, and show that actually some variables cannot figure in the lagrangian. Then any lagrangian, expressed in the remaining variables, will automatically meet the conditions of covariance and equivariance. This is the topic of this subsection. It will necessitate some computations, but they will provide general results, which can be implemented for any lagrangian, and have far reaching consequences.

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 : $\psi^{ij}, G^a_\alpha, \grave{A}^a_\alpha, P^\alpha_i, \delta_\beta\psi^{ij}, \delta_\beta G^a_\alpha, \delta_\beta \grave{A}^a_\alpha, \delta_\beta P^\alpha_i, V^\alpha$ in a section of the 1st jet extension. For the purpose at hand we will use the partial derivatives, $\psi^{ij}, G^a_\alpha, \grave{A}^a_\alpha, P^\alpha_i, \partial_\beta\psi^{ij}, \partial_\beta G^a_\alpha, \partial_\beta \grave{A}^a_\alpha, \partial_\beta P^\alpha_i, V^\alpha$, as the equivalent quantities transform similarly in a change of gauge or charts.

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 $\psi^{ij}, \partial_\alpha\psi^{ij}$ 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 :

$\partial_\alpha\psi^{ij} \to \nabla_\alpha\psi^{ij} = \partial_\alpha\psi^{ij} + \sum_{k=1}^4 \sum_{a=1}^6 \left[\gamma C\left(G^a_\alpha\right)\right]^i_k \psi^{kj} + \sum_{k=1}^n \psi^{ik}\left[\grave{A}_\alpha\right]^k_j$

$\partial_\beta G^a_\alpha \to \mathcal{F}^a_{G\alpha\beta} = \partial_\alpha G^a_\beta - \partial_\beta G^a_\alpha + 2\left[G_\alpha, G_\beta\right]^a$ and $F_{G\alpha\beta} = \partial_\alpha G^a_\beta + \partial_\beta G^a_\alpha$

$\partial_\beta \grave{A}^a_\alpha \to \mathcal{F}^a_{A\alpha\beta} = \partial_\alpha \grave{A}^a_\beta - \partial_\beta \grave{A}^a_\alpha + 2\left[\grave{A}_\alpha, \grave{A}_\beta\right]^a$ and $F_{G\alpha\beta} = \partial_\alpha \grave{A}^a_\beta + \partial_\beta \grave{A}^a_\alpha$

And the lagrangian is then a function :

$\mathcal{L}\left(\psi^{ij}, G^a_\alpha, \grave{A}^a_\alpha, P^\alpha_i, \nabla_\alpha\psi^{ij}, \mathcal{F}_{G\alpha\beta}, F^a_{G\alpha\beta}, \mathcal{F}_{A\alpha\beta}, F_{A\alpha\beta}, \partial_\beta P^\alpha_i, V^\alpha\right)$

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

Equivariance in a change of gauge

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) :

$\kappa \in \mathfrak{X}\left(P_G\left[T_1 Spin(3,1), \mathbf{Ad}\right]\right)$

$\theta \in \mathfrak{X}\left(P_U\left[T_1 U, Ad\right]\right)$

$\kappa = \upsilon\left(\kappa_r, \kappa_w\right), \theta$ 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.1978) :

$\exp : \mathbb{R} \times T_1 Spin\left(3,1\right) \rightarrow Spin\left(3,1\right) :: \exp\left(\tau\kappa\left(m\right)\right)$

$\exp : \mathbb{R} \times T_1 U \rightarrow U :: \exp\left(\tau\theta\left(m\right)\right)$

The exponential on $T_1 Spin\left(3,1\right)$ is expressed by :

$\exp t\kappa = \exp \tau\upsilon\left(\kappa_r, \kappa_w\right) = \sigma_w\left(\tau, \kappa_w\right) \cdot \sigma_r\left(\tau, \kappa_r\right)$

$\sigma_w\left(\tau, \kappa_w\right) = a_w\left(\tau, \kappa_w\right) + \sinh \frac{1}{2}\tau\sqrt{\kappa_w^t \kappa_w}\upsilon\left(0, \kappa_w\right)$

$a_w\left(\tau, \kappa_w\right) = \sqrt{1 + \frac{1}{4}\left(\kappa_w^t \kappa_w \sinh^2 \frac{1}{2}\tau\sqrt{\kappa_w^t \kappa_w}\right)}$

$\sigma_r\left(\tau, \kappa_r\right) = a_r\left(\tau, \kappa_w\right) + \sin t\frac{1}{2}\sqrt{\kappa_r^t \kappa_r}\upsilon\left(\kappa_r, 0\right)$

$a_r\left(\tau, \kappa_w\right) = \sqrt{1 - \frac{1}{4}\kappa_r^r \kappa_r \sin^2 t\frac{1}{2}\sqrt{\kappa_r^t \kappa_r}}$

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 :

$\frac{d}{d\tau}\exp\left(\tau\kappa\left(m\right)\right)|_{\tau=0} = \kappa\left(m\right)$

$\frac{d}{d\tau}\exp\left(\tau\theta\left(m\right)\right)|_{\tau=0} = \theta\left(m\right)$

With the change of gauge :

$\mathbf{p}_G\left(m\right) \rightarrow \widetilde{\mathbf{p}}_G\left(m, \tau\right) \cdot \exp\left(-\tau\kappa\left(m\right)\right)$

$\mathbf{p}_U\left(m\right) \rightarrow \widetilde{\mathbf{p}}_U\left(m\right) \cdot \exp\left(-\tau\theta\left(m\right)\right)$

The components of the variables become :

$P_i^\alpha \rightarrow \widetilde{P}_i^\alpha\left(m, \tau\right) = \sum_{j=0}^3 \left[h\left(\exp\left(-\tau\kappa\right)\right)\right]_i^j P_j^\alpha$ where $[h]$ is the $SO(3,1)$ corresponding matrix

$\psi^{ij} \rightarrow \widetilde{\psi}^{ij}\left(m, \tau\right) = \sum_{k=1}^4 \sum_{l=1}^n \left[\gamma C\left(\exp\left(\tau\kappa\right)\right)\right]_k^i \left[\varrho\left(\exp\left(\tau\theta\right)\right)\right]_l^j \psi^{kl}$

$G_\alpha\left(m\right) \rightarrow \widetilde{G}_\alpha\left(m\right) = \mathbf{Ad}_{\exp\tau\kappa}\left(G_\alpha - \exp\left(-\tau\kappa\right)\left(\exp\tau\kappa\right)'\tau\partial_\alpha\kappa\right)$

$\grave{A}_\alpha \rightarrow \widetilde{\grave{A}}_\alpha\left(m, \tau\right) = Ad_{\exp\tau\theta}\left(\grave{A}_\alpha - \exp\left(-\tau\theta\right)\exp\left(\tau\theta\right)'\tau\partial_\alpha\theta\right)$

$\nabla_\alpha\psi \rightarrow \widetilde{\nabla_\alpha\psi^{ij}}\left(m, \tau\right) = \sum_{k=1}^4 \sum_{l=1}^n \left[\gamma C\left(\exp\left(\tau\kappa\right)\right)\right]_k^i \left[\varrho\left(\exp\left(\tau\theta\right)\right)\right]_l^j \nabla_\alpha\psi^{kl}$

All these expressions depend on m, as well as $\kappa\left(m\right), \theta\left(m\right)$, so they can be differentiated with respect to the coordinates of m to get :

$\partial_\beta P_i^\alpha \rightarrow \widetilde{\partial_\beta P}_i^\alpha\left(m, \tau\right) = \sum_{j=0}^3 \left(\left[h\left(\exp\left(-\tau\kappa\right)'\partial_\beta\kappa\right)\right]_i^j P_j^\alpha + \left[h\left(\exp\left(-\tau\kappa\right)\right)\right]_i^j \partial_\beta P_i^\alpha\right)$

$\partial_\beta\widetilde{G}_\alpha\left(m, \tau\right)$

$= \left[\left(\exp -\tau\kappa\right)\left(\exp\tau\kappa\right)'\tau\partial_\beta\kappa, G_\alpha - \tau\partial_\alpha\kappa\right]$

$+Ad_{\exp\tau\kappa}\{\partial_\beta G_\alpha -$

$\{\left(\exp -\tau\kappa\right)'\tau\partial_\beta\kappa \circ \left(\exp\tau\kappa\right)'\tau\partial_\alpha\kappa$

$+\exp\left(-\tau\kappa\right) \circ \left(\exp\tau\kappa\right)"\left(\tau\partial_\beta\kappa, \tau\partial_\alpha\kappa\right) + \exp\left(-\tau\kappa\right) \circ \exp\left(\tau\kappa\right)'\tau\partial_{\alpha\beta}^2\kappa\}\}$

$\partial_\beta\widetilde{\grave{A}}_\alpha\left(m, \tau\right) = \left[\left(\exp -\tau\theta\right)\left(\exp\tau\theta\right)'\tau\partial_\beta\theta, \grave{A}_\alpha - \tau\partial_\alpha\theta\right]$

$+Ad_{\exp\tau\theta}(\partial_\beta\grave{A}_\alpha - \left(\exp\left(-\tau\theta\right)'\tau\partial_\beta\theta \circ \left(\exp\tau\theta\right)'\tau\partial_\alpha\theta\right)$

$+\exp\left(-\tau\theta\right) \circ \left(\exp\tau\kappa\right)"\left(\tau\partial_\beta\theta, \tau\partial_\alpha\theta\right) + \exp\left(-\tau\theta\right) \circ \exp\left(\tau\theta\right)'\tau\partial_{\alpha\beta}^2\theta)$

$\mathcal{F}_{G\alpha\beta}^a \rightarrow \widetilde{\mathcal{F}}_{G\alpha\beta}\left(\tau\right) = Ad_{\exp\tau\kappa}\mathcal{F}_{G\alpha\beta}$

$\mathcal{F}_{A\alpha\beta}^a \rightarrow \widetilde{\mathcal{F}}_{A\alpha\beta}\left(\tau\right) = Ad_{\exp\tau\theta}\mathcal{F}_{A\alpha\beta}$

$F_{G\alpha\beta} \rightarrow Ad_{\exp\tau\kappa}F_{G\alpha\beta}$

$+\left[\left(\exp -\tau\kappa\right)\left(\exp\tau\kappa\right)'\tau\partial_\beta\kappa, G_\alpha - t\partial_\alpha\kappa\right] + \left[\left(\exp -\tau\kappa\right)\left(\exp\tau\kappa\right)'\tau\partial_\alpha\kappa, G_\beta - \tau\partial_\beta\kappa\right]$

$-Ad_{\exp\tau\kappa}(\left(\exp -\tau\kappa\right)'\tau\partial_\beta\kappa \circ \left(\exp\tau\kappa\right)'\tau\partial_\alpha\kappa + \exp\left(-\tau\kappa\right) \circ \left(\exp\tau\kappa\right)"\left(\tau\partial_\beta\kappa, \tau\partial_\alpha\kappa\right)$

$+\exp\left(-\tau\kappa\right) \circ \exp\left(\tau\kappa\right)'\tau\partial_{\alpha\beta}^2\kappa)$

$-Ad_{\exp\tau\kappa}(\left(\exp -\tau\kappa\right)'\tau\partial_\alpha\kappa \circ \left(\exp\tau\kappa\right)'\tau\partial_\beta\kappa + \exp\left(-\tau\kappa\right) \circ \left(\exp\tau\kappa\right)"\left(\tau\partial_\alpha\kappa, \tau\partial_\beta\kappa\right)$

$+\exp(-\tau\kappa)\circ\exp(\tau\kappa)'\tau\partial^2_{\alpha\beta}\kappa)$
$F_{A\alpha\beta}\to Ad_{\exp(\tau\theta}F_{A\alpha\beta}$
$+\left[(\exp-\tau\theta)(\exp\tau\theta)'\tau\partial_\beta\theta,\grave{A}_\alpha-\tau\partial_\alpha\theta\right]+\left[(\exp-\tau\theta)(\exp\tau\theta)'\tau\partial_\alpha\theta,\grave{A}_\beta-\tau\partial_\beta\theta\right]$
$-Ad_{\exp\tau\theta}(\exp(-\tau\theta)'\tau\partial_\beta\theta\circ(\exp\tau\theta)'\tau\partial_\alpha\theta+\exp(-\tau\theta)\circ(\exp\tau\kappa)"(\tau\partial_\beta\theta,\tau\partial_\alpha\theta)$
$+\exp(-\tau\theta)\circ\exp(\tau\theta)'\tau\partial^2_{\alpha\beta}\theta)$
$-Ad_{\exp\tau\theta}(\exp(-\tau\theta)'\tau\partial_\alpha\theta\circ(\exp\tau\theta)'\tau\partial_\beta\theta+\exp(-\tau\theta)\circ(\exp\tau\kappa)"(\tau\partial_\alpha\theta,\tau\partial_\beta\theta)$
$+\exp(-\tau\theta)\circ\exp(\tau\theta)'\tau\partial^2_{\alpha\beta}\theta)$

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 :

$\frac{d}{d\tau}\widetilde{P}^\alpha_i(m,\tau)|_{\tau=0}=-\sum_a\kappa^a([P][\kappa_a])^\alpha_i$

$\frac{d}{d\tau}\operatorname{Re}\widetilde{\psi}^{ij}(m,\tau)|_{\tau=0}$
$=\sum_a\kappa^a\sum_{k=1}^4\left(\operatorname{Re}\left([\gamma C(\kappa_a)]^i_k\right)\operatorname{Re}\psi^{kj}-\operatorname{Im}\left([\gamma C(\kappa_a)]^i_k\right)\operatorname{Im}\psi^{kj}\right)$
$=\sum_a\kappa^a\operatorname{Re}([\gamma C(\kappa_a)][\psi])^i_j$

$\frac{d}{d\tau}\operatorname{Im}\widetilde{\psi^{ij}}(m,\tau)|_{\tau=0}$
$=\sum_a\kappa^a\sum_{k=1}^4\left(\operatorname{Re}\left([\gamma C(\kappa_a)]^i_k\right)\operatorname{Im}\psi^{kj}+\operatorname{Im}\left([\gamma C(\kappa_a)]^i_k\right)\operatorname{Re}\psi^{kj}\right)$
$=\sum_a\kappa^a\operatorname{Im}([\gamma C(\kappa_a)][\psi])^i_j$

$\frac{d}{d\tau}\partial_\beta\widetilde{P}(m,t)^\alpha_j|_{\tau=0}=-\sum_a\kappa^a([\partial_\beta P][\kappa_a])^\alpha_i+\partial_\beta\kappa^a([P][\kappa_a])^\alpha_i$

$\frac{d}{d\tau}\operatorname{Re}\widetilde{\nabla_\alpha\psi}^{ij}(m,\tau)|_{\tau=0}$
$=\sum_a\kappa^a\sum_{k=1}^4\left(\operatorname{Re}\left([\gamma C(\kappa_a)]^i_k\right)\operatorname{Re}\nabla_\alpha\psi^{kj}-\operatorname{Im}\left([\gamma C(\kappa_a)]^i_k\right)\operatorname{Im}\nabla_\alpha\psi^{kj}\right)$
$=\sum_a\kappa^a\operatorname{Re}([\gamma C(\kappa_a)][\nabla_\alpha\psi])^i_j$

$\frac{d}{d\tau}\operatorname{Im}\widetilde{\nabla_\alpha\psi^{ij}}(m,\tau)|_{\tau=0}$
$=\sum_a\kappa^a\sum_{k=1}^4\left(\operatorname{Re}\left([\gamma C(\kappa_a)]^i_k\right)\operatorname{Im}\nabla_\alpha\psi^{kj}+\operatorname{Im}\left([\gamma C(\kappa_a)]^i_k\right)\operatorname{Re}\nabla_\alpha\psi^{kj}\right)$
$=\sum_a\kappa^a\operatorname{Im}([\gamma C(\kappa_a)][\nabla_\alpha\psi])^i_j$

$\frac{d}{d\tau}\widetilde{G_\alpha^a}(m)|_{\tau=0} = \sum_b \kappa^b [\overrightarrow{\kappa}_b, G_\alpha]^a - \partial_\alpha \kappa^a$
$\frac{d}{d\tau}\partial_\beta \widetilde{G}_\alpha^a(m,\tau)|_{\tau=0} = \sum_b \kappa^b [\overrightarrow{\kappa}_b, \partial_\beta G_\alpha]^a + \partial_\beta \kappa^b [\overrightarrow{\kappa}_b, G_\alpha]^a - \partial_{\alpha\beta}\kappa^a$
$\frac{d}{d\tau}\widetilde{\mathcal{F}}_{G\alpha\beta}^a(\tau)|_{\tau=0} = \sum_b \kappa^b [\overrightarrow{\kappa}_b, \mathcal{F}_{G\alpha\beta}]^a$
$\frac{d}{d\tau}\widetilde{F}_{G\alpha\beta}^a|_{\tau=0} = \sum_b \kappa^b [\overrightarrow{\kappa}_b, F_{G\alpha\beta}]^a + \partial_\beta\kappa^b [\overrightarrow{\kappa}_b, G_\alpha]^a + \partial_\alpha\kappa^b [\overrightarrow{\kappa}_b, G_\beta]^a - 2\partial_{\alpha\beta}\kappa^a$
So we have the identity :
$\forall \kappa_a, \partial_\beta\kappa^a, \partial_{\alpha\beta}\kappa^a$:
$0 =$
$\sum_a \kappa^a \{\sum_{ij} \frac{\partial L}{\partial \operatorname{Re}\psi^{ij}} \operatorname{Re}([\gamma C(\kappa_a)][\psi])_j^i + \frac{\partial L}{\partial \operatorname{Im}\psi^{ij}} \operatorname{Im}([\gamma C(\kappa_a)][\psi])_j^i$
$+\sum_{\alpha ij} \frac{\partial L}{\partial \operatorname{Re}\nabla_\alpha\psi^{ij}} \operatorname{Re}([\gamma C(\kappa_a)][\nabla_\alpha\psi])_j^i + \frac{\partial L}{\partial \operatorname{Im}\nabla_\alpha\psi^{ij}} \operatorname{Im}([\gamma C(\kappa_a)][\nabla_\alpha\psi])_j^i\}$
$+\sum_{i\alpha} \frac{\partial L}{\partial P_i^\alpha}(-\sum_a \kappa^a ([P][\kappa_a])_i^\alpha)$
$+\sum_{i\alpha\beta} \frac{\partial \mathcal{L}}{\partial \partial_\beta P_i^\alpha}\left(-\sum_a \kappa^a ([\partial_\beta P][\kappa_a])_j^\alpha + \partial_\beta\kappa^a ([P][\kappa_a])_j^\alpha\right)$
$+\sum_{a\alpha} \frac{\partial L}{\partial G_\alpha^a}\left(\sum_b \kappa^b [\overrightarrow{\kappa}_b, G_\alpha]^a - \partial_\alpha\kappa^a\right)$
$+\sum_{a\alpha\beta} \frac{\partial L}{\partial \mathcal{F}_{G\alpha\beta}^a}\left(\sum_b \kappa^b [\overrightarrow{\kappa}_b, \mathcal{F}_{G\alpha\beta}]^a\right)$
$+\frac{\partial L}{\partial F_{G\alpha\beta}^a}\left(\sum_b \kappa^b [\overrightarrow{\kappa}_b, F_{G\alpha\beta}]^a + \partial_\beta\kappa^b [\overrightarrow{\kappa}_b, G_\alpha]^a + \partial_\alpha\kappa^b [\overrightarrow{\kappa}_b, G_\beta]^a - 2\partial_{\alpha\beta}\kappa^a\right)$
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 $\partial_\alpha\kappa^a : \forall a, \alpha : \sum_{\beta i} \frac{\partial L}{\partial \partial_\alpha P_i^\beta}([P][\kappa_a])_i^\beta = -\frac{\partial L}{\partial G_\alpha^a}$
And we are left with :
$\forall a = 1..6$:
$0 =$
$\sum_{ij} \frac{\partial L}{\partial \psi^{ij}}([\gamma C(\kappa_a)][\psi])_j^i + \sum_{\alpha ij} \frac{\partial L}{\partial \nabla_\alpha \psi^{ij}}([\gamma C(\kappa_a)][\nabla_\alpha\psi])_j^i$
$-\sum_{i\alpha} \frac{\partial L}{\partial P_i^\alpha}([P][\kappa_a])_i^\alpha - \sum_{i\alpha\beta} \frac{\partial \mathcal{L}}{\partial \partial_\beta P_i^\alpha}([\partial_\beta P][\kappa_a])_j^\alpha$
$+\sum_{b\alpha} \frac{\partial L}{\partial G_\alpha^b}[\overrightarrow{\kappa}_a, G_\alpha]^b + \sum_{a\alpha\beta} \frac{\partial L}{\partial \mathcal{F}_{G\alpha\beta}^b}[\overrightarrow{\kappa}_a, \mathcal{F}_{G\alpha\beta}]^b$

Moreover, by taking the derivative with respect to the initial variables we get :
$\sum_{k=1}^4 [\gamma C(\exp(\tau\kappa(m)))]_i^k \frac{\partial \widetilde{L}}{\partial \widetilde{\psi^{kj}}} = \frac{\partial L}{\partial \psi^{ij}}$
$\sum_{k=1}^4 [\gamma C(\exp(\tau\kappa(m)))]_i^k \frac{\partial \widetilde{L}}{\partial \widetilde{\nabla_\alpha\psi^{kj}}} = \frac{\partial L}{\partial \nabla\psi^{ij}}$
$\sum_j [h(\exp(-\tau\kappa(m)))]_i^j \frac{\partial \widetilde{L}}{\partial \widetilde{P_j^\alpha}} = \frac{\partial L}{\partial P_i^\alpha}$
$L\left([Ad_{\exp\tau\kappa}]_b^a \mathcal{F}_{G\alpha\beta}^b\right) = L(\mathcal{F}_{G\alpha\beta})$
$\sum_b [Ad_{\exp\tau\kappa}]_a^b \frac{\partial \widetilde{L}}{\partial \widetilde{\mathcal{F}_{G\alpha\beta}^b}} = \frac{\partial L}{\partial \mathcal{F}_{G\alpha\beta}^a}$
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 $T_1 Spin(3,1)$: $\overrightarrow{\kappa}^a(\overrightarrow{\kappa}_b) = \delta_b^a$.

Equivariance on P_U

We have similarly :
$\frac{d}{d\tau}\widetilde{\psi^{ij}}(m,\tau)|_{\tau=0} = \sum_{k=1}^n \sum_{a=1}^m \theta^a \psi^{ik} [\theta_a]_j^k$
$\frac{d}{d\tau}\operatorname{Re}\widetilde{\psi^{ij}}(m,\tau)|_{\tau=0} = \sum_{a=1}^m \theta^a \operatorname{Re}(\psi[\theta_a])^{ij}$
$\frac{d}{d\tau}\operatorname{Im}\widetilde{\psi^{ij}}(m,\tau)|_{\tau=0} = \sum_{a=1}^m \theta^a \operatorname{Im}(\psi[\theta_a])^{ij}$
$\frac{d}{d\tau}\widetilde{\grave{A}}_\alpha^a(m,\tau)|_{\tau=0} = \sum_{b=1}^m \theta^b \left[\overrightarrow{\theta}_b, \grave{A}_\alpha\right]^a - \partial_\alpha\theta^a$
$\frac{d}{d\tau}\operatorname{Re}\widetilde{\nabla_\alpha\psi^{ij}}(m,\tau)|_{\tau=0} = \sum_{a=1}^m \theta^a \operatorname{Re}(\nabla_\alpha\psi[\theta_a])^{ij}$

$\frac{d}{d\tau}\operatorname{Im}\widetilde{\nabla_\alpha\psi^{ij}}(m,\tau)|_{\tau=0}=\sum_{a=1}^m\theta^a\operatorname{Im}(\nabla_\alpha\psi[\theta_a])^{ij}$
$\frac{d}{d\tau}\partial_\beta\widetilde{\grave{A}}^a_\alpha(m,\tau)|_{\tau=0}=\sum_{b=1}^m\theta^b\left[\theta_b,\partial_\beta\grave{A}_\alpha\right]^a+\partial_\beta\theta^b\left[\theta_b,\grave{A}_\alpha\right]^a-\partial_{\alpha\beta}\theta^a$
$\frac{d}{d\tau}\widetilde{\mathcal{F}}_{A\alpha\beta}(\tau)|_{\tau=0}=\sum_{b=1}^m\theta^b\left[\overrightarrow{\theta}_b,\mathcal{F}_{A\alpha\beta}\right]^a$
$\frac{d}{d\tau}\widetilde{F}_{A\alpha\beta}|_{\tau=0}=\sum_{b=1}^m\theta^b[\theta_b,F_{A\alpha\beta}]^a+\partial_\beta\theta^b\left[\theta_b,\grave{A}_\alpha\right]^a++\partial_\alpha\theta^b\left[\theta_b,\grave{A}_\beta\right]^a-2\partial_{\alpha\beta}\theta^a$

$$\begin{aligned}&\sum_{ij}\frac{\partial L}{\partial\operatorname{Re}\psi^{ij}}\sum_{a=1}^m\theta^a\operatorname{Re}(\psi[\theta_a])^{ij}+\frac{\partial L}{\partial\operatorname{Im}\psi^{ij}}\sum_{a=1}^m\theta^a\operatorname{Im}(\psi[\theta_a])^{ij}\\&+\sum_{ij\alpha}\frac{\partial L}{\partial\operatorname{Re}\nabla_\alpha\psi^{ij}}\sum_{a=1}^m\theta^a\operatorname{Re}(\psi[\theta_a])^{ij}+\sum_{ij\alpha}\frac{\partial L}{\partial\operatorname{Im}\nabla_\alpha\psi^{ij}}\sum_{a=1}^m\theta^a\operatorname{Im}(\psi[\theta_a])^{ij}\\&+\sum_{a\alpha}\frac{\partial L}{\partial\grave{A}^a_\alpha}\left(\sum_{b=1}^m\theta^b\left[\overrightarrow{\theta}_b,\grave{A}_\alpha\right]^a-\partial_\alpha\theta^a\right)+\sum_{a\alpha\beta}\frac{\partial L}{\partial\mathcal{F}^a_{A\alpha\beta}}\left(\sum_{b=1}^m\theta^b\left[\overrightarrow{\theta}_b,\mathcal{F}_{A\alpha\beta}\right]^a\right)\\&+\sum_{a\alpha\beta}\frac{\partial L}{\partial F^a_{A\alpha\beta}}\left(\sum_{b=1}^m\theta^b[\theta_b,F_{A\alpha\beta}]^a+\partial_\beta\theta^b\left[\theta_b,\grave{A}_\alpha\right]^a+\partial_\alpha\theta^b\left[\theta_b,\grave{A}_\beta\right]^a-2\partial_{\alpha\beta}\theta^a\right)\\&=0\end{aligned}$$

Which implies :
$\forall a,\alpha,\beta:\frac{\partial L}{\partial F^a_{A\alpha\beta}}=0,\frac{\partial L}{\partial\grave{A}^a_\alpha}=0$
$\forall a=1..m:$
$\sum_{ij}\frac{\partial L}{\partial\psi^{ij}}(\psi[\theta_a])^{ij}+\sum_{ij\alpha}\frac{\partial L}{\partial\nabla_\alpha\psi^{ij}}(\nabla_\alpha\psi[\theta_a])^{ij}+\sum_{b\alpha\beta}\frac{\partial L}{\partial\mathcal{F}^b_{A\alpha\beta}}\left(\left[\overrightarrow{\theta}_a,\mathcal{F}_{A\alpha\beta}\right]^b\right)=0$

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 : $\sum_i\frac{\partial L}{\partial\psi^{ij}}\mathbf{f}^j,\sum_i\frac{\partial L}{\partial\nabla_\alpha\psi^{ij}}\mathbf{f}^j,\sum_a\frac{\partial L}{\partial\mathcal{F}^a_{A\alpha\beta}}\overrightarrow{\theta}^a$ with $\overrightarrow{\theta}^a$ the basis vector of the dual of T_1U : $\overrightarrow{\theta}^a\left(\overrightarrow{\theta}_b\right)=\delta^a_b$

Covariance

In a change of charts on M with the jacobian : $J=\left[J^\alpha_\beta\right]=\left[\frac{\partial\widetilde{\xi}^\alpha}{\partial\xi^\beta}\right]$ 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 :
ψ^{ij} do not change
The covariant derivatives are one form :
$\nabla_\alpha\psi^{ij}\to\widetilde{\nabla_\alpha\psi^{ij}}=\sum_\beta K^\beta_\alpha\nabla_\beta\psi^{ij}$
P,V are vectors, but their components are functions :
$V^\alpha\to\widetilde{V}^\alpha=\sum_\gamma J^\alpha_\gamma V^\gamma$
$P^\alpha_i\to\widetilde{P}^\alpha_i=\sum_\gamma J^\alpha_\gamma P^\gamma_i$
$\widetilde{\partial_\beta P^\alpha_i}=\frac{\partial}{\partial\widetilde{\xi}^\beta}\left(\sum_\gamma J^\alpha_\gamma(\xi)P^\gamma_i(\xi)\right)=\sum_\gamma\left(\frac{\partial}{\partial\widetilde{\xi}^\beta}J^\alpha_\gamma(\xi)\right)P^\gamma_i(\xi)+J^\alpha_\gamma(\xi)\frac{\partial}{\partial\widetilde{\xi}^\beta}P^\gamma_i(\xi)$
$\widetilde{\partial_\beta P^\alpha_i}=\sum_{\gamma\eta}\left(\partial_\eta J^\alpha_\gamma\right)K^\eta_\beta P^\gamma_i+\left(\left(\partial_\eta P^\gamma_i\right)J^\alpha_\gamma K^\eta_\beta\right)$
The potentials are 1-form :
$G^a_\alpha\to\widetilde{G}^a_\alpha=\sum_\beta K^\beta_\alpha G^a_\beta$
$\grave{A}^a_\alpha\to\widetilde{\grave{A}}^a_\alpha=\sum_\beta K^\beta_\alpha\grave{A}^a_\beta$
The strengths of the fields are 2-forms. They change as :
$\mathcal{F}^a_{G\alpha\beta}\to\widetilde{\mathcal{F}}^a_{G\alpha\beta}=\sum_{\{\gamma\eta\}=0}^3\mathcal{F}^a_{G\gamma\eta}\det[K]^{\{\gamma\eta\}}_{\{\alpha\beta\}}=\sum_{\gamma\eta=0}^3\mathcal{F}^a_{G\gamma\eta}K^\gamma_\alpha K^\eta_\beta$

So we have the identity :
$L\left(z^i, z^i_\alpha, z^i_{\alpha\beta}\right) = \widetilde{L}\left(\widetilde{z}^i, \widetilde{z}^i_\alpha, \widetilde{z}^i_{\alpha\beta}\right)$
$= \widetilde{L}\left(\widetilde{z}^i\left(z^i_\lambda, J^\lambda_\mu\right), \widetilde{z}^i_\alpha\left(z^i_\lambda, J^\lambda_\mu, \partial_\gamma J^\lambda_\mu\right), \widetilde{z}^i_{\alpha\beta}\left(z^i_\lambda, J^\lambda_\mu, \partial_\gamma J^\lambda_\mu, \partial^2_{\gamma\varepsilon} J^\lambda_\mu\right)\right).$
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 $\left(\partial_\eta J^\lambda_\mu\right)$:
$0 = \sum_{i\alpha\beta} \frac{\partial L}{\partial \partial_\beta P^\alpha_i} \sum_{\gamma\eta} K^\eta_\beta P^\gamma_i \delta^\alpha_\lambda \delta^\mu_\gamma = \sum_{\alpha\beta\eta i} \frac{\partial L}{\partial \partial_\beta P^\lambda_i} P^\mu_i K^\eta_\beta$
take $J^\lambda_\mu = \delta^\lambda_\mu \Rightarrow K^\eta_\beta = \delta^\eta_\beta$
$\sum_i \frac{\partial L}{\partial \partial_\eta P^\lambda_i} P^\mu_i = 0$
$\forall \alpha, \beta, \gamma : \sum_i \frac{\partial L}{\partial \partial_\alpha P^\beta_i} P^\gamma_i = 0$
by product with $P^{j\prime}_\gamma$ and summation : $\forall \alpha, \beta, j : \frac{\partial L}{\partial \partial_\alpha P^\beta_j} = 0$
and as we had :
$\forall a, \alpha : \sum_{\beta i} \frac{\partial L}{\partial \partial_\alpha P^\beta_i} \left([P][\kappa_a]\right)^\beta_i = -\frac{\partial L}{\partial G^a_\alpha} \Rightarrow \forall a, \alpha : \frac{\partial L}{\partial G^a_\alpha} = 0$
The derivative with respect to J^λ_μ :
$\sum_{i\alpha} \frac{\partial L}{\partial P^\alpha_i} \sum_\gamma P^\gamma_i \delta^\lambda_\alpha \delta^\mu_\gamma + \sum_{i\alpha} \frac{\partial L}{\partial \operatorname{Re} \nabla_\alpha \psi^{ij}} \sum_\beta \left(\frac{\partial}{\partial J^\lambda_\mu} K^\beta_\alpha\right) \operatorname{Re} \nabla_\beta \psi^{ij}$
$+ \sum_{i\alpha} \frac{\partial L}{\partial \operatorname{Im} \nabla_\alpha \psi^{ij}} \sum_\beta \left(\frac{\partial}{\partial J^\lambda_\mu} K^\beta_\alpha\right) \operatorname{Im} \nabla_\beta \psi^{ij}$
$+ \sum_{a\alpha\beta} \frac{\partial L}{\partial \mathcal{F}^a_{G\alpha\beta}} \sum_{\gamma\eta} \left(\left(\frac{\partial}{\partial J^\lambda_\mu} K^\gamma_\alpha\right) K^\eta_\beta + K^\gamma_\alpha \frac{\partial}{\partial J^\lambda_\mu} K^\eta_\beta\right) \mathcal{F}^a_{G\gamma\eta}$
$+ \frac{\partial L}{\partial \mathcal{F}^a_{A\alpha\beta}} \sum_{\gamma\eta} \left(\left(\frac{\partial}{\partial J^\lambda_\mu} K^\gamma_\alpha\right) K^\eta_\beta + K^\gamma_\alpha \frac{\partial}{\partial J^\lambda_\mu} K^\eta_\beta\right) \mathcal{F}^a_{A\gamma\eta} + \frac{\partial L}{\partial V^\alpha} \sum_\gamma V^\gamma \delta^\lambda_\alpha \delta^\mu_\gamma = 0$
with $\frac{\partial}{\partial J^\lambda_\mu} K^\beta_\alpha = -K^\beta_\lambda K^\mu_\alpha$
$\sum_{i\alpha} \frac{\partial L}{\partial P^\lambda_i} P^\mu_i + \sum_{ij\alpha} \frac{\partial L}{\partial \nabla_\alpha \psi^{ij}} \sum_\beta \left(-K^\beta_\lambda K^\mu_\alpha\right) \nabla_\beta \psi^{ij}$
$+ \sum_{a\alpha\beta} \frac{\partial L}{\partial \mathcal{F}^a_{G\alpha\beta}} \sum_{\gamma\eta} \left(\left(\left(-K^\gamma_\lambda K^\mu_\alpha\right)\right) K^\eta_\beta + K^\gamma_\alpha \left(-K^\eta_\lambda K^\mu_\beta\right)\right) \mathcal{F}^a_{G\gamma\eta}$
$+ \frac{\partial L}{\partial \mathcal{F}^a_{A\alpha\beta}} \sum_{\gamma\eta} \left(\left(\left(-K^\gamma_\lambda K^\mu_\alpha\right)\right) K^\eta_\beta + K^\gamma_\alpha \left(-K^\eta_\lambda K^\mu_\beta\right)\right) \mathcal{F}^a_{A\gamma\eta} + \frac{\partial L}{\partial V^\lambda} V^\mu = 0$
Let us take $J^\lambda_\mu = \delta^\lambda_\mu \Rightarrow K^\lambda_\mu = \delta^\lambda_\mu$
$\sum_i \frac{\partial L}{\partial P^\lambda_i} P^\mu_i - \sum_{i\alpha} \frac{\partial L}{\partial \nabla_\mu \psi^{ij}} \nabla_\lambda \psi^{ij} - \sum_{a\eta} \frac{\partial L}{\partial \mathcal{F}^a_{G\mu\eta}} \mathcal{F}^a_{G\lambda\eta} - \sum_{a\gamma} \frac{\partial L}{\partial \mathcal{F}^a_{G\gamma\mu}} \mathcal{F}^a_{G\gamma\lambda} - \sum_{a\eta} \frac{\partial L}{\partial \mathcal{F}^a_{A\mu\eta}} \mathcal{F}^a_{A\lambda\eta} - \sum_{a\gamma} \frac{\partial L}{\partial \mathcal{F}^a_{A\gamma\mu}} \mathcal{F}^a_{A\gamma\lambda} + \frac{\partial L}{\partial V^\lambda} V^\mu = 0$
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 :
$\widetilde{L}\left(\widetilde{P^\alpha_i}, \widetilde{\psi^{ij}}, \widetilde{\nabla_\alpha \psi^{ij}}, \widetilde{\mathcal{F}^a_{A\alpha\beta}}, \widetilde{\mathcal{F}^a_{G\alpha\beta}}, \widetilde{V}^\alpha\right)$
$= \widetilde{L}\left(\widetilde{P^\alpha_i}\left(P^\lambda_i\right), \psi^{ij}, \widetilde{\nabla_\alpha \psi^{ij}}\left(\nabla_\lambda \psi^{pq}\right), \widetilde{\mathcal{F}^a_{A\alpha\beta}}\left(\mathcal{F}^b_{A\lambda\mu}\right), \widetilde{\mathcal{F}^a_{G\alpha\beta}}\left(\mathcal{F}^b_{G\lambda\mu}\right), \widetilde{V}^\alpha\left(V^\lambda\right)\right)$
$= L\left(P^\alpha_i, \psi^{ij}, \nabla_\alpha \psi^{ij}, \mathcal{F}^a_{A\alpha\beta}, \mathcal{F}^a_{G\alpha\beta}, V^\alpha\right)$
$\frac{\partial \widetilde{L}}{\partial \widetilde{P^\alpha_i}} \frac{\partial \widetilde{P^\alpha_i}}{\partial P^\lambda_i} = \frac{\partial \widetilde{L}}{\partial \widetilde{P^\alpha_i}} J^\alpha_\lambda = \frac{\partial L}{\partial P^\lambda_i}$
$\frac{\partial \widetilde{L}}{\partial \widetilde{\nabla_\alpha \psi^{ij}}} \frac{\partial \widetilde{\nabla_\alpha \psi^{ij}}}{\partial \nabla_\lambda \psi^{ij}} = \frac{\partial \widetilde{L}}{\partial \widetilde{\nabla_\alpha \psi^{ij}}} K^\lambda_\alpha = \frac{\partial L}{\partial \nabla_\lambda \psi^{ij}}$
$\frac{\partial \widetilde{L}}{\partial \widetilde{\mathcal{F}^a_{G\alpha\beta}}} \frac{\partial \widetilde{\mathcal{F}^a_{G\alpha\beta}}}{\partial \mathcal{F}^a_{B\lambda\mu}} = \frac{\partial \widetilde{L}}{\partial \widetilde{\mathcal{F}^a_{G\alpha\beta}}} K^\lambda_\alpha K^\mu_\beta = \frac{\partial L}{\partial \mathcal{F}^a_{G\lambda\mu}}$
$\frac{\partial \widetilde{L}}{\partial \widetilde{\mathcal{F}^a_{A\alpha\beta}}} \frac{\partial \widetilde{\mathcal{F}^a_{G\alpha\beta}}}{\partial \mathcal{F}^a_{A\lambda\mu}} = \frac{\partial \widetilde{L}}{\partial \widetilde{\mathcal{F}^a_{A\alpha\beta}}} K^\lambda_\alpha K^\mu_\beta = \frac{\partial L}{\partial \mathcal{F}^a_{A\lambda\mu}}$
$\frac{\partial \widetilde{L}}{\partial \widetilde{V^\alpha}} \frac{\partial \widetilde{V^\alpha}}{\partial V^\beta} = \frac{\partial \widetilde{L}}{\partial \widetilde{V^\alpha}} J^\alpha_\beta = \frac{\partial L}{\partial V^\beta}$

which shows that the corresponding quantities are tensors : in TM^* for $\frac{\partial L}{\partial P_i^\lambda}, \frac{\partial L}{\partial V^\alpha}$ and in $TM \otimes TM$ for $\frac{\partial L}{\partial \nabla_\lambda \psi^{ij}}, \frac{\partial L}{\partial \mathcal{F}^a_{G\lambda\mu}}, \frac{\partial L}{\partial \mathcal{F}^a_{A\lambda\mu}}$.

Conclusion

i) The potentials $\grave{A}, G$, and the derivatives $\partial_\beta P_i^\alpha$ 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 :

$$L = L\left(\psi, \nabla_\alpha \psi, P_i^\alpha, \mathcal{F}_{G\alpha\beta}, \mathcal{F}_{A\alpha\beta}, V^\alpha\right) \tag{6.4}$$

ii) The following quantities are tensors :
$\Pi_\psi = \sum_{ij} \frac{\partial L}{\partial \psi^{ij}} \mathbf{e}^i \otimes \mathbf{f}^i$
$\Pi_\nabla = \sum_\alpha \frac{\partial L}{\partial \nabla_\alpha \psi^{ij}} \partial\xi_\alpha \otimes \mathbf{e}^i \otimes \mathbf{f}^i$
$\Pi_P = \sum_\alpha \frac{\partial L}{\partial P_i^\alpha} d\xi^\alpha \otimes \varepsilon^i$
$\Pi_A = \sum_{\alpha\beta} \frac{\partial L}{\partial \mathcal{F}^a_{A\alpha\beta}} \partial\xi_\alpha \wedge \partial\xi_\beta \otimes \overrightarrow{\theta}^a$
$\Pi_G = \sum_{\alpha\beta} \frac{\partial L}{\partial \mathcal{F}^a_{G\alpha\beta}} \partial\xi_\alpha \wedge \partial\xi_\beta \otimes \overrightarrow{\kappa}^a$
$\Pi_V = \sum_\alpha \frac{\partial L}{\partial V^\alpha} d\xi^\alpha$
and similarly $\sum_{\alpha\beta} \upsilon^* \left(\frac{\partial L}{\partial \mathcal{F}_{r\alpha\beta}}, \frac{\partial L}{\partial \mathcal{F}_{w\alpha\beta}}\right) \partial\xi_\alpha \wedge \partial\xi_\beta$

$\Pi_\nabla, \Pi_P, \Pi_A, \Pi_G, \Pi_V$ are associated to the variables $\psi, P, \grave{A}, G, V$ and appear in the Energy-Momentum tensor. Notice that these quantities, when $\det[P']$ is added to L, are no longer covariant.

iii) We have the identities
$\forall a = 1..6 : \Pi_\psi \left[\gamma C\left(\overrightarrow{\kappa}_a\right)\right]\psi + \sum_\alpha \Pi_\nabla^\alpha \left[\gamma C\left(\overrightarrow{\kappa}_a\right)\right] \nabla_\alpha \psi - \Pi_P [P][\kappa_a] + \sum_{b\alpha\beta} \Pi_{Gb}^{\alpha\beta} \left[\overrightarrow{\kappa}_a, \mathcal{F}_{G\alpha\beta}\right]^b = 0$
$\forall a = 1..m : \left(\Pi_\psi \psi + \sum_\alpha \Pi_\nabla^\alpha \nabla_\alpha \psi\right)[\theta_a] + \sum_{b\alpha\beta} \Pi_{Ab}^{\alpha\beta} \left[\overrightarrow{\theta}_a, \mathcal{F}_{A\alpha\beta}\right]^b = 0$
$\forall \alpha, \beta : \Pi_\nabla^\beta \nabla_\alpha \psi + \sum_{a\gamma} \Pi_{Ga}^{\beta\gamma} \mathcal{F}^a_{G\alpha\gamma} + \Pi_{Aa}^{\beta\gamma} \mathcal{F}^a_{A\alpha\gamma} - \sum_i \Pi^i_{P\alpha} P_i^\beta = \frac{\partial L}{\partial V^\alpha} V^\beta$

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 :

$\int_\Omega L_1\left(P_i^\alpha, \mathcal{F}_{G\alpha\beta}, \mathcal{F}_{A\alpha\beta}\right)\varpi_4 + \sum_{p=1}^N \int_0^T L_2\left(\psi_p, \nabla_\alpha \mathcal{M}_p, P_i^\alpha, V_p^\alpha\right)(t)\,dt$

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 $\mathcal{F}_A, \mathcal{F}_G$, and the scalar lagrangian is summed with a density. 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 and provides the equations for the propagations of the field. Combined with the more general results of the previous chapter they give the value of the field in the vacuum, with respect to initial conditions.

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 comprises particles moving in known fields, or when the impact of the particles on the value of the field can be neglected, actually only the second part of the action is involved. We have a classic variational problem over the interval $[0, T]$ of the experiment. 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 $u = \frac{c}{\sqrt{-\langle V,V\rangle}} V$ which is the generalized Lorentz equation.

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.

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 [3]. 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 conservation of matter is accounted for by a continuity equation for the density μ.

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 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).

[3] 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.

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.

6.2.4 Two solutions

Our framework provides actually two potential solutions.

Continuous distribution of particles

If there is only one kind of particle (corresponding to a single vector ψ_0) they have the same behavior, both kinematic and under the action of the field. With not too many collisions one can then expect that their trajectories are similar, they should be the integral curves of a common vector field and the particles can be represented by a matter field, with a density μ following the continuity equation. The particles keep their intrinsic properties through ψ_0 and it is not necessary to introduce additional constraints for the conservation of charge or mass. The model can deal with the two components of motion : translation and rotation. The second integral takes the form :

$\int_\Omega L_2 \left(\psi, \nabla_\alpha \psi, P_i^\alpha, V^\alpha\right) \mu \varpi_4$

and solutions can be found with the usual Lagrange equations. The section ψ is defined through $\sigma(r, w)$ and the key variables are then maps : $r, w : M \to \mathbb{R}^3$.

However these equations give only general solutions, which shall be adjusted to the initial conditions (covering the local density and velocities).

Individual particles

If the number of particles vary (there are creations and annihilations of particles), or if there are collisions, the process is discontinuous, and other methods must be used (Chapter 8).

For a fixed, number N of particles, with known fundamental state ψ_{0p}, the second integral of the action reads :

$\sum_{p=1}^N \int_0^T L_2 \left(\psi_p, \widehat{\nabla}_\alpha \mathcal{M}_p, P_i^\alpha, V_p^\alpha\right)(t)\, dt$

There is no need for a density along the trajectory $q(t)$: a given particle follows an integral curve fixed by the origin of the trajectory and its tangent, and occupies a single location on $\Omega_3(t)$.

The first problem is mathematical : the fields and the particles are defined over domains which are manifolds of different dimensions, thus we have 2 integrals which are not of the same order, which precludes the usual method by Euler-Lagrange equations. It is common to put a Dirac's function in the integral for particles, but this, naive, solution is just a formal way to rewrite the same integral without any added value. There is actually a rigorous mathematical solution, by functional derivatives. However its results must be understood to be used properly.

The second problem is to find an adequate representation of the particles. If we use maps such as :

$[0, T] \to J^1 Q[E \otimes F, \vartheta] :: (q(t), \psi(t), \delta\psi(t))$

one cannot vary the trajectory $q(t)$, which is of course part of the variables.

In a given environment particles with the same fundamental state ψ_0 should behave the same way, their motion can then be represented by a single section $\sigma \in \mathfrak{X}(P_G)$, which defines a unique vector field V of trajectories and similarly for a section $\psi \in \mathfrak{X}(Q[E \otimes F, \vartheta])$. So wee can use a section as a "blue print" for the map $\psi(t)$. It sums up to take : $\psi(t) = \psi(q(t))$.

The parameter on the integral curves of V is the time of the observer, the variables $r, w : M \to \mathbb{R}^3$ become then $r, w : [0, T] \to \mathbb{R}^3$.

As in the first model we get general solutions which must be adjusted to the initial conditions.

In the next Chapter we will see a model with a matter field and a density, and a model with a fixed number of individual particles, both in the most general context, using the formalism presented in the book.

6.3 FINDING A SOLUTION

The implementation of the Principle of Least Action leads to the problem of finding sections on jet extension of 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. $\det[P'] = \sqrt{-\det[g]}$ is necessary to account for ϖ_4 which involves g.

In the lagrangian as well as in the derivatives $\frac{d\left(L \det P'\right)}{dz^i}, \frac{d\left(L \det P'\right)}{dz^i_\beta}$ the quantities which are involved are the components of a section of the 1st jet extension : z^i, z^i_α, seen as independent variables. The next step is to replace in the Lagrange equations the quantities z^i_α by the partial derivatives : $z^i_\alpha = \partial_\alpha z^i$ to get PDE in $Z = \left(z^i\right)_{i=1..n}$, section of E.

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) :

$div X = \sum_{\alpha=0}^3 \frac{\partial X^\alpha}{\partial \xi^\alpha} + \frac{1}{2} X^\alpha \sum_{\beta\gamma=0}^3 g^{\beta\gamma}\partial_\alpha g_{\beta\gamma}$ which reads in SR geometry : $div(X) = \sum_\alpha \partial_\alpha\left(X^\alpha\right)$

$\frac{dL}{dz^i_\beta}$ is a vector : $Z_i = \sum_\beta \frac{dL}{dz^i_\beta}\partial\xi_\beta$ 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 complex equation :

$\frac{\partial L \det P'}{\partial z^i} = \sum_\beta \frac{d}{d\xi^\beta} \left(\frac{\partial L \det P'}{\partial z^i_\beta} \right)$

Conservation laws

If for a variable we have $\frac{d\mathcal{L}}{dz^i} = 0$, then at equilibrium $\frac{1}{\det P'} \sum_\beta \frac{d}{d\xi^\beta} \frac{d\left(L \det P'\right)}{dz^i_\beta} = 0 = div\left(Z_i\right)$ with the vector $Z_i = \sum_\beta \frac{dL}{dz^i_\beta} \partial\xi_\beta$. The quantity Z_i is conserved at equilibrium. In particular with maps depending on t only : $\frac{d}{dt} \left(\frac{dL}{d\frac{dz^i}{dt}} \right) = 0$. Notice that $\frac{d\mathcal{L}}{dz^i}$ are total derivatives, meaning that the variable cannot appear as part of another variable, so this does not apply to the potentials.

If for a variable $\frac{d\left(L \det P'\right)}{dz^i_\beta} = 0$ then $\frac{d\mathcal{L}}{dz^i} = 0$ at equilibrium.

This is the case, in the more general lagrangian, for the tetrad P. The equation reads :

$\frac{1}{\det P'} \frac{d\left(L \det P'\right)}{dP^i_\beta} = 0 = \frac{dL}{dP^i_\beta} + L \frac{\partial \det P'}{\partial P^i_\alpha}$

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 :

$\forall i, \alpha : \frac{dL}{dP^\alpha_i} \det P' - L\left(\det P'\right) P'^i_\alpha = 0$

By product with P^β_i and summation on i the equations sum up to :

$$\forall \alpha, \beta = 0...3 : \sum_i \frac{dL}{dP^\alpha_i} P^\beta_i - L\delta^\alpha_\beta = 0 \tag{6.6}$$

The associated moment is a tensor : $\Pi_P = \sum_{\beta,j} \frac{\partial L}{\partial P^\beta_j} d\xi^\beta \otimes \varepsilon^j$.

The equation is equivalent to the conservation of energy (see below).

6.3.2 Functional derivatives

Whenever the system comprises force fields or matter fields on one hand, and individual particles on the other hand, such as :

$\int_\Omega L_1\left(P^\alpha_i, \mathcal{F}_{G\alpha\beta}, \mathcal{F}_{A\alpha\beta}\right) \varpi_4 + \sum_{p=1}^N \int_0^T L_2\left(\psi_p, \widehat{\nabla}_\alpha \mathcal{M}_p, P^\alpha_i, V^\alpha_p\right)(t)\, dt$

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 of variational problems, but is also a powerful tool to study the neighborhood of an equilibrium.

A functional : $\ell : J^r E \to \mathbb{R}$ defined on a normed subspace of sections $\mathfrak{X}(J^r E)$ of a vector bundle E has a functional derivative $\frac{\delta \ell}{\delta z}(Z_0)$ with respect to a section $Z \in \mathfrak{X}(E)$ in Z_0 if there is a distribution $\frac{\delta \ell}{\delta z}$ such that for any smooth, compactly supported $\delta Z \in \mathfrak{X}_{c\infty}(E)$:

$\lim_{\|\delta Z\| \to 0} \left\| \ell(Z_0 + \delta Z) - \ell(Z_0) - \frac{\delta \ell}{\delta z}(Z_0) Z \right\| = 0$

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 $\frac{\delta \ell}{\delta z}(Z_0) = 0$.

When the functional is linear in Z then $\frac{\delta \ell}{\delta z} = \ell$

When the functional is given by an integral : $\int_\Omega \lambda(J^r Z) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$ the functional derivative is the distribution :

$\frac{\delta \ell}{\delta z}(\delta Z) = \int_\Omega \sum_{s=0}^r \sum_{\alpha_1 \ldots \alpha_s} (-1)^s D_{\alpha_1 \ldots \alpha_s} \frac{\partial \lambda}{\partial Z_{\alpha_1 \ldots \alpha_s}} \delta Z d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

The functional can be the sum of integrals of different orders, then the method applies to the sum of the derivatives. Even if the result is usually expressed as an equality at each point, the equations must be understood “in the meaning of distributions” : for each measure of the quantities z^i, z^i_α done by the computation of the functional with any section $\delta Z \in \mathfrak{X}_{c\infty}(E)$ the equation holds. This is consistent with the physical process of measures, done by testing the value of the unknown variables with known quantities not pointwise, but over any compact area of Ω.

For a 1st order lagrangian the equations read :

$$\forall i : \frac{dL}{dZ^i} = \frac{d}{dt}\left(\frac{\partial L}{\partial \frac{dZ^i}{dt}}\right) \tag{6.7}$$

We will see how to implement this method in the next chapter.

6.4 ENERGY-MOMENTUM TENSOR

6.4.1 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.

The system is represented by variables, which are sections of some 1st jet prolongation of a fiber bundle E :

$\delta Z : \Omega \to J^1 E :: \delta Z(m) = \left(m, z^i(m), z^i_\alpha(m)\right)$

The action is then a functional :

$\ell : \mathfrak{X}\left(J^1 E\right) \to \mathbb{R} :: \ell\left(z^i, z^i_\alpha\right) = \int_\Omega \mathcal{L}\left(z^i, z^i_\alpha\right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

At equilibrium : $\widehat{Z} = \left(\widehat{z}^i, \widehat{z}^i_\alpha\right)$, the variational derivative $\frac{\delta \ell}{\delta z}\left(\widehat{\delta Z}\right)$, which is a distribution, is null for any smooth section δZ with compact support :

$\delta Z \in \mathfrak{X}_{\infty,c}\left(J^1 E\right) : \frac{\delta \ell}{\delta z}\left(\widehat{Z}\right)(\delta Z) = 0$

Consider a change $\delta Z = \left(\delta\zeta^i, \delta\zeta^i_\alpha\right) \in \mathfrak{X}_c\left(J^1 E\right)$ for a section with compact support, in the neighborhood of the equilibrium :

$\delta \ell = \ell\left(\widehat{\delta Z} + \delta\zeta\right) - \ell\left(\widehat{\delta Z}\right) \simeq \int_\Omega \left(\sum_i \frac{\partial \mathcal{L}}{\partial z^i}\delta\zeta^i + \sum_{i,\alpha} \frac{\partial \mathcal{L}}{\partial z^i_\alpha}\delta\zeta^i_\alpha\right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

Whenever δZ is the prolongation of a section of E, that is : $\delta\zeta^i_\alpha = \partial_\alpha \zeta^i$, by definition of the variational derivative :

$\delta \ell = \frac{\delta \ell}{\delta z^i}(\delta\zeta) = \int_\Omega \sum_i \left(\frac{\partial \mathcal{L}}{\partial z^i} - \sum_\alpha \frac{d}{d\xi^\alpha}\left(\frac{\partial \mathcal{L}}{\partial z^i_\alpha}\right)\right)\left(\delta\zeta^i\right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

At equilibrium : $\forall \delta Z : \frac{\delta \ell}{\delta z^i}\left(\widehat{\delta Z}\right)(\delta\zeta) = 0$

$\int_\Omega \sum_i \frac{\partial \mathcal{L}}{\partial z^i}\left(\delta\zeta^i\right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3 = \int_\Omega \sum_i \left(\sum_\alpha \frac{d}{d\xi^\alpha}\left(\frac{\partial \mathcal{L}}{\partial z^i_\alpha}\right)\right)\left(\delta\zeta^i\right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

Thus, for any $\delta Z \in \mathfrak{X}_{\infty,c}(E)$:

$\delta \ell = \int_\Omega \sum_i \left(\sum_\alpha \frac{d}{d\xi^\alpha}\left(\frac{\partial \mathcal{L}}{\partial z^i_\alpha}\right)\right)\left(\delta\zeta^i\right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3 + \int_\Omega \left(\sum_{i,\alpha} \frac{\partial \mathcal{L}}{\partial z^i_\alpha}\delta\zeta^i_\alpha\right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

Let $\mathcal{L} = L \det P'$

$\sum_i \left(\sum_\alpha \frac{d}{d\xi^\alpha}\left(\frac{\partial \mathcal{L}}{\partial z^i_\alpha}\right)\right)\left(\delta\zeta^i\right) = \sum_i \sum_\alpha \frac{d}{d\xi^\alpha}\left(\frac{\partial L}{\partial z^i_\alpha}\det P'\right)\left(\delta\zeta^i\right)$

$S = \sum_{\alpha,i} \frac{\partial L}{\partial z^i_\alpha}\partial\xi_\alpha \otimes \varepsilon^i$ is a tensor (see Lagrangian)

$S(\delta\zeta) = \sum_i \delta\zeta^i \sum_\alpha \frac{\partial L}{\partial z^i_\alpha}\partial\xi_\alpha$ is a vector field

$div S(\delta\zeta) = \frac{1}{\sqrt{-\det g}}\sum_{\alpha=0}^3 \frac{d}{d\xi^\alpha}\left(S^\alpha(\delta\zeta)\sqrt{-\det g}\right)$

$\int_\Omega \sum_i \left(\sum_\alpha \frac{d}{d\xi^\alpha}\left(\frac{\partial \mathcal{L}}{\partial z^i_\alpha}\right)\right)\left(\delta\zeta^i\right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$= \int_\Omega (div S(\delta\zeta)) \det P' d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$= \sum_i \int_\Omega (div S(\delta\zeta))\varpi_4$

$\sum_{i,\alpha} \frac{\partial \mathcal{L}}{\partial z^i_\alpha}\delta\zeta^i_\alpha = \sum_{i,\alpha} S^\alpha_i \delta\zeta^i_\alpha \det P'$

$\int_\Omega \left(\sum_{i,\alpha} \frac{\partial \mathcal{L}}{\partial z^i_\alpha}\delta\zeta^i_\alpha\right) d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$= \int_\Omega \sum_{i,\alpha} S^\alpha_i \delta\zeta^i_\alpha \det P' d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$= \int_\Omega \sum_{i,\alpha} S^\alpha_i \delta\zeta^i_\alpha \varpi_4$

$\sum_{i,\alpha} S^\alpha_i \delta\zeta^i_\alpha$ is a function

$\delta \ell = \int_\Omega \left(div(S(\delta\zeta)) + \sum_{i,\alpha} S^\alpha_i \delta\zeta^i_\alpha\right)\varpi_4$

and $\delta \ell$ is the integral of 2 functions, with the volume form ϖ_4.

Because δZ is the prolongation of a section of E, $\delta\zeta^i_\alpha = \partial_\alpha \zeta^i$, the variation of z^i along a vector field V is : $\delta\zeta^i = \sum_{\alpha=0}^3 \partial_\beta \zeta^i V^\beta$ thus $S_i(\delta\zeta) = \sum_{\alpha\beta} \frac{\partial L}{\partial z^i_\alpha}\partial_\beta \zeta^i V^\beta \partial\xi_\alpha$

The quantity : $T=\sum_{i\alpha\beta}\frac{\partial L}{\partial z_{\alpha}^{i}}z_{\beta}^{i}\partial\xi_{\alpha}\otimes d\xi^{\beta}$ is a tensor. The quantity $S_{i}\left(\delta\zeta\right)=\sum_{\alpha\beta}\frac{\partial L}{\partial z_{\alpha}^{i}}\partial_{\beta}\zeta^{i}V^{\beta}\partial\xi_{\alpha}$ can be seen as the forces opposed by the system to a change of the variable z^{i} in the direction given by V, expressed in the holonomic basis at each point :

$\delta F_{i}\left(V\right)=\sum_{\alpha\beta}\frac{\partial L}{\partial z_{\alpha}^{i}}\partial_{\beta}\zeta^{i}V^{\beta}\partial\xi_{\alpha}$

For a small compact area $\omega\in\Omega$ with boundary $\partial\omega$:

$\int_{\Omega}div\left(S\left(\delta\zeta\right)\right)\varpi_{4}=\int_{\Omega}div\left(T\left(V\right)\right)\varpi_{4}=\int_{\partial\omega}i_{T\left(V\right)}\varpi_{4}$

$\int_{\partial\omega}i_{T\left(V\right)}\varpi_{4}$ can be seen as the resultant of the forces exercised on $\partial\omega$, weighted by the density on $\partial\omega$.

The lagrangian has the meaning of a density of energy for the whole system.

The trace of the tensor T is the tensor : $Tr\left(T\right)=\sum_{i\alpha}\frac{\partial L}{\partial z_{\alpha}^{i}}z_{\alpha}^{i}d\xi^{\alpha}\in\mathfrak{X}\left(TM^{*}\right)$

$\sum_{i,\alpha}S_{i}^{\alpha}\delta\zeta_{\alpha}^{i}=\sum_{i,\alpha}\frac{\partial L}{\partial z_{\alpha}^{i}}\partial_{\alpha}\zeta^{i}=Tr\left(T\right)\left(V\right)$ has the meaning of the a variation of this energy due to the action of these forces.

Definition 100 *The Energy-Momentum of a system with action :*

$\ell:\mathfrak{X}\left(J^{1}E\right)\rightarrow\mathbb{R}::\ell\left(z^{i},z_{\alpha}^{i}\right)=\int_{\Omega}L\left(z^{i},z_{\alpha}^{i}\right)\varpi_{4}$

is the tensor :

$$T:\mathfrak{X}\left(J^{1}E\right)\rightarrow\mathfrak{X}\left(TM^{*}\otimes TM\otimes E^{*}\right)::T=\sum_{i\alpha\beta}\frac{\partial L}{\partial z_{\alpha}^{i}}z_{\beta}^{i}\partial\xi_{\alpha}\otimes d\xi^{\beta}\otimes e^{i}\qquad(6.8)$$

The quantities

$$\Pi_{i}=\sum_{\alpha\beta}\frac{\partial L}{\partial z_{\alpha}^{i}}z_{\beta}^{i}\partial\xi_{\alpha}\otimes d\xi^{\beta}\otimes e^{i}\in\mathfrak{X}\left(TM^{*}\otimes TM\otimes E^{*}\right)\qquad(6.9)$$

are the momenta associated to the variable Z^{t}.

To a change of the variable z^{i} in the direction V the system opposes at each point m a force :

$\delta F_{i}\left(V\right)=\Pi_{i}\left(V\right)\in T_{m}M$

and the change of energy of the system due to the action of these forces is given by the trace $Tr\left(T\right)$

$$\delta\ell=\int_{\Omega}\left(div\left(T\left(V\right)\right)+Tr\left(T\right)\left(V\right)\right)\varpi_{4}\qquad(6.10)$$

With any lagrangian one can compute an explicit energy-momentum tensor. It is usually assumed that the energy momentum tensor is symmetric : $T_{\beta}^{\alpha}=T_{\alpha}^{\beta}$, which it should be, in the Einstein equation, because the Ricci tensor, with the Lévy-Civita connection, is symmetric, but there is no reason why it should be so, and it is common to use a substitute of T in order to get a symmetric tensor.

If the equilibrium is kept : $\delta\ell=0$. Usually $\delta\ell\neq0$ for a change $\delta\zeta$ of the variable. However if the divergence of the vector $T\left(V\right)$ is null then $\delta\ell=0$. If such vectors V exist they show privileged directions over which the system can be deformed without energy spent, that is equivalent states of equilibrium.

6.4.2 General expression

With the more general lagrangian $\mathcal{L}=L\left(\psi,\nabla_{\alpha}\psi_{p},P_{i}^{\alpha},\mathcal{F}_{G\alpha\beta},\mathcal{F}_{A\alpha\beta},V^{\alpha}\right)\det P'$ the energy momentum tensor T reads :

$T=\sum_{\alpha\beta}\{\sum_{ij}\frac{\partial L}{\partial\partial_{\alpha}\psi^{ij}}\delta_{\beta}\psi^{ij}+\sum_{a\gamma}\frac{\partial L}{\partial\partial_{\alpha}\grave{A}_{\gamma}^{a}}\delta_{\beta}\grave{A}_{\gamma}^{a}+\sum_{a\gamma}\frac{\partial L}{\partial\partial_{\alpha}G_{\gamma}^{a}}\delta_{\beta}G_{\gamma}^{a}\}\partial\xi_{\alpha}\otimes d\xi^{\beta}$

Notice that $P_{i}^{\alpha},V^{\alpha}$ do not appear.

$\frac{\partial L}{\partial\partial_{\alpha}\psi^{ij}}=\frac{\partial L}{\partial\nabla_{\alpha}\psi^{ij}}$ and $\Pi_{\nabla}=\sum_{\alpha}\frac{\partial L}{\partial\nabla_{\alpha}\psi^{ij}}\partial\xi_{\alpha}\otimes\mathbf{e}^{i}\otimes\mathbf{f}^{i}$ is a tensor

$\frac{\partial L}{\partial \partial_\alpha \grave{A}^a_\gamma} = 2\frac{\partial L}{\partial \mathcal{F}^a_{A\alpha\gamma}}$ and $\Pi_A = \sum_{\alpha\gamma} \frac{\partial L}{\partial \mathcal{F}^a_{A\alpha\gamma}} \partial\xi_\alpha \wedge \partial\xi_\gamma \otimes \overrightarrow{\theta}^a$ is a tensor

$\frac{\partial L}{\partial \partial_\alpha G^a_\gamma} = 2\frac{\partial L}{\partial \mathcal{F}^a_{G\alpha\gamma}}$ and

$\Pi_G = \sum_{\alpha\beta} \frac{\partial L}{\partial \mathcal{F}^a_{G\alpha\gamma}} \partial\xi_\alpha \wedge \partial\xi_\gamma \otimes \overrightarrow{\kappa}^a$ as well as $\sum_{\alpha\beta} \upsilon^* \left(\frac{\partial L}{\partial \mathcal{F}_{r\alpha\gamma}}, \frac{\partial L}{\partial \mathcal{F}_{w\alpha\gamma}} \right) \partial\xi_\alpha \wedge \partial\xi_\gamma$

$$T = \sum_{\alpha\beta} \{ \sum_{ij} \frac{\partial L}{\partial \nabla_\alpha \psi^{ij}} \delta_\beta \psi^{ij} + 2 \sum_{a,\gamma} \left(\frac{\partial L}{\partial \mathcal{F}^a_{A\alpha\gamma}} \delta_\beta \grave{A}^a_\gamma + \frac{\partial L}{\partial \mathcal{F}^a_{G\alpha\gamma}} \delta_\beta G^a_\gamma \right) \} \partial\xi_\alpha \otimes d\xi^\beta \tag{6.11}$$

Conversely, the momenta can be derived from the Energy-Momentum tensor, in a way which is usual in fluid mechanics : $\Pi^\alpha_{\nabla ij} = \frac{\partial T}{\partial \partial_\beta \psi^{ij}}, ...$ This is the generalized version of the Hamilton equations. Moreover the generalized momenta Π are related (see Lagrangian).

The computation above is quite general, holds for any lagrangian, 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, defined on the same support. This remark will be useful when studying discontinuous processes.

Variables function of t

The method of functional derivatives allows to consider variables $z^i : [0,T] \to E$ which are defined over some interval $[0,T]$. The corresponding lagrangian is then $L\left(z^i, \frac{\delta z^i}{\delta t}\right)$ with the action $\ell\left(Z, \frac{\delta Z}{\delta t}\right) = \int_0^T L\left(z^i, \frac{\delta z^i}{\delta t}\right) dt$.

ℓ has a functional derivative $\frac{\delta \ell}{\delta z^i}$ with respect to z^i in $\widehat{Z}$ if there is a distribution $\frac{\delta \ell}{\delta z}$ such that, for any smooth map with compact support $\delta\zeta \in C_{\infty c}([0,T], E)$:

$\lim_{\|\delta Z\| \to 0} \left\| \ell\left(\widehat{Z} + \delta\zeta\right) - \ell\left(\widehat{Z}\right) - \frac{\delta \ell}{\delta z}\left(\widehat{Z}\right) \delta\zeta \right\| = 0$

and for a linear defined by an integral :

$\forall \delta\zeta \in C_{\infty c}([0,T]; E) : \frac{\delta \ell}{\delta z^i}(\delta\zeta) = \int_0^T \sum_i \left(\frac{\partial L}{\partial z^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \frac{dz^i}{dt}} \right) \right) (\delta\zeta^i)\, dt$

Then the condition for an equilibrium is given by : $\frac{\delta \ell}{\delta z^i}\left(\widehat{Z}\right) = 0.$

The computation done previously can be extended to variables depending on t.

For any variation $\delta\zeta^i$, smooth map $[0,T] \to E$ with compact support :

$\delta\ell = \ell\left(\widehat{Z} + \delta\zeta\right) - \ell\left(\widehat{Z}\right) \simeq \int_0^T \left(\sum_i \frac{\partial \mathcal{L}}{\partial z^i} \delta\zeta^i + \sum_{i,\alpha} \frac{\partial L}{\partial \frac{dz^t}{dt}} \frac{\delta z^i}{\delta t} \right) dt$

and from the functional derivative :

$\frac{\delta \ell}{\delta z^i}(\delta\zeta) = \int_0^T \sum_i \left(\frac{\partial \mathcal{L}}{\partial z^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \frac{dz^t}{dt}} \right) \right) (\delta\zeta^i)\, dt$

At equilibrium $\widehat{Z} : \forall \delta\zeta : \frac{\delta \ell}{\delta z^i}(\delta\zeta) = 0$

$\delta\ell \simeq \int_0^T \left(\sum_i \frac{d}{dt}\left(\frac{\partial L}{\partial \frac{dz^t}{dt}} \right) (\delta\zeta^i) + \sum_{i,\alpha} \frac{\partial L}{\partial \frac{dz^t}{dt}} \frac{\delta z^i}{\delta t} \right) dt$

because $\delta\zeta^i$ is a section its first jet extension : $\frac{\delta \zeta^i}{\delta t} = \frac{d}{dt} \delta z^i$

$\delta\ell \simeq \int_0^T \left(\sum_i \frac{d}{dt}\left(\frac{\partial L}{\partial \frac{dz^t}{dt}} \right) (\delta\zeta^i) + \sum_{i,\alpha} \frac{\partial L}{\partial \frac{dz^t}{dt}} \frac{d}{dt} \delta z^i \right) dt$

$= \int_0^T \frac{d}{dt} \left(\sum_i \left(\frac{\partial L}{\partial \frac{dz^t}{dt}} \right) (\delta\zeta^i) \right) dt = \left[\sum_i \delta\zeta^i \frac{\partial L}{\partial \frac{dz^t}{dt}} \right]_0^T$

$\sum_i \delta\zeta^i \frac{\partial L}{\partial \frac{dz^t}{dt}}$ is the equivalent of the Energy-Momentum tensor, for a change $\delta\zeta^i$ of z^i during the time interval δt, or equivalently of the forces of the system in resistance to the change.

6.4.3 Conservation of Momentum and Energy

The Principle of Least Action is complementary to the Conservation of Momentum or Energy. The lagrangian represents usually the sum of all the exchanges of energy between the objects in the system. And the Energy-Momentum tensor represents the inertial forces of the system. The lagrangian formalism gives a comprehensive picture of these Principles.

Conservation of Energy

Usually, and in the models which will be used in this book, the lagrangian represents the balance of energy between the components of the system. The energy of the system can then be defined as :

$\mathcal{E} = \int_\Omega L\left(z^i, z^i_\alpha\right)\varpi_4$

accounting for the variation of the volume measure with the metric.

The Principle of Least Action gives the conditions for an equilibrium, it is different from the Principle of Conservation of Energy, which states that the balance of energy of a system must be even at each time for an observer. So, if all the physical objects interacting in the system have been accounted for, we must have the additional condition [4]:

$\mathcal{E}(t) = \int_{\Omega(t)} L\left(z^i, z^i_\alpha\right)\varpi_3 = Ct = \int_{\Omega(t)} i_{\varepsilon_0}\left(L\left(z^i, z^i_\alpha\right)\varpi_4\right)$

The integral is on $\Omega(t)$ and not the whole of Ω

Consider the manifold $\Omega([t_1, t_2])$ with borders $\Omega(t_1), \Omega(t_2)$:

$\mathcal{E}(t_2) - \mathcal{E}(t_1) = \int_{\partial\Omega([t_1,t_2])} i_{\varepsilon_0}(L\varpi_4) = \int_{\Omega([t_1,t_2])} d\left(i_{\varepsilon_0}L\varpi_4\right)$

$d\left(i_{\varepsilon_0}L\varpi_4\right) = \pounds_{\varepsilon_0}(L\varpi_4) - i_{\varepsilon_0}d(L\varpi_4)$

$= \left(\pounds_{\varepsilon_0}L\right)\varpi_4 + L\pounds_{\varepsilon_0}\varpi_4 - i_{\varepsilon_0}(dL\wedge\varpi_4) - i_{\varepsilon_0}Ld\varpi_4$

$= L'(\varepsilon_0)\varpi_4 + L(div\varepsilon_0)\varpi_4 - i_{\varepsilon_0}(dL\wedge\varpi_4)$

$= div(L\varepsilon_0)\varpi_4$

$\mathcal{E}(t_2) - \mathcal{E}(t_1) = \int_{\Omega([t_1,t_2])} div(L\varepsilon_0)\varpi_4$

The conservation of energy for the observer imposes an additional condition : $div(L\varepsilon_0) = 0$ to the solutions, *specific to each observer* (through ε_0).

With the general lagrangian $\mathcal{L} = L\left(\psi, \nabla_\alpha\psi, P^\alpha_i, \mathcal{F}_{G\alpha\beta}, \mathcal{F}_{A\alpha\beta}, V^\alpha\right)\det P'$ the derivative of P^α_i does not appear, and the corresponding equation for the tetrad reads :

$\forall\alpha,\beta = 0...3 : \sum_i \frac{dL}{dP^\alpha_i}P^\beta_i - L\delta^\alpha_\beta = 0$

But $div(\varepsilon_0 L) = \sum_{\alpha=0}^3 \frac{\partial}{\partial\xi^\alpha}\left(\varepsilon^\alpha_0 L\det P'\right) = \frac{\partial}{\partial\xi^0}(L\det P')$

$= \frac{\partial}{c\partial t}(L\det P') = \det P'\frac{\partial L}{c\partial t} + L\frac{\partial}{c\partial t}(\det P')$

$= \det P'\frac{\partial L}{cdt} - L\sum_{i\alpha} P'^i_\alpha\frac{\partial P^i_\alpha}{\partial t}(\det P')$

thus :

$div(\varepsilon_0 L) = 0 \Leftrightarrow \frac{\partial L}{\partial t} = L\sum_{i\alpha} P'^i_\alpha\frac{\partial P^i_\alpha}{\partial t}$

The tetrad equation reads

$\frac{\partial L}{\partial P^\alpha_i} - LP'^i_\alpha = 0$

and on shell the condition sums up to the identity :

$\frac{\partial L}{\partial t} = \sum_{i\alpha} LP'^i_\alpha\frac{\partial P^i_\alpha}{\partial t} = \sum_{i\alpha}\frac{\partial L}{\partial P^\alpha_i}\frac{\partial P^i_\alpha}{\partial t}$

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),
- it can be derived by the use of functional derivatives, and does not require all the smoothness conditions imposed by the Lagrange equations,

[4] 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.

- it is based upon the variation of the metric, which appears as the quantity through which this balance of energy is kept.

Conservation of the momenta

In Newtonian Mechanics the Conservation of Momentum is actually the expression, in specific cases, of the general laws for the evolution of the system, and it could be expressed in a simple way because it is possible to define a center of mass, and so to give a physical meaning to the sum of the forces exercised on the system.

The momenta are vectorial quantities, defined in different vector spaces, and at different points, so their aggregation has no meaning in the GR picture. The Energy-Momentum tensor gives the equivalent of inertial forces opposed to the system at any change, if the equilibrium is kept the divergence of the vectors $T(V)$ should be null but it does not tell us anything pointwise.

6.5 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.

6.5.1 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 $\mathbf{R}$ 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 : $\mathbf{R}$ 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 $\mathbf{R}$ 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 in the next Chapter. It is more pedagogical, and opens the possibility to study a dissymmetric gravitational field. So we will take in a perturbative lagrangian :

$$\int_\Omega \left(\sum_{\alpha\beta} C_G \left(\sum_{a=1}^{3} \mathcal{F}^a_{G\alpha\beta} \mathcal{F}^{a\alpha\beta}_G - \sum_{a=4}^{6} \mathcal{F}^a_{G\alpha\beta} \mathcal{F}^{a\alpha\beta}_G \right) + C_A \sum_{a=1}^{m} \mathcal{F}^a_{A\alpha\beta} \mathcal{F}^{a\alpha\beta}_A \right) \varpi_4 (m) \tag{6.12}$$

where C_G, C_A are real constant scalars, which depend on the units. Notice that, for the convenience of computations, the quantities are defined as non ordered sum of indices. Moreover usually the EM field will be incorporated in the "other fields". More precisely we have :

$\sum_{\alpha\beta} \left\{ C_G \left(\sum_{a=1}^{3} \mathcal{F}^a_{G\alpha\beta} \mathcal{F}^{a\alpha\beta}_G - \sum_{a=4}^{6} \mathcal{F}^a_{G\alpha\beta} \mathcal{F}^{a\alpha\beta}_G \right) + C_A \sum_{a=1}^{m} \mathcal{F}^a_{A\alpha\beta} \mathcal{F}^{a\alpha\beta}_A + C_{EM} \mathcal{F}_{EM\alpha\beta} \mathcal{F}^{\alpha\beta}_{EM} \right\}$

$= \sum_{\alpha\beta} 4C_G \left\langle \mathcal{F}^{\alpha\beta}_G, \mathcal{F}_{G\alpha\beta} \right\rangle_{Cl} + C_A \left\langle \mathcal{F}^{\alpha\beta}_A, \mathcal{F}_{A\alpha\beta} \right\rangle_{T_1U} + C_{EM} \left\langle \mathcal{F}_{EM\alpha\beta}, \mathcal{F}^{\alpha\beta}_{EM} \right\rangle_{T_1U(1)}$

$= 8C_G \left\langle \mathcal{F}_G, \mathcal{F}_G \right\rangle_G + 2C_A \left\langle \mathcal{F}_A, \mathcal{F}_A \right\rangle_U + 2C_{EM} \left\langle \mathcal{F}_{EM\alpha\beta}, \mathcal{F}^{\alpha\beta}_{EM} \right\rangle_{T_1U(1)}$

the factor 2 accounting for the non ordered indices.

6.5.2 Interactions Particles /Fields

The logical term in the lagrangian is the variation of energy :

$\delta E = C_I \frac{1}{M_p} \frac{1}{i} \left\langle \psi, \nabla_V \psi \right\rangle = - C_I \frac{1}{2} M_p \left\{ k_0^t \operatorname{Re} \mathbf{Ad}_{\sigma^{-1}} \left(\upsilon (X_r, X_w) + \widehat{G} \right) + k_c^t \left(Ad_\varkappa \widehat{\grave{A}} \right) \right\}$

The fields act on the momentum of the particles
$\mathcal{M} = (m, \psi = \vartheta(\sigma, 1)\psi_0, \delta_\alpha\psi = \vartheta(\upsilon(X_{r\alpha}, X_{w\alpha})\cdot\sigma, 1)\psi_0) \in J^1Q[E\otimes F, \vartheta]$
through the covariant derivative :
$[\nabla_\alpha\psi] = \vartheta(\sigma, 1)\left(\left[\gamma C\left(\nabla_\alpha^G\sigma\right)\right][\psi_0] + [\psi_0]\left[\grave{A}_\alpha\right]\right)$
$\nabla_\alpha^G\sigma = \mathbf{Ad}_{\sigma^{-1}}(\upsilon(X_{r\alpha}, X_{w\alpha}) + G_\alpha)$
and along the trajectories, with vector $V = \sum_{\alpha=0}^3 V^\alpha\partial\xi_\alpha$ defined through σ :
$V = \frac{dq}{dt} = c\varepsilon_0 + \overrightarrow{v} = \sum_{\alpha,i}[P]_j^\alpha [U]^j \partial\xi_\alpha$
$U = -\frac{c}{\langle \mathbf{Ad}_\sigma\varepsilon_0, \varepsilon_0\rangle_{Cl}}\mathbf{Ad}_\sigma\varepsilon_0$
In a continuous motion : $\upsilon(X_{r\alpha}, X_{w\alpha}) = \partial_\alpha\sigma\cdot\sigma^{-1}$.
For the EM field : $k_c = -2q$

In a model with a density of particles each type of particles is represented by a matter field, a section $\psi \in \mathfrak{X}(Q[E\otimes F, \vartheta])$, with fixed ψ_0, and a density μ which follows the continuity equation $\frac{d\mu}{dt} + \mu divV = 0$. The key variable is then σ, defined through 2 maps $r, w : M \to \mathbb{R}^3$.
The action is then :
$\int_\Omega C_I \frac{1}{M}\frac{1}{i}\langle\psi(r(m), w(m)), \nabla_V\psi r(m), w(m)\rangle \mu(m)\varpi_4(m)$

In a model with a fixed number N of particles $p = 1...N$, each particle is represented in two steps
- the particles of the same type are represented by a matter field $\psi_p \in \mathfrak{X}(Q[E\otimes F, \vartheta])$, with fixed ψ_{0p}, corresponding to a common section $\sigma_p \in \mathfrak{X}(P_G)$;
- within the section the particle is such that σ_p depends on 2 maps $r_p, w_p : [0,T] \to \mathbb{R}^3$
So the state of the particle p is given by
$\psi_p(t) = \psi_p(q_p(t)) = \vartheta(\sigma_p(r_p(t), w_p(t)), 1)\psi_{0p}$
its velocity by $\sigma(r_p(t), w_p(t)), P_\alpha^i(q_p(t))$
$U_p^i = -\frac{c}{\left\langle \mathbf{Ad}_{\sigma_p}\varepsilon_0, \varepsilon_0\right\rangle_{Cl}}\mathbf{Ad}_{\sigma_p}\varepsilon_0$
$V_p^\alpha = \sum_{i=0}^3 P_i^\alpha(q_p(t))U_p^i$
the motion is assumed to be continuous
$\sum_{\alpha=0}^3 V^\alpha\partial_\alpha\sigma_p(r_p(t), w_p(t)) = \frac{d\sigma_p}{dt}$
which defines the trajectory $q_p(t)$.
The time t is the time of the observer, which is common to all the sections and particles.
The action is then :
$\sum_{p=1}^N \int_0^T C_I\frac{1}{M_p}\frac{1}{i}\langle\psi(r_p(t), w_p(t)), \nabla_V\psi(r_p(t), w_p(t))\rangle dt$

The lagrangian of the Standard Model [5] is similar, with the Dirac operator and $\grave{A}$ is identified with the bosons as force carriers (which requires the introduction of the Higgs boson).

6.5.3 Units

As said in the 2nd Chapter, any equation should be actually unitless, in order to be fully compliant with the rules in a change of gauge. However we use different, specialized units, in our measures. So there is the need for universal constants to make the bridge between them. The most obvious is the constant c : the basic unit in geometry is the length, acknowledge that "time" is a geometric measure leads to $\partial\xi_0 = c\partial t$.

The lagrangian, in its perturbative formulation, is the good place to look at the problem, as it involves all the quantities. The scalar lagrangian is a density of energy, with unit $[E] = [M][L]^2[T]^{-2}$.

[5] Of course the tools used in QTF to find solutions are quite different (the key variables are locat operators), but they are based on a pertubative lagrangian.

For particles $\frac{dK}{dt} = -C_I \frac{M_p}{2} k_0^t \operatorname{Re} \mathbf{Ad}_{\sigma^{-1}} \upsilon (X_r, X_w)$, the constant C_I has the dimension $[E]$ of energy, and $\upsilon (X_r, X_w)$ the dimension $[T]^{-1}$ so that $\epsilon \int_0^T -C_I \frac{1}{2} k_0^t M_p \operatorname{Re} (\mathbf{Ad}_{\sigma^{-1}} \upsilon (X_\alpha, Y_\alpha))\, dt$ has the dimension of energy.

The gravitational constant $\mathcal{G}$ has for units $[M]^{-1} [L]^3 [T]^{-2} \sim [E]^{-1} [L]^5 [T]^{-4}$ so that $\int_0^T -C_I \frac{1}{2} k_0^t M_p \operatorname{Re} \left(\mathbf{Ad}_{\sigma^{-1}} \widehat{G} \right) dt$ can equivalently be expressed as : $\frac{c^4}{\mathcal{G}} \int_0^T -\frac{1}{2} k_0^t M_p \operatorname{Re} \left(\mathbf{Ad}_{\sigma^{-1}} \widehat{G} \right) dt$ with G unitless

$\frac{c^4}{\mathcal{G}} = 1.208 \times 10^{44}\ m \times kg \times s^{-2}$ is the Planck's unit of force

$-\int_0^T C_I \frac{1}{2} M_p k_c^t \left(Ad_{\varkappa} \widehat{\grave{A}} \right) dt = -\int_0^T C_I M_p q \widehat{\grave{A}} dt$ for the EM field so that $q\grave{A}_\alpha$ has the dimension $[M]^{-1} [L]^{-1}$, $\grave{A}_\alpha$ has the dimension $[L]$ and q the dimension $[M]^{-1} [L]^{-2}$.

$\mathcal{F}_G = dG + j(G) G, \mathcal{F}_A = d\grave{A}$ are then unitless and C_G, C_A have the dimension of energy as C_I.

Chapter 7

CONTINUOUS MODELS

Continuous models represent systems where no discontinuous process occurs : the particles keep their fundamental state, without creation or annihilation, the trajectories do not cross, the motion is continuous, the maps are smooth. Continuous models correspond to an ideal situation, they are nevertheless useful to study the basic relations between the variables. 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 maps on vector bundles. 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 static or periodic 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 dwell on 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.

We will study 2 models : matter field with a density, individual particles. The main purpose is to show the computational methods, and introduce the currents.

7.1 MODEL WITH A MATTER FIELD

The action is : $\int_\Omega \left(C_I \mu \frac{1}{i} \langle \psi, \nabla_V \psi \rangle + C_G \langle \mathcal{F}_G, \mathcal{F}_G \rangle + C_A \langle \mathcal{F}_A, \mathcal{F}_A \rangle\right) \varpi_4$.

7.1.1 Equation for the Matter Field

The action is :

$\int_\Omega C_I \frac{1}{M} \frac{1}{i} \langle \psi, \nabla_V \psi \rangle \mu \varpi_4$

$[\nabla_\alpha \psi] = \vartheta(\sigma, 1) \left(\left[\gamma C \left(\nabla_\alpha^G \sigma\right)\right] [\psi_0] + [\psi_0] \left[\grave{A}_\alpha\right] \right)$

$\nabla_\alpha^G \sigma = \mathbf{Ad}_{\sigma^{-1}} \left(\partial_\alpha \sigma \cdot \sigma^{-1} + G_\alpha\right)$

We will not specify the chart on the Clifford Algebra : $\sigma(r, w)$ with $r, w : M \to \mathbb{R}^3$ represent

- either r, w with : $\sigma = \sigma_w \cdot \sigma_r = (a_w(m) + \upsilon(0, w(m))) \cdot (a_r(m) + \upsilon(r(m), 0))$

- or $r(m) = \operatorname{Re} Z(m), w(m) = \operatorname{Im} Z(m)$ with $\sigma = A + Z$

The tangent to the trajectory is deduced from :

$V = \frac{dq}{dt} = c\varepsilon_0 + \overrightarrow{v} = -\frac{c}{\langle \mathbf{Ad}_\sigma \varepsilon_0, \varepsilon_0 \rangle_{Cl}} \mathbf{Ad}_\sigma \varepsilon_0$

It is assumed that the density $\mu\left(\xi^0, \xi^1, \xi^2; \xi^3\right)$ meets the continuity equation : $\frac{d\mu}{dt} + \mu div V = 0$.

The equations are :

$\forall a = 1, 2, 3$:

$\frac{dL}{dr_a} = \frac{1}{\det P'} \sum_{\alpha=0}^3 \frac{d}{d\xi^\alpha} \left(\frac{\partial L}{\partial \partial_\alpha r_a} \det P' \right)$

$\frac{dL}{dw_a} = \frac{1}{\det P'} \sum_{\alpha=0}^3 \frac{d}{d\xi^\alpha} \left(\frac{\partial L}{\partial \partial_\alpha w_a} \det P' \right)$

Equation for r

i) First step : computation of the right hand side

$\frac{1}{\det P'} \sum_{\alpha=0}^3 \frac{d}{d\xi^\alpha} \left(\frac{\partial L}{\partial \partial_\alpha r_a} \det P' \right)$

$= \frac{1}{\det P'} C_I \frac{1}{i} \frac{1}{M} \sum_{\alpha=0}^3 \frac{d}{d\xi^\alpha} \left(\mu \det P' \sum_{\beta=0}^3 V^\beta \left(\frac{\partial}{\partial \partial_\alpha r_a} \langle \psi, \nabla_\beta \psi \rangle \right) \right)$

$\frac{\partial}{\partial \partial_\alpha r_a} \langle \psi, \nabla_\beta \psi \rangle$

$= \frac{\partial}{\partial \partial_\alpha r_a} \left\langle \psi_0, \left[\gamma C \left(\left(\sigma^{-1} \cdot \partial_\beta \sigma + \mathbf{Ad}_{\sigma^{-1}} G_\alpha\right)\right)\right] [\psi_0] \right\rangle$

$= \left\langle \psi_0, \left[\gamma C \left(\left(\sigma^{-1} \cdot \frac{\partial}{\partial \partial_\alpha r_a} (\partial_\beta \sigma)\right)\right)\right] [\psi_0] \right\rangle$

$\partial_\beta \sigma = \sum_{a=1}^3 \frac{\partial \sigma}{\partial w_a} \frac{\partial w_a}{\partial \xi^\beta} + \frac{\partial \sigma}{\partial r_a} \frac{\partial r_a}{\partial \xi^\beta} = \sum_{b=1}^3 \frac{\partial \sigma}{\partial w_b} \partial_\beta w_b + \frac{\partial \sigma}{\partial r_b} \partial_\beta r_b$

$\frac{\partial}{\partial \partial_\alpha r_a} (\partial_\beta \sigma) = \delta_\alpha^\beta \frac{\partial \sigma}{\partial r_a}$

$\frac{\partial}{\partial \partial_\alpha r_a} \langle \psi, \nabla_\beta \psi \rangle = \delta_\alpha^\beta \left\langle \psi_0, \left[\gamma C \left(\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a}\right)\right] [\psi_0] \right\rangle$

$\frac{\partial L}{\partial \partial_\alpha r_a} \det P' = C_I \frac{1}{i} \mu \frac{1}{M_p} V^\alpha \left\langle \psi_0, \left[\gamma C \left(\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a}\right)\right] [\psi_0] \right\rangle \det P'$

$\frac{1}{\det P'} \sum_{\alpha=0}^3 \frac{d}{d\xi^\alpha} \left(\frac{\partial L}{\partial \partial_\alpha r_a} \det P' \right)$

$= \frac{1}{\det P'} C_I \frac{1}{i} \frac{1}{M_p} \sum_{\alpha=0}^3 \frac{d}{d\xi^\alpha} \left(\mu \det P' V^\alpha \left\langle \psi_0, \gamma C \left(\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a}\right) \psi_0 \right\rangle \right)$

$= C_I \frac{1}{i} \frac{1}{M_p} \left\langle \psi_0, \gamma C \left(\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a}\right) \psi_0 \right\rangle \left(\frac{1}{\det P'} \sum_{\alpha=0}^3 \frac{d}{d\xi^\alpha} (\mu \det P' V^\alpha) \right)$

$+ \frac{1}{\det P'} C_I \frac{1}{i} \sum_{\alpha=0}^3 \mu \det P' V^\alpha \frac{d}{d\xi^\alpha} \left\langle \psi_0, \gamma C \left(\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a}\right) \psi_0 \right\rangle$

$\frac{1}{\det P'} \sum_{\alpha=0}^3 \frac{d}{d\xi^\alpha} (\mu \det P' V^\alpha)$

$= \frac{1}{\det P'} \sum_{\alpha=0}^3 \det P' V^\alpha \frac{d}{d\xi^\alpha} (\mu) + \frac{1}{\det P'} \sum_{\alpha=0}^3 \mu \frac{d}{d\xi^\alpha} (\det P' V^\alpha)$

$= \frac{d\mu}{dt} + \mu div V = 0$

$\frac{1}{\det P'} \sum_{\alpha=0}^3 \frac{d}{d\xi^\alpha} \left(\frac{\partial L}{\partial \partial_\alpha r_a} \det P' \right) = C_I \frac{1}{i} \frac{1}{M} \mu \left\langle \psi_0, \gamma C \left(\frac{d}{dt} \left(\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} \right) \right) \psi_0 \right\rangle$

ii) Computation of the left hand side and the derivative of V

$\frac{dL}{dr_a} = C_I \mu \frac{1}{i} \frac{1}{M} \frac{\partial}{\partial r_a} \sum_{\beta=0}^{3} V^\beta \left\langle \psi_0, \left[\gamma C \left(\nabla_\beta^G \sigma \right) \right] [\psi_0] + [\psi_0] \left[\grave{A}_\beta \right] \right\rangle$

$= C_I \mu \frac{1}{i} \frac{1}{M} \sum_{\beta=0}^{3} \left(\frac{\partial}{\partial r_a} V^\beta \right) \langle \psi_0, \nabla_\beta \psi \rangle + C_I \mu \frac{1}{i} \frac{1}{M} \sum_{\beta=0}^{3} V^\beta \left\langle \psi_0, \left[\gamma C \left(\frac{\partial}{\partial r_a} \nabla_\beta^G \sigma \right) \right] [\psi_0] \right\rangle$

Derivative of V :c

We have seen (Chapter 3 motion) the general formula :

$\alpha > 0 : \partial_\beta V^\alpha = \sum_{j=0}^{3} \left(P_j^\alpha - \frac{1}{c} P_j^0 V^\alpha \right) \left[\partial_\beta \sigma \cdot \sigma^{-1}, U \right]^j$

Here

$\beta > 0 : \frac{\partial V^\beta}{\partial r_a} = \sum_{j=0}^{3} \left(P_j^\beta - \frac{1}{c} P_j^0 V^\beta \right) \left[\frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1}, U \right]^j$

$C_I \mu \frac{1}{i} \frac{1}{M} \sum_{\beta=0}^{3} \left(\frac{\partial}{\partial r_a} V^\beta \right) \langle \psi_0, \nabla_\beta \psi \rangle$

$= C_I \mu \frac{1}{i} \frac{1}{M} \sum_{\beta=0}^{3} \sum_{j=0}^{3} P_j^\beta \left[\frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1}, U \right]^j \langle \psi, \nabla_\beta \psi \rangle - C_I \mu \frac{1}{i} \frac{1}{M} \sum_{j=0}^{3} \frac{1}{c} P_j^0 \left[\frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1}, U \right]^j \langle \psi, \nabla_V \psi \rangle$

$\frac{dL}{dr_a} = C_I \mu \frac{1}{i} \frac{1}{M} \{ \sum_{\beta=0}^{3} V^\beta \left\langle \psi_0, \left[\gamma C \left(\frac{\partial}{\partial r_a} \nabla_\beta^G \sigma \right) \right] [\psi_0] \right\rangle$

$+ \sum_{\beta,j=0}^{3} P_j^\beta \left[\frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1}, U \right]^j \langle \psi, \nabla_\beta \psi \rangle - \sum_{j=0}^{3} \frac{1}{c} P_j^0 \left[\frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1}, U \right]^j \langle \psi, \nabla_V \psi \rangle \}$

iii) Assembling the equations

We are left with the equations, for $a = 1, 2, 3$:

$\sum_{\beta=0}^{3} V^\beta \left\langle \psi_0, \left[\gamma C \left(\frac{\partial}{\partial r_a} \nabla_\beta^G \sigma \right) \right] [\psi_0] \right\rangle$

$+ \sum_{\beta=0}^{3} \sum_{j=0}^{3} P_j^\beta \left[\frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1}, U \right]^j \langle \psi, \nabla_\beta \psi \rangle - \sum_{j=0}^{3} \frac{1}{c} P_j^0 \left[\frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1}, U \right]^j \langle \psi, \nabla_V \psi \rangle$

$= \left\langle \psi_0, \gamma C \left(\frac{d}{dt} \left(\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} \right) \right) \psi_0 \right\rangle$

$\left\langle \psi_0, \left[\gamma C \left(\sum_{\beta=0}^{3} V^\beta \frac{\partial}{\partial r_a} \left(\nabla_\beta^G \sigma \right) - \frac{d}{dt} \left(\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} \right) \right) \right] [\psi_0] \right\rangle$

$= - \sum_{\beta=0}^{3} \sum_{j=0}^{3} P_j^\beta \left[\frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1}, U \right]^j \langle \psi, \nabla_\beta \psi \rangle + \sum_{j=0}^{3} \frac{1}{c} P_j^0 \left[\frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1}, U \right]^j \langle \psi, \nabla_V \psi \rangle$

iv) Computation of the derivatives of σ

$\frac{\partial}{\partial r_a} \left(\nabla_\beta^G \sigma \right) = \frac{\partial}{\partial r_a} \left(\sigma^{-1} \cdot \partial_\beta \sigma + \mathbf{Ad}_{\sigma^{-1}} G_\beta \right)$

$= -\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1} \cdot \partial_\beta \sigma - \sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1} \cdot G_\beta \cdot \sigma + \sigma^{-1} \cdot \frac{\partial}{\partial r_a} \partial_\beta \sigma + \sigma^{-1} \cdot G_\beta \cdot \frac{\partial \sigma}{\partial r_a}$

$\sum_{\beta=0}^{3} V^\beta \frac{\partial}{\partial r_a} \left(\nabla_\beta^G \sigma \right)$

$= -\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1} \cdot \frac{d\sigma}{dt} - \sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1} \cdot \widehat{G} \cdot \sigma + \sigma^{-1} \cdot \sum_{\beta=0}^{3} V^\beta \frac{\partial}{\partial r_a} \partial_\beta \sigma + \sigma^{-1} \cdot \widehat{G} \cdot \frac{\partial \sigma}{\partial r_a}$

$= -\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1} \cdot \frac{d\sigma}{dt} + \sigma^{-1} \cdot \sum_{\beta=0}^{3} V^\beta \frac{\partial}{\partial r_a} \partial_\beta \sigma + \mathbf{Ad}_{\sigma^{-1}} \left[\widehat{G}, \frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1} \right]$

$\frac{d}{dt} \left(\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} \right) = -\sigma^{-1} \cdot \frac{d\sigma}{dt} \cdot \sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} + \sigma^{-1} \cdot \frac{d}{dt} \frac{\partial \sigma}{\partial r_a}$

$\sum_{\beta=0}^{3} V^\beta \frac{\partial}{\partial r_a} \left(\nabla_\beta^G \sigma \right) - \frac{d}{dt} \left(\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} \right)$

$= \mathbf{Ad}_{\sigma^{-1}} \left[\frac{d\sigma}{dt} \cdot \sigma^{-1} + \widehat{G}, \frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1} \right] + \sigma^{-1} \cdot \left(\sum_{\beta=0}^{3} V^\beta \frac{\partial}{\partial r_a} \partial_\beta \sigma - \frac{d}{dt} \frac{\partial \sigma}{\partial r_a} \right)$

$\sum_{\beta=0}^{3} V^\beta \frac{\partial}{\partial r_a} \partial_\beta \sigma - \frac{d}{dt} \frac{\partial \sigma}{\partial r_a} = \sum_{\beta=0}^{3} V^\beta \left(\frac{\partial}{\partial r_a} \partial_\beta \sigma - \partial_\beta \frac{\partial \sigma}{\partial r_a} \right)$

$\frac{\partial}{\partial r_a} \partial_\beta \sigma - \partial_\beta \frac{\partial \sigma}{\partial r_a} = \frac{\partial}{\partial r_a} \left(\sum_{b=1}^{3} \frac{\partial \sigma}{\partial w_b} \partial_\beta w_b + \frac{\partial \sigma}{\partial r_b} \partial_\beta r_b \right) - \sum_{b=1}^{3} \left(\frac{\partial}{\partial r_b} \left(\frac{\partial \sigma}{\partial r_a} \right) \partial_\beta r_b + \frac{\partial}{\partial w_b} \left(\frac{\partial \sigma}{\partial r_a} \right) \partial_\beta w_b \right)$

$= \sum_{b=1}^{3} \frac{\partial}{\partial r_a} \frac{\partial \sigma}{\partial w_b} \partial_\beta w_b + \frac{\partial}{\partial r_a} \frac{\partial \sigma}{\partial r_b} \partial_\beta r_b - \frac{\partial}{\partial r_b} \left(\frac{\partial \sigma}{\partial r_a} \right) \partial_\beta r_b - \frac{\partial}{\partial w_b} \left(\frac{\partial \sigma}{\partial r_a} \right) \partial_\beta w_b = 0$

$\sum_{\beta=0}^{3} V^\beta \frac{\partial}{\partial r_a} \left(\nabla_\beta^G \sigma \right) - \frac{d}{dt} \left(\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} \right) = \mathbf{Ad}_{\sigma^{-1}} \left[\frac{d\sigma}{dt} \cdot \sigma^{-1} + \widehat{G}, \frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1} \right] = \left[\nabla_V^G \sigma, \sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} \right]$

v) Result

$\left\langle \psi_0, \left[\gamma C \left(\left[\nabla_V^G \sigma, \sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} \right] \right) \right] [\psi_0] \right\rangle$

$= - \sum_{\beta=0}^{3} \sum_{j=0}^{3} P_j^\beta \left[\frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1}, U \right]^j \langle \psi, \nabla_\beta \psi \rangle + \sum_{j=0}^{3} \frac{1}{c} P_j^0 \left[\frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1}, U \right]^j \langle \psi, \nabla_V \psi \rangle$

The equation reads :

$$a=1,2,3:$$
$$\left\langle \psi_0,\gamma C\left(\left[\sigma^{-1}\cdot\frac{\partial\sigma}{\partial r_a},\nabla_V^G\sigma\right]\right)\psi_0\right\rangle=\sum_{j=0}^3\frac{1}{c}P_j^0\left[\frac{\partial\sigma}{\partial r_a}\cdot\sigma^{-1},U\right]^j\langle\psi,\nabla_\beta\psi\rangle$$
$$-\sum_{\beta,j=0}^3 P_j^\beta\left[\frac{\partial\sigma}{\partial r_a}\cdot\sigma^{-1},U\right]^j\langle\psi,\nabla_V\psi\rangle \tag{7.1}$$

Equation for w

The computation is identical.

$$a=1,2,3:$$
$$\left\langle \psi_0,\gamma C\left(\left[\sigma^{-1}\cdot\frac{\partial\sigma}{\partial w_a},\nabla_V^G\sigma\right]\right)\psi_0\right\rangle=\sum_{j=0}^3\frac{1}{c}P_j^0\left[\frac{\partial\sigma}{\partial w_a}\cdot\sigma^{-1},U\right]^j\langle\psi,\nabla_\beta\psi\rangle$$
$$-\sum_{\beta,j=0}^3 P_j^\beta\left[\frac{\partial\sigma}{\partial w_a}\cdot\sigma^{-1},U\right]^j\langle\psi,\nabla_V\psi\rangle \tag{7.2}$$

7.1.2 Equations for the gravitational field

The equations are, with the full lagrangian.

$\forall a=1..6,\alpha=0...3:$

$\frac{d\left(L\det P'\right)}{dG_\alpha^a}=\sum_\beta\frac{d}{d\xi^\beta}\frac{d\left(L\det P'\right)}{d\partial_\beta G_\alpha^a}$

Derivatives

i) $\frac{dL}{dG_\alpha^a}=C_I\mu\frac{1}{i}\frac{1}{M}\frac{\partial}{\partial G_\alpha^a}\langle\psi,\nabla_V\psi\rangle+C_G\frac{\partial}{\partial G_\alpha^a}\sum_{\lambda\mu}\sum_{b=1}^3\mathcal{F}_{r\lambda\mu}^b\mathcal{F}_r^{b\lambda\mu}-\mathcal{F}_{w\lambda\mu}^b\mathcal{F}_w^{b\lambda\mu}$

ii) Computation of the derivative for the particles

$\frac{\partial}{\partial G_\alpha^a}\langle\psi,\nabla_V\psi\rangle=V^\alpha\frac{\partial}{\partial G_\alpha^a}\langle\psi,\nabla_\alpha\psi\rangle=V^\alpha\left\langle\psi,\frac{\partial}{\partial G_\alpha^a}\nabla_\alpha\psi\right\rangle=V^\alpha\langle\psi,[\gamma C(\overrightarrow{\kappa}_a)][\psi]\rangle$

$=V^\alpha\langle\vartheta(\sigma,1)\psi_0,[\gamma C(\overrightarrow{\kappa}_a)]\vartheta(\sigma,1)\psi_0\rangle$

$=V^\alpha\langle\psi_0,\gamma C(\mathbf{Ad}_{\sigma^{-1}}\overrightarrow{\kappa}_a)\psi_0\rangle$

Let be $Z_a=\mathbf{Ad}_{\sigma^{-1}}\overrightarrow{\kappa}_a=\sum_{b=1}^6\left[Ad\left(\sigma^{-1}\right)\right]_a^b\overrightarrow{\kappa}_b$

$\langle\psi_0,\gamma C(\mathbf{Ad}_{\sigma^{-1}}\overrightarrow{\kappa}_a)\psi_0\rangle=\langle\psi_0,\gamma C(Z_a)\psi_0\rangle=-i\frac{M^2}{2}k_0^t\operatorname{Re}Z_a=-i\frac{M^2}{2}\operatorname{Re}k_0^tZ_a$

$k_0^tZ_a=\sum_{b=1}^3\left[Ad\left(\sigma^{-1}\right)\right]_a^b[k_0]^b=\sum_{b=1}^6\left(\left[Ad\left(\sigma^{-1}\right)\right]^t\right)_b^a[k_0]^b$

$=\sum_{b=1}^6[Ad(\sigma)]_b^a[k_0]^b=\sum_{b=1}^6([Ad(\sigma)][k_0])^a=[\mathbf{Ad}_\sigma\upsilon(k_0,0)]^a$

$\frac{\partial}{\partial G_\alpha^a}\langle\psi,\nabla_V\psi\rangle=-i\frac{M^2}{2}\operatorname{Re}[\mathbf{Ad}_\sigma\upsilon(k_0,0)]^a$

$\frac{\partial}{\partial G_\alpha^a}C_I\mu\frac{1}{i}\frac{1}{M}\langle\psi,\nabla_V\psi\rangle=-C_I\mu\frac{M}{2}V^\alpha[\mathbf{Ad}_\sigma\upsilon(k_0,0)]^a$

iii) Computation of the derivative for the potential

$\frac{\partial}{\partial G_\alpha^a}\left(\sum_{\lambda\mu}\sum_{b=1}^3\mathcal{F}_{r\lambda\mu}^b\mathcal{F}_r^{b\lambda\mu}-\mathcal{F}_{w\lambda\mu}^b\mathcal{F}_w^{b\lambda\mu}\right)$

$=\frac{\partial}{\partial G_\alpha^a}\left(\sum_{b=1}^3\sum_{pr\lambda\mu}\mathcal{F}_{r\lambda\mu}^b g^{p\lambda}g^{q\mu}\mathcal{F}_{rpq}^b-\mathcal{F}_{w\lambda\mu}^b g^{p\lambda}g^{q\mu}\mathcal{F}_{wpq}^b\right)$

$=2\sum_{b=1}^3\sum_{\lambda\mu}\left(\frac{\partial}{\partial G_\alpha^a}\mathcal{F}_{r\lambda\mu}^b\right)\mathcal{F}_r^{b\lambda\mu}-\left(\frac{\partial}{\partial G_\alpha^a}\mathcal{F}_{w\lambda\mu}^b\right)\mathcal{F}_w^{b\lambda\mu}$

$\frac{\partial}{\partial G_\alpha^a}\mathcal{F}_{r\lambda\mu}^b=2\frac{\partial}{\partial G_\alpha^a}[j(G_{r\lambda})G_{r\mu}-j(G_{w\lambda})G_{w\mu}]^b=2\frac{\partial}{\partial G_\alpha^a}\sum_{p,q=1}^3\epsilon(b,p,q)\left[G_{r\lambda}^pG_{r\mu}^q-G_{w\lambda}^pG_{w\mu}^q\right]$

$\frac{\partial}{\partial G_\alpha^a}\mathcal{F}_{w\lambda\mu}^b=2\frac{\partial}{\partial G_\alpha^a}[j(G_{w\lambda})G_{r\mu}+j(G_{r\lambda})G_{w\mu}]^b=2\frac{\partial}{\partial G_\alpha^a}\sum_{p,q=1}^3\epsilon(b,p,q)\left[G_{w\lambda}^pG_{r\mu}^q+G_{r\lambda}^pG_{w\mu}^q\right]$

$a=1,2,3:$

$\frac{\partial}{\partial G_{r\alpha}^a}\mathcal{F}_{r\lambda\mu}^b=2\frac{\partial}{\partial G_{r\alpha}^a}\sum_{p,q=1}^3\epsilon(b,p,q)\left[G_{r\lambda}^pG_{r\mu}^q-G_{w\lambda}^pG_{w\mu}^q\right]$

$=2\sum_{c=1}^3\epsilon(b,a,c)\left(\delta_\alpha^\lambda G_{r\mu}^c-\delta_\alpha^\mu G_{r\lambda}^c\right)$

$\frac{\partial}{\partial G_\alpha^a}\mathcal{F}_{w\lambda\mu}^b=2\frac{\partial}{\partial G_{r\alpha}^a}\sum_{p,q=1}^3\epsilon(b,p,q)\left[G_{w\lambda}^pG_{r\mu}^q+G_{r\lambda}^pG_{w\mu}^q\right]$

$= 2\sum_{c=1}^{3}\epsilon(b,a,c)\left(-\delta_{\alpha}^{\mu}G_{w\lambda}^{c}+\delta_{\alpha}^{\lambda}G_{w\mu}^{c}\right)$
$\frac{\partial}{\partial G_{\alpha}^{a}}\left(\sum_{\lambda\mu}\sum_{b=1}^{3}\mathcal{F}_{r\lambda\mu}^{b}\mathcal{F}_{r}^{b\lambda\mu}-\mathcal{F}_{w\lambda\mu}^{b}\mathcal{F}_{w}^{b\lambda\mu}\right)$
$= -8\sum_{b,c=1}^{3}\sum_{\lambda}\epsilon(a,b,c)\left(\mathcal{F}_{r}^{b\alpha\lambda}G_{r\lambda}^{c}-\mathcal{F}_{w}^{b\alpha\lambda}G_{w\lambda}^{c}\right)$
$= -8\sum_{\lambda}\left(j\left(\mathcal{F}_{r}^{\alpha\lambda}\right)G_{r\lambda}-j\left(\mathcal{F}_{w}^{\alpha\lambda}\right)G_{w\lambda}\right)^{a}$
$\frac{\partial}{\partial G_{r\alpha}^{a}}\mathcal{F}_{r\lambda\mu}^{b}+i\frac{\partial}{\partial G_{\alpha}^{a}}\mathcal{F}_{w\lambda\mu}^{b}=2\sum_{c=1}^{3}\epsilon(b,a,c)\left(\delta_{\alpha}^{\lambda}G_{r\mu}^{c}-\delta_{\alpha}^{\mu}G_{r\lambda}^{c}\right)+i2\sum_{c=1}^{3}\epsilon(b,a,c)\left(-\delta_{\alpha}^{\mu}G_{w\lambda}^{c}+\delta_{\alpha}^{\lambda}G_{w\mu}^{c}\right)$
$\frac{\partial\mathcal{F}_{\lambda\mu}^{b}}{\partial G_{r\alpha}^{a}}=2\sum_{c=1}^{3}\epsilon(b,a,c)\left(\delta_{\alpha}^{\lambda}G_{\mu}^{c}-\delta_{\alpha}^{\mu}G_{\lambda}^{c}\right)$
$a=4,5,6:$
$\frac{\partial}{\partial G_{w\alpha}^{a}}\mathcal{F}_{r\lambda\mu}^{b}=2\frac{\partial}{\partial G_{w\alpha}^{a}}\sum_{p,q=1}^{3}\epsilon(b,p,q)\left[G_{r\lambda}^{p}G_{r\mu}^{q}-G_{w\lambda}^{p}G_{w\mu}^{q}\right]$
$= -2\sum_{c=1}^{3}\epsilon(b,a,c)\left(\delta_{\alpha}^{\lambda}G_{w\mu}^{c}-\delta_{\alpha}^{\mu}G_{w\lambda}^{c}\right)$
$\frac{\partial}{\partial G_{\alpha}^{a}}\mathcal{F}_{w\lambda\mu}^{b}=2\frac{\partial}{\partial G_{w\alpha}^{a}}\sum_{p,q=1}^{3}\epsilon(b,p,q)\left[G_{w\lambda}^{p}G_{r\mu}^{q}+G_{r\lambda}^{p}G_{w\mu}^{q}\right]$
$= 2\sum_{c=1}^{3}\epsilon(b,a,c)\left(\delta_{\alpha}^{\lambda}G_{r\mu}^{c}-G_{r\lambda}^{c}\delta_{\alpha}^{\mu}\right)$
$\frac{\partial}{\partial G_{\alpha}^{a}}\mathcal{F}_{r\lambda\mu}^{b}+i\frac{\partial}{\partial G_{\alpha}^{a}}\mathcal{F}_{w\lambda\mu}^{b}=-2\sum_{c=1}^{3}\epsilon(b,a,c)\left(\delta_{\alpha}^{\lambda}G_{w\mu}^{c}-\delta_{\alpha}^{\mu}G_{w\lambda}^{c}\right)+i2\sum_{c=1}^{3}\epsilon(b,a,c)\left(\delta_{\alpha}^{\lambda}G_{r\mu}^{c}-G_{r\lambda}^{c}\delta_{\alpha}^{\mu}\right)$
$\frac{\partial\mathcal{F}_{\lambda\mu}^{b}}{\partial G_{r\alpha}^{a}}=2\sum_{c=1}^{3}\epsilon(b,a,c)\left(-\delta_{\alpha}^{\lambda}G_{w\mu}^{c}+i\delta_{\alpha}^{\lambda}G_{r\mu}^{c}+\delta_{\alpha}^{\mu}G_{w\lambda}^{c}-iG_{r\lambda}^{c}\delta_{\alpha}^{\mu}\right)$
$= 2i\sum_{c=1}^{3}\epsilon(b,a,c)\left(\delta_{\alpha}^{\lambda}G_{\mu}^{c}-G_{\lambda}^{c}\delta_{\alpha}^{\mu}\right)$
$\frac{\partial}{\partial G_{\alpha}^{a}}\left(\sum_{\lambda\mu}\sum_{b=1}^{3}\mathcal{F}_{r\lambda\mu}^{b}\mathcal{F}_{r}^{b\lambda\mu}-\mathcal{F}_{w\lambda\mu}^{b}\mathcal{F}_{w}^{b\lambda\mu}\right)$
$= 8\sum_{b,c=1}^{3}\sum_{\lambda}\epsilon(a,b,c)\left(\mathcal{F}_{r}^{b\alpha\lambda}G_{w\lambda}^{c}+\mathcal{F}_{w}^{b\alpha\lambda}G_{r\lambda}^{c}\right)$
$= 8\sum_{\lambda}\left(j\left(\mathcal{F}_{r}^{\alpha\lambda}\right)G_{w\lambda}+j\left(\mathcal{F}_{w}^{\alpha\lambda}\right)G_{r\lambda}\right)^{a}$
$\frac{\partial}{\partial G_{w\alpha}^{a}}\left(\sum_{\lambda\mu}\sum_{b=1}^{3}\left(\mathcal{F}_{r\lambda\mu}^{b}\mathcal{F}_{r}^{b\lambda\mu}-\mathcal{F}_{w\lambda\mu}^{b}\mathcal{F}_{w}^{b\lambda\mu}\right)\right)=8\sum_{\lambda=0}^{3}\left[j\left(\mathcal{F}_{r}^{\alpha\lambda}\right)G_{w\lambda}+j\left(\mathcal{F}_{w}^{\alpha\lambda}\right)G_{r\lambda}\right]^{a}$
$\frac{\partial L_{G}}{\partial G_{r\alpha}^{a}}=-8\sum_{\lambda}\left(j\left(\mathcal{F}_{r}^{\alpha\lambda}\right)G_{r\lambda}-j\left(\mathcal{F}_{w}^{\alpha\lambda}\right)G_{w\lambda}\right)^{a}$
$\frac{\partial L_{G}}{\partial G_{w\alpha}^{a}}=8\sum_{\lambda=0}^{3}\left[j\left(\mathcal{F}_{r}^{\alpha\lambda}\right)G_{w\lambda}+j\left(\mathcal{F}_{w}^{\alpha\lambda}\right)G_{r\lambda}\right]^{a}$
$\frac{\partial L\det P'}{\partial z^{i}}=\frac{\partial L\det P'}{\partial\operatorname{Re}z^{i}}+\frac{1}{i}\frac{\partial L\det P'}{\partial\operatorname{Im}z^{i}}$
iv) Computation of the right hand side
$\frac{dL}{d\partial_{\beta}G_{\alpha}^{a}}=C_{G}\frac{\partial}{\partial\partial_{\beta}G_{\alpha}^{a}}\left(\sum_{b=1}^{3}\sum_{pr\lambda\mu}\mathcal{F}_{r\lambda\mu}^{b}g^{p\lambda}g^{q\mu}\mathcal{F}_{rpq}^{b}-\mathcal{F}_{w\lambda\mu}^{b}g^{p\lambda}g^{q\mu}\mathcal{F}_{wpq}^{b}\right)$
$= 2C_{G}\sum_{b=1}^{3}\sum_{\lambda\mu}\left(\frac{\partial}{\partial\partial_{\beta}G_{\alpha}^{a}}\mathcal{F}_{r\lambda\mu}^{b}\right)\mathcal{F}_{r}^{b\lambda\mu}-\left(\frac{\partial}{\partial\partial_{\beta}G_{\alpha}^{a}}\mathcal{F}_{w\lambda\mu}^{b}\right)\mathcal{F}_{w}^{b\lambda\mu}$
$a=1,2,3:$
$\frac{\partial}{\partial\partial_{\beta}G_{\alpha}^{a}}\mathcal{F}_{w\lambda\mu}^{b}=0$
$\frac{\partial L}{\partial\partial_{\beta}G_{r\alpha}^{a}}=-4C_{G}\mathcal{F}_{r}^{a\alpha\beta}$
$a=4,5,6:$
$\frac{\partial}{\partial\partial_{\beta}G_{w\alpha}^{a}}\mathcal{F}_{r\lambda\mu}^{b}=0$
$\frac{dL}{d\partial_{\beta}G_{w\alpha}^{a}}=4C_{G}\mathcal{F}_{w}^{a\alpha\beta}$

Equations :

$\forall\alpha=0,...3,$
$\forall a=1,2,3$
$-C_{I}\mu\frac{M}{2}V^{\alpha}\operatorname{Re}\left[\mathbf{Ad}_{\sigma}\upsilon(k_{0},0)\right]^{a}-8C_{G}\sum_{\lambda=0}^{3}\left(j\left(\mathcal{F}_{r}^{\alpha\lambda}\right)G_{r\lambda}-j\left(\mathcal{F}_{w}^{\alpha\lambda}\right)G_{w\lambda}\right)^{a}$
$= -4C_{G}\frac{1}{\det P'}\sum_{\beta}\frac{d}{d\xi^{\beta}}\left(\mathcal{F}_{r}^{a\alpha\beta}\det P'\right)$
$\frac{C_{I}}{8C_{G}}\mu\frac{M}{2}V^{\alpha}\operatorname{Re}\left[\mathbf{Ad}_{\sigma}\upsilon(k_{0},0)\right]^{a}+\sum_{\lambda=0}^{3}\left(j\left(\mathcal{F}_{r}^{\alpha\lambda}\right)G_{r\lambda}-j\left(\mathcal{F}_{w}^{\alpha\lambda}\right)G_{w\lambda}\right)^{a}$
$= \frac{1}{2}\frac{1}{\det P'}\sum_{\beta}\frac{d}{d\xi^{\beta}}\left(\mathcal{F}_{r}^{a\alpha\beta}\det P'\right)$
$\forall a=4,5,6$
$C_{I}\mu\frac{M}{2}V^{\alpha}\operatorname{Im}\left[\mathbf{Ad}_{\sigma}\upsilon(k_{0},0)\right]^{a}+8C_{G}\sum_{\lambda=0}^{3}\left(j\left(\mathcal{F}_{r}^{\alpha\lambda}\right)G_{w\lambda}+j\left(\mathcal{F}_{w}^{\alpha\lambda}\right)G_{r\lambda}\right)^{a}$
$= 4C_{G}\frac{1}{\det P'}\sum_{\beta}\frac{d}{d\xi^{\beta}}\left(\mathcal{F}_{w}^{a\alpha\beta}\det P'\right)$

$\frac{C_I}{8C_G}\mu\frac{M}{2}V^\alpha \operatorname{Im}\left[\mathbf{Ad}_\sigma \upsilon\left(k_0,0\right)\right]^a+\sum_{\lambda=0}^3\left(j\left(\mathcal{F}_r^{\alpha\lambda}\right)G_{w\lambda}+j\left(\mathcal{F}_w^{\alpha\lambda}\right)G_{r\lambda}\right)^a$
$=\frac{1}{2}\frac{1}{\det P'}\sum_\beta\frac{d}{d\xi^\beta}\left(\mathcal{F}_w^{a\alpha\beta}\det P'\right)$
Multiply each equation by $\overrightarrow{\kappa}_a$ and sum.
$\sum_{a=1}^3\left(j\left(\mathcal{F}_r^{\alpha\lambda}\right)G_{r\lambda}-j\left(\mathcal{F}_w^{\alpha\lambda}\right)G_{w\lambda}\right)^a\overrightarrow{\kappa}_a+\left(j\left(\mathcal{F}_r^{\alpha\lambda}\right)G_{w\lambda}+j\left(\mathcal{F}_w^{\alpha\lambda}\right)G_{r\lambda}\right)^a\overrightarrow{\kappa}_{a+3}$
$=\upsilon\left(j\left(\mathcal{F}_r^{\alpha\lambda}\right)G_{r\lambda}-j\left(\mathcal{F}_w^{\alpha\lambda}\right)G_{w\lambda},j\left(\mathcal{F}_w^{\alpha\lambda}\right)G_{r\lambda}+j\left(\mathcal{F}_r^{\alpha\lambda}\right)G_{w\lambda}\right)=\left[\upsilon\left(\mathcal{F}_r^{\alpha\lambda},\mathcal{F}_w^{\alpha\lambda}\right),\upsilon\left(G_{r\lambda},G_{w\lambda}\right)\right]$
$\sum_{a=1}^3\mathcal{F}_r^{a\alpha\beta}\overrightarrow{\kappa}_a+\mathcal{F}_w^{a\alpha\beta}\overrightarrow{\kappa}_{a+3}=\mathcal{F}_G^{\alpha\beta}$
$\frac{C_I}{8C_G}\mu\frac{M}{2}\sum_{a=1}^3\left(\operatorname{Re}\left[\mathbf{Ad}_\sigma\upsilon\left(k_0,0\right)\right]^a\overrightarrow{\kappa}_a+\operatorname{Im}\left[\mathbf{Ad}_\sigma\upsilon\left(k_0,0\right)\right]^a\overrightarrow{\kappa}_{a+3}\right)$
$=\frac{C_I}{8C_G}\mu\epsilon\frac{M}{2}\sum_{a=1}^3\left[\mathbf{Ad}_\sigma\upsilon\left(k_0,0\right)\right]^a\overrightarrow{\kappa}_a=\mathbf{Ad}_\sigma\upsilon\left(k_0,0\right)$
The equations read :
$\frac{C_I}{16C_G}\mu MV^\alpha\mathbf{Ad}_\sigma\upsilon\left(k_0,0\right)+\sum_{\beta=0}^3\left[\upsilon\left(\mathcal{F}_r^{\alpha\beta},\mathcal{F}_w^{\alpha\beta}\right),\upsilon\left(G_{r\beta},G_{w\beta}\right)\right]$
$=\frac{1}{2}\frac{1}{\det P'}\sum_{\beta=0}^3\frac{d}{d\xi^\beta}\left(\mathcal{F}^{\alpha\beta}\det P'\right)$
That is :

$$\begin{aligned}\forall\alpha=0,...3:\phi_G^\alpha-J_G^\alpha&=\frac{1}{2}\frac{1}{\det P'}\sum_\beta\frac{d}{d\xi^\beta}\left(\mathcal{F}_G^{\alpha\beta}\det P'\right)\\ J_G^\alpha&=-\frac{C_I}{16C_G}\mu MV^\alpha\mathbf{Ad}_\sigma\upsilon\left(k_0,0\right)\\ \phi_G^\alpha&=\sum_{\beta=0}^3\left[\mathcal{F}^{\alpha\beta},G_\beta\right]_{T_1Spin(3,1)}\end{aligned}\qquad(7.3)$$

7.1.3 Equation for the other fields

Derivatives :

$\frac{dL}{d\grave{A}_\alpha^a}=C_I\mu\frac{1}{i}\frac{1}{M}\left\langle\psi,\frac{\partial}{\partial\grave{A}_\alpha^a}\nabla_V\psi\right\rangle+C_A\frac{\partial}{\partial\grave{A}_\alpha^a}\sum_{\lambda\mu}\sum_{b=1}^m\left(\mathcal{F}_{A\lambda\mu}^b\mathcal{F}_A^{a\lambda\mu}\right)$

$\left\langle\psi,\frac{\partial}{\partial\grave{A}_\alpha^a}\nabla_V\psi\right\rangle=V^\alpha\left\langle\vartheta\left(\sigma,1\right)\psi_0,\frac{\partial}{\partial\grave{A}_\alpha^a}\vartheta\left(\sigma,1\right)\left(\left[\psi_0\right]\left[Ad_\varkappa\left(\grave{A}_\alpha\right)\right]\right)\right\rangle$
$=V^\alpha\left\langle\psi_0,\left[\psi_0\right]\left[\frac{\partial}{\partial\grave{A}_\alpha^a}\left(\grave{A}_\alpha\right)\right]\right\rangle=V^\alpha\left\langle\psi_0,\left[\psi_0\right]\left[\theta_a\right]\right\rangle$
$\frac{\partial}{\partial\grave{A}_\alpha^a}\sum_{\lambda\mu}\sum_{b=1}^m\left(\mathcal{F}_{A\lambda\mu}^b\mathcal{F}_A^{a\lambda\mu}\right)=\frac{\partial}{\partial\grave{A}_\alpha^a}\left(\sum_{b=1}^m\sum_{pr\lambda\mu}\mathcal{F}_{A\lambda\mu}^bg^{p\lambda}g^{q\mu}\mathcal{F}_{Apq}^b\right)$
$=\sum_{b=1}^m\sum_{pr\lambda\mu}\left(\frac{\partial}{\partial\grave{A}_\alpha^a}\mathcal{F}_{A\lambda\mu}^b\right)g^{p\lambda}g^{q\mu}\mathcal{F}_{Apq}^b+\mathcal{F}_{A\lambda\mu}^bg^{p\lambda}g^{q\mu}\left(\frac{\partial}{\partial\grave{A}_\alpha^a}\mathcal{F}_{Apq}^b\right)$
$=\sum_{b=1}^m\sum_{pr\lambda\mu}\left(\frac{\partial}{\partial\grave{A}_\alpha^a}\mathcal{F}_{A\lambda\mu}^b\right)\mathcal{F}_A^{b\lambda\mu}+\left(\frac{\partial}{\partial\grave{A}_\alpha^a}\mathcal{F}_{Apq}^b\right)\mathcal{F}_A^{bpq}$
$=2\sum_{b=1}^m\sum_{\lambda\mu}\left(\frac{\partial}{\partial\grave{A}_\alpha^a}\mathcal{F}_{A\lambda\mu}^b\right)\mathcal{F}_A^{b\lambda\mu}$
$\frac{\partial}{\partial\grave{A}_\alpha^a}\mathcal{F}_{A\lambda\mu}^b=2\frac{\partial}{\partial\grave{A}_\alpha^a}\left[\grave{A}_\lambda,\grave{A}_\mu\right]^b$
$=2\frac{\partial}{\partial\grave{A}_\alpha^a}\left[\sum_{c=1}^m\grave{A}_\lambda^c\overrightarrow{\theta}_c,\sum_{d=1}^m\grave{A}_\mu^d\overrightarrow{\theta}_d\right]^b$
$=2\left(\left[\delta_\alpha^\lambda\overrightarrow{\theta}_a,\sum_{d=1}^m\grave{A}_\mu^d\overrightarrow{\theta}_d\right]^b+\left[\sum_{c=1}^m\grave{A}_\lambda^c\overrightarrow{\theta}_c,\delta_\alpha^\mu\overrightarrow{\theta}_a\right]^b\right)$
$=2\left(\delta_\alpha^\lambda\left[\overrightarrow{\theta}_a,\grave{A}_\mu\right]^b+\delta_\alpha^\mu\left[\grave{A}_\lambda,\overrightarrow{\theta}_a\right]^b\right)$
$\frac{\partial}{\partial\grave{A}_\alpha^a}\sum_{\lambda\mu}\sum_{b=1}^m\left(\mathcal{F}_{A\lambda\mu}^b\mathcal{F}_A^{a\lambda\mu}\right)$
$=4\sum_{b=1}^m\sum_{\lambda\mu}\left(\delta_\alpha^\lambda\left[\overrightarrow{\theta}_a,\grave{A}_\mu\right]^b\mathcal{F}_A^{b\lambda\mu}+\delta_\alpha^\mu\left[\grave{A}_\lambda,\overrightarrow{\theta}_a\right]^b\mathcal{F}_A^{b\lambda\mu}\right)$
$=4\sum_{b=1}^m\sum_{\lambda\mu}\left(\left[\overrightarrow{\theta}_a,\grave{A}_\mu\right]^b\mathcal{F}_A^{b\alpha\mu}+\left[\grave{A}_\lambda,\overrightarrow{\theta}_a\right]^b\mathcal{F}_A^{b\lambda\alpha}\right)$
$=8\sum_{b=1}^m\sum_{\lambda=0}^3\left[\overrightarrow{\theta}_a,\grave{A}_\lambda\right]^b\mathcal{F}_A^{b\alpha\lambda}$
$=8\sum_{\lambda=0}^3\left\langle\mathcal{F}_A^{\alpha\lambda},\left[\overrightarrow{\theta}_a,\grave{A}_\lambda\right]\right\rangle_{T_1U}$

$= 8\sum_{\lambda=0}^{3}\left\langle \overrightarrow{\theta}_a, \left[\grave{A}_\lambda, \mathcal{F}_A^{\alpha\lambda}\right]\right\rangle_{T_1U}$

$\frac{\partial}{\partial \grave{A}_\alpha^a}\sum_{\lambda\mu}\sum_{b=1}^{m}\left(\mathcal{F}_{A\lambda\mu}^b \mathcal{F}_A^{a\lambda\mu}\right) = 8\sum_{\lambda=0}^{3}\left[\grave{A}_\lambda, \mathcal{F}_A^{\alpha\lambda}\right]^a$

Using : $\forall X, Y, Z \in T_1U : \langle X, [Y,Z]\rangle = \langle [X,Y], Z\rangle$ and the fact that the basis is orthonormal.

ii) $\frac{dL}{d\partial_\beta \grave{A}_\alpha^a} = C_A \frac{\partial}{\partial\partial_\beta \grave{A}_\alpha^a}\sum_{b=1}^{m}\sum_{\lambda\mu}\mathcal{F}_{A\lambda\mu}^b \mathcal{F}_A^{b\lambda\mu}$

$= C_A \frac{\partial}{\partial\partial_\beta \grave{A}_\alpha^a}\left(\sum_{b=1}^{m}\sum_{pr\lambda\mu}\mathcal{F}_{A\lambda\mu}^b g^{p\lambda} g^{q\mu}\mathcal{F}_{Apq}^b\right)$

$= 2C_A\sum_{b=1}^{m}\sum_{\lambda\mu}\left(\frac{\partial}{\partial\partial_\beta \grave{A}_\alpha^a}\mathcal{F}_{A\lambda\mu}^b\right)\mathcal{F}_A^{b\lambda\mu}$

$\frac{d\mathcal{F}_{A\lambda\mu}^b}{d\partial_\beta \grave{A}_\alpha^a} = \left(\delta_\beta^\lambda\delta_\alpha^\mu - \delta_\alpha^\lambda\delta_\beta^\mu\right)\delta_a^b$

$\frac{dL}{d\partial_\beta \grave{A}_\alpha^a} = -4C_A\mathcal{F}_A^{a\alpha\beta}$

Equation

The equation reads in the vacuum :

$\forall\alpha = 0...3, \forall a = 1,...m$

$C_I\mu\frac{1}{i}\frac{1}{M}V^\alpha\langle\psi_0, [\psi_0][\theta_a]\rangle + 8C_A\sum_{\lambda=0}^{3}\left[\grave{A}_\lambda, \mathcal{F}_A^{\alpha\lambda}\right]^a = -4C_A\frac{1}{\det P'}\sum_\beta\frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{a\alpha\beta}\det P'\right)$

$-\frac{C_I}{8C_A}\mu\frac{1}{i}\frac{1}{M}V^\alpha\langle\psi_0, [\psi_0][\theta_a]\rangle + \sum_{\beta=0}^{3}\left[\mathcal{F}_A^{\alpha\beta}, \grave{A}_\beta\right]^a = \frac{1}{2}\frac{1}{\det P'}\sum_{\beta=0}^{3}\frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{a\alpha\beta}\det P'\right)$

That is :

$$\begin{gathered}\forall\alpha = 0...3 : \phi_A^\alpha - J_A^\alpha = \frac{1}{2}\frac{1}{\det P'}\sum_\beta\frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{a\alpha\beta}\det P'\right)\\ \phi_A^\alpha = \sum_{\beta=0}^{3}\left[\mathcal{F}_A^{\alpha\beta}, \grave{A}_\beta\right]_{T_1U}\\ J_A^\alpha = \frac{C_I}{8C_A}\mu V^\alpha\frac{1}{M}\sum_{a=1}^{m}\frac{1}{i}\left\langle\psi_0, [\psi_0]\left[\overrightarrow{\theta_a}\right]\right\rangle\overrightarrow{\theta}_a\end{gathered} \tag{7.4}$$

7.1.4 Equation for the tetrad

We have seen previously the equations :

$\forall\alpha,\beta = 0...3 : \sum_i\frac{dL}{dP_i^\alpha}P_i^\beta - L\delta_\beta^\alpha = 0$

Derivatives

For the part related to the fields :

$\frac{dL_1}{dP_i^\alpha} = \sum_{\rho\theta\lambda\mu}\frac{\partial}{\partial P_i^\alpha}\left(g^{\lambda\rho}g^{\mu\theta}\right)\left(C_G\sum_{a=1}^{3}\left(\mathcal{F}_{r\lambda\mu}^a\mathcal{F}_{r\rho\theta}^a - \mathcal{F}_{w\lambda\mu}^a\mathcal{F}_{w\rho\theta}^a\right) + C_A\sum_{a=1}^{m}\mathcal{F}_{A\lambda\mu}^a\mathcal{F}_{A\rho\theta}^a\right)$

$\frac{\partial}{\partial P_i^\alpha}\left(g^{\lambda\rho}g^{\mu\theta}\right) = \sum_{pqjk}\eta^{ij}\left(\delta_\alpha^\lambda g^{\mu\theta}P_j^\rho + \delta_\alpha^\rho g^{\mu\theta}P_j^\lambda + \delta_\alpha^\mu g^{\lambda\rho}P_j^\theta + g^{\lambda\rho}\delta_\alpha^\theta P_j^\mu\right)$

$\frac{dL_1}{dP_i^\alpha} = 4\sum_{\theta\lambda\mu}\sum_j\eta^{ij}g^{\mu\lambda}P_j^\theta\{C_G\sum_{a=1}^{3}\mathcal{F}_{r\alpha\mu}^a\mathcal{F}_{r\theta\lambda}^a - \mathcal{F}_{w\alpha\mu}^a\mathcal{F}_{w\theta\lambda}^a + C_A\sum_{a=1}^{m}\mathcal{F}_{A\alpha\mu}^a\mathcal{F}_{A\theta\lambda}^a\}$

For the interaction, V is defined by : $V = \sum_{\alpha=0}^{3}\sum_{i=0}^{3}U^iP_i^\alpha\partial\xi_\alpha$

$\frac{d}{dP_i^\alpha}\left(\sum_{\alpha,i=0}^{3}C_I\mu\frac{1}{i}\frac{1}{M}U^iP_i^\alpha\langle\psi, \nabla_\alpha\psi\rangle\right) = C_I\mu\frac{1}{i}\frac{1}{M}U^i\langle\psi, \nabla_\alpha\psi\rangle$

Equations :

$\forall\alpha,\beta = 0...3 :$

$4\sum_{\theta\lambda\mu}\sum_{ij}\eta^{ij}g^{\mu\lambda}P_i^\beta P_j^\theta\{C_G\sum_{a=1}^{3}\mathcal{F}_{r\alpha\mu}^a\mathcal{F}_{r\theta\lambda}^a - \mathcal{F}_{w\alpha\mu}^a\mathcal{F}_{w\theta\lambda}^a + C_A\sum_{a=1}^{m}\mathcal{F}_{A\alpha\mu}^a\mathcal{F}_{A\theta\lambda}^a\}$

$+C_I\mu\frac{1}{i}\frac{1}{M}\sum_iU^iP_i^\beta\langle\psi, \nabla_\alpha\psi\rangle - L\delta_\beta^\alpha = 0$

$C_I\mu\frac{1}{i}\frac{1}{M}V^\beta\langle\psi, \nabla_\alpha\psi\rangle + 4\sum_{\lambda=0}^{3}\{4C_G\left\langle\mathcal{F}_G^{\beta\lambda}, \mathcal{F}_{G\alpha\lambda}\right\rangle_{Cl} + C_A\left\langle\mathcal{F}_A^{\beta\lambda}, \mathcal{F}_{A\alpha\lambda}\right\rangle\} = L\delta_\beta^\alpha$

$L = C_I \mu \frac{1}{i} \frac{1}{M} \sum_{\lambda=0}^{3} V^\lambda \langle \psi, \nabla_\lambda \psi \rangle + \sum_{\lambda\mu=0}^{3} 4C_G \left\langle \mathcal{F}_G^{\lambda\mu}, \mathcal{F}_{G\lambda\mu} \right\rangle_{Cl} + C_A \left\langle \mathcal{F}_A^{\lambda\mu}, \mathcal{F}_{A\lambda\mu} \right\rangle_{T_1 U}$

The equation reads :

$\forall \alpha, \beta = 0...3$:

$C_I \mu \frac{1}{i} \frac{1}{M} V^\beta \langle \psi, \nabla_\alpha \psi \rangle + 4 \sum_{\gamma=0}^{3} \{4C_G \left\langle \mathcal{F}_{G\alpha\gamma}, \mathcal{F}_G^{\beta\gamma} \right\rangle_{Cl} + C_A \left\langle \mathcal{F}_{A\alpha\gamma}, \mathcal{F}_A^{\beta\gamma} \right\rangle_{T_1 U} \}$

$= \delta_\beta^\alpha \sum_{\lambda\mu} \{ C_I \mu \frac{1}{i} \frac{1}{M_p} \sum_{\lambda=0}^{3} V^\lambda \langle \psi, \nabla_\lambda \psi \rangle + \sum_{\lambda\mu=0}^{3} 4C_G \left\langle \mathcal{F}_G^{\lambda\mu}, \mathcal{F}_{G\lambda\mu} \right\rangle_{Cl} + C_A \left\langle \mathcal{F}_A^{\lambda\mu}, \mathcal{F}_{A\lambda\mu} \right\rangle_{T_1 U} \}$

By taking $\alpha = \beta$ and summing :

$C_I \mu \frac{1}{i} \frac{1}{M} \langle \psi, \nabla_V \psi \rangle + 4 \sum_{\alpha\gamma} \{4C_G \langle \mathcal{F}_{G\alpha\gamma}, \mathcal{F}_G^{\alpha\gamma} \rangle_G + C_A \langle \mathcal{F}_{A\alpha\gamma}, \mathcal{F}_A^{\alpha\gamma} \rangle_{T_1 U} \}$

$= 4\{ C_I \mu \frac{1}{i} \frac{1}{M} \langle \psi, \nabla_V \psi \rangle + \sum_{\lambda\mu=0}^{3} 4C_G \left\langle \mathcal{F}_G^{\lambda\mu}, \mathcal{F}_{G\lambda\mu} \right\rangle_{Cl} + C_A \left\langle \mathcal{F}_A^{\lambda\mu}, \mathcal{F}_{A\lambda\mu} \right\rangle_{T_1 U} \}$

$\Rightarrow$

$$\langle \psi, \nabla_V \psi \rangle = 0 \tag{7.5}$$

$\sum_{\lambda\mu=0}^{3} 4C_G \left\langle \mathcal{F}_G^{\lambda\mu}, \mathcal{F}_{G\lambda\mu} \right\rangle_{Cl} + C_A \left\langle \mathcal{F}_A^{\lambda\mu}, \mathcal{F}_{A\lambda\mu} \right\rangle_{T_1 U} = 8C_G \langle \mathcal{F}_G, \mathcal{F}_G \rangle_G + 2C_A \langle \mathcal{F}_A, \mathcal{F}_A \rangle_U$

the factor 2 accounting for the fact that the indices λ, μ are not ordered in the lagrangian.

The equation reads :

$$\begin{gathered} \forall \alpha, \beta = 0...3 : \\ C_I \mu \frac{1}{i} \frac{1}{M} V^\beta \langle \psi, \nabla_\alpha \psi \rangle + 4 \sum_{\gamma\hat{}0}^{3} \{4C_G \left\langle \mathcal{F}_{G\alpha\gamma}, \mathcal{F}_G^{\beta\gamma} \right\rangle_{Cl} + C_A \left\langle \mathcal{F}_{A\alpha\gamma}, \mathcal{F}_A^{\beta\gamma} \right\rangle_{T_1 U} \} \\ = \delta_\alpha^\beta \left(8C_G \langle \mathcal{F}_G, \mathcal{F}_G \rangle_G + 2C_A \langle \mathcal{F}_A, \mathcal{F}_A \rangle_U \right) \end{gathered} \tag{7.6}$$

The density is deduced from the continuity equation :

$\frac{d\mu}{dt} + \mu \sum_{\alpha,j=1}^{3} Q_j^\alpha \left\{ \left[\partial_\alpha \sigma \cdot \sigma^{-1}, U \right]^j + \frac{1}{2} \frac{1}{\det g} U^j \partial_\alpha (\det g) \right\} = 0$

7.2 MODEL WITH INDIVIDUAL PARTICLES

We consider a system of a fixed number N of particles $p = 1...N$ interacting with the fields, with an action of the general form :

$\int_\Omega \left(\sum_{\alpha\beta} C_G \left(\mathcal{F}^t_{r\alpha\beta} \mathcal{F}^{\alpha\beta}_r - \mathcal{F}^t_{w\alpha\beta} \mathcal{F}^{\alpha\beta}_w \right) + C_A \mathcal{F}^t_{A\alpha\beta} \mathcal{F}^{\alpha\beta}_A \right) \varpi_4$
$+ \sum_{p=1}^N \int_0^T C_I \frac{1}{i} \frac{1}{M_p} \left\langle \psi_p, \nabla_{V_p} \psi_p \right\rangle dt$

7.2.1 Equations for the particles

The variables ψ_p are involved in the last integral only :

$\int_0^T C_I \frac{1}{i} \frac{1}{M_p} \left\langle \psi_p, \nabla_{V_p} \psi_p \right\rangle dt$

so the equations can be deduced from the Euler-Lagrange equations with the variables $r_p(t), w_p(t)$.

The computation holds for each particle and we will drop the index p.

The equations are :

$\forall a = 1, 2, 3 :$

$\frac{dL}{dr_a} = \sum_{\alpha=0}^3 \frac{d}{dt} \left(\frac{dL}{d\frac{dr_a}{st}} \right)$

$\frac{dL}{dw_a} = \sum_{\alpha=0}^3 \frac{d}{dt} \left(\frac{dL}{d\frac{dw_a}{dt}} \right)$

Equation for r

i) Computation of the right hand side

$\frac{dL}{d\frac{dr_a}{dt}} = \frac{d}{d\frac{dr_a}{dt}} C_I \frac{1}{M} \frac{1}{i} \left\{ \left\langle \psi_0, \left[\gamma C \left(\sigma^{-1} \cdot \frac{d\sigma}{dt} + \mathbf{Ad}_{\sigma^{-1}} \widehat{G} \right) \right] \psi_0 \right\rangle + \left\langle \psi_0, [\psi_0] \left[\widehat{\grave{A}} \right] \right\rangle \right\}$

$= C_I \frac{1}{i} \frac{1}{M} \left\langle \psi_0, \gamma C \left(\frac{d}{d\frac{dr_a}{dt}} \left(\sigma^{-1} \cdot \frac{d\sigma}{dt} \right) \right) \psi_0 \right\rangle$

$\frac{d\sigma}{dt} = \sum_{b=1}^3 \frac{dr_b}{dt} \frac{\partial \sigma}{\partial r_b} + \frac{dw_b}{dt} \frac{\partial \sigma}{\partial w_b}$

$\frac{dL}{d\frac{dr_a}{dt}} = C_I \frac{1}{i} \frac{1}{M} \left\langle \psi_0, \gamma C \left(\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} \right) \psi_0 \right\rangle$

$\frac{d}{dt} \frac{dL}{d\frac{dr_a}{dt}} = C_I \frac{1}{i} \frac{1}{M} \left\langle \psi_0, \gamma C \left(\frac{d}{dt} \left(\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} \right) \right) \psi_0 \right\rangle$

ii) Computation of the left hand side

$\frac{dL}{dr_a} = C_I \frac{1}{i} \frac{1}{M} \frac{\partial}{\partial r_a} \sum_{\beta=0}^3 V^\beta \left\langle \psi_0, \left[\gamma C \left(\mathbf{Ad}_{\sigma^{-1}} \left(\partial_\beta \sigma \cdot \sigma^{-1} + G_\beta \right) \right) \right] [\psi_0] + [\psi_0] \left[\grave{A}_\beta \right] \right\rangle$

$= C_I \frac{1}{i} \frac{1}{M} \sum_{\beta=0}^3 \left(\frac{\partial}{\partial r_a} V^\beta \right) \left\langle \psi_0, \left[\gamma C \left(\mathbf{Ad}_{\sigma^{-1}} \left(\partial_\beta \sigma \cdot \sigma^{-1} + G_\beta \right) \right) \right] [\psi_0] + [\psi_0] \left[\grave{A}_\beta \right] \right\rangle$

$+ C_I \frac{1}{i} \frac{1}{M} \sum_{\beta=0}^3 V^\beta \left\langle \psi_0, \left[\gamma C \left(\frac{\partial}{\partial r_a} \left\{ \mathbf{Ad}_{\sigma^{-1}} \left(\partial_\beta \sigma \cdot \sigma^{-1} + G_\beta \right) \right\} \right) \right] [\psi_0] \right\rangle$

$= C_I \frac{1}{i} \frac{1}{M} \sum_{\beta=0}^3 \left(\sum_{j=0}^3 \left(P_j^\beta - \frac{1}{c} P_j^0 V^\beta \right) \left[\frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1}, U \right]^j \right) \langle \psi_0, \nabla_\beta \psi \rangle$

$+ C_I \frac{1}{i} \frac{1}{M} \left\langle \psi_0, \left[\gamma C \left(\frac{\partial}{\partial r_a} \left\{ \mathbf{Ad}_{\sigma^{-1}} \left(\frac{d\sigma}{dt} + \widehat{G} \right) \right\} \right) \right] [\psi_0] \right\rangle$

$= C_I \frac{1}{i} \frac{1}{M} \sum_{\beta=0}^3 \sum_{j=0}^3 P_j^\beta \left[\frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1}, U \right]^j \langle \psi, \nabla_\beta \psi \rangle - C_I \frac{1}{i} \frac{1}{M_p} \sum_{j=0}^3 \frac{1}{c} P_j^0 \left[\frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1}, U \right]^j \langle \psi, \nabla_V \psi \rangle$

$+ C_I \frac{1}{i} \frac{1}{M} \left\langle \psi_0, \left[\gamma C \left(\frac{\partial}{\partial r_a} \nabla_V^G \sigma \right) \right] [\psi_0] \right\rangle$

iii) The equations read :

$\sum_{\beta=0}^3 \sum_{j=0}^3 P_j^\beta \left[\frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1}, U \right]^j \langle \psi, \nabla_\beta \psi \rangle - \sum_{j=0}^3 \frac{1}{c} P_j^0 \left[\frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1}, U \right]^j \langle \psi, \nabla_V \psi \rangle$

$+ \left\langle \psi_0, \left[\gamma C \left(\frac{\partial}{\partial r_a} \nabla_V^G \sigma \right) \right] [\psi_0] \right\rangle = \left\langle \psi_0, \gamma C \left(\frac{d}{dt} \left(\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} \right) \right) \psi_0 \right\rangle$

They are the same as in the first model.

$$a=1,2,3:$$
$$\left\langle \psi_0, \gamma C\left(\left[\sigma^{-1}\cdot\frac{\partial\sigma}{\partial r_a},\nabla_V^G\sigma\right]\right)\psi_0\right\rangle = \sum_{j=0}^3 \frac{1}{c}P_j^0\left[\frac{\partial\sigma}{\partial r_a}\cdot\sigma^{-1},U\right]^j\left\langle\psi,\nabla_\beta\psi\right\rangle$$
$$-\sum_{\beta,j=0}^3 P_j^\beta\left[\frac{\partial\sigma}{\partial r_a}\cdot\sigma^{-1},U\right]^j\left\langle\psi,\nabla_V\psi\right\rangle \quad (7.7)$$

Equation for w

The computation is identical.

$$a=1,2,3:$$
$$\left\langle \psi_0, \gamma C\left(\left[\sigma^{-1}\cdot\frac{\partial\sigma}{\partial w_a},\nabla_V^G\sigma\right]\right)\psi_0\right\rangle = \sum_{j=0}^3 \frac{1}{c}P_j^0\left[\frac{\partial\sigma}{\partial w_a}\cdot\sigma^{-1},U\right]^j\left\langle\psi,\nabla_\beta\psi\right\rangle$$
$$-\sum_{\beta,j=0}^3 P_j^\beta\left[\frac{\partial\sigma}{\partial w_a}\cdot\sigma^{-1},U\right]^j\left\langle\psi,\nabla_V\psi\right\rangle \quad (7.8)$$

7.2.2 Equation for the fields

The equation for the fields is computed by the method of variational derivative. 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 with $d\xi = d\xi^0\wedge d\xi^1\wedge d\xi^2\wedge d\xi^3$:

$$\frac{\delta}{\delta\grave{A}_\alpha^a}\left(\int_\Omega C_A\sum_{\alpha\beta}\mathcal{F}_{A\alpha\beta}^t\mathcal{F}_A^{\alpha\beta}\varpi_4\right)\left(\delta\grave{A}_\alpha^a\right)$$
$$=\int_\Omega C_A\left(\frac{\partial}{\partial\grave{A}_\alpha^a}\sum_{\alpha\beta}\mathcal{F}_{A\alpha\beta}^t\mathcal{F}_A^{\alpha\beta}\det P'-\sum_\beta\frac{d}{d\xi^\beta}\frac{\partial}{\partial\partial_\beta\grave{A}_\alpha^a}\left(\sum_{\alpha\beta}\mathcal{F}_{A\alpha\beta}^t\mathcal{F}_A^{\alpha\beta}\det P'\right)\right)\left(\delta\grave{A}_\alpha^a\right)d\xi$$
$$=\int_\Omega C_A\left(8\sum_\beta\left[\grave{A}_\beta,\mathcal{F}_A^{\alpha\beta}\right]^a\det P'-\sum_\beta\frac{d}{d\xi^\beta}\left(-4C_A\mathcal{F}_A^{a\alpha\beta}\det P'\right)\right)\left(\delta\grave{A}_\alpha^a\right)d\xi$$
$$=C_A\int_\Omega\left(8\sum_\beta\left[\grave{A}_\beta,\mathcal{F}_A^{\alpha\beta}\right]^a+\frac{1}{\det P'}\sum_\beta\frac{d}{d\xi^\beta}\left(4\mathcal{F}_A^{a\alpha\beta}\det P'\right)\right)\left(\delta\grave{A}_\alpha^a\right)\varpi_4$$

For the simple integral a direct computation gives the functional integral :

$$\frac{\delta}{\delta\grave{A}_\alpha^a}\left(\sum_{p=1}^N\int_0^T C_I\frac{1}{i}\frac{1}{M}\left\langle\psi_p,\nabla_{V_p}\psi_p\right\rangle dt\right)\left(\delta\grave{A}_\alpha^a\right)$$
$$=C_I\frac{1}{i}\sum_{p=1}^N\frac{1}{M_p}\int_0^T\left\langle\psi_p,\psi_pV_p^\alpha\delta\grave{A}_\alpha^a\left(q_p\right)\left[\theta_a\right]\right\rangle dt$$
$$=C_I\frac{1}{i}\sum_{p=1}^N\frac{1}{M_p}\int_0^T V_p^\alpha\delta\grave{A}_\alpha^a\left(q_p\right)\left\langle\psi_p,\psi_p\left[\theta_a\right]\right\rangle dt$$

The equation reads :

$\forall\delta\grave{A}_\alpha^a$:

$$C_A\int_\Omega\left(8\sum_\beta\left[\grave{A}_\beta,\mathcal{F}_A^{\alpha\beta}\right]^a+4\frac{1}{\det P'}\sum_\beta\frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{a\alpha\beta}\det P'\right)\right)\left(\delta\grave{A}_\alpha^a\right)\varpi_4$$
$$+\sum_{p=1}^N\int_0^T C_I\frac{1}{i}\frac{1}{M_p}\left(V_p^\alpha\delta\grave{A}_\alpha^a\left(q_p\right)\right)\left\langle\psi_p,\left[\psi_p\right]\left[\theta_a\right]\right\rangle dt=0$$

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 :

$$C_A\int_\Omega\left(8\sum_\beta\left[\grave{A}_\beta,\mathcal{F}_A^{\alpha\beta}\right]^a+4\frac{1}{\det P'}\sum_\beta\frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{a\alpha\beta}\det P'\right)\right)\left(\delta\grave{A}_\alpha^a\right)\varpi_4$$
$$\to C_A\int_0^T\left(8\sum_\beta\left[\grave{A}_\beta,\mathcal{F}_A^{\alpha\beta}\right]^a+4\frac{1}{\det P'}\sum_\beta\frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{a\alpha\beta}\det P'\right)\right)\left(\delta\grave{A}_\alpha^a\left(q_p\left(t\right)\right)\right)dt$$

and the equation reads :

$\forall\delta\grave{A}_\alpha^a$:

$$C_A\int_0^T\left(8\sum_\beta\left[\grave{A}_\beta,\mathcal{F}_A^{\alpha\beta}\right]^{a\prime}+4\frac{1}{\det P'}\sum_\beta\frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{a\alpha\beta}\det P'\right)\right)\left(\delta\grave{A}_\alpha^a\left(m_p\left(t\right)\right)\right)dt$$
$$+\int_0^T C_I\frac{1}{i}\frac{1}{M_p}\left(\delta\grave{A}_\alpha^a\left(m_p\left(t\right)\right)\right)V_p^\alpha\left\langle\psi_p,\left[\psi_p\right]\left[\theta_a\right]\right\rangle dt=0$$

$\forall a,\alpha : C_A \sum_\beta 8\left[\grave{A}_\beta, \mathcal{F}_A^{\alpha\beta}\right]^a + C_I \frac{1}{i}\frac{1}{M_p} V_p^\alpha \left\langle \psi_p, [\psi_p]\,[\theta_a]\right\rangle + 4C_A \frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{\alpha\beta}\det P'\right) = 0$

$\sum_\beta \left[\mathcal{F}_A^{\alpha\beta}, \grave{A}_\beta\right]^a - \frac{C_I}{8C_A}\frac{1}{i}\frac{1}{M_p} V_p^\alpha \left\langle \psi_p, [\psi_p]\,[\theta_a]\right\rangle = \frac{1}{2}\frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{\alpha\beta}\det P'\right)$

$$\phi_A^\alpha - J_A^\alpha = \frac{1}{2}\frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{\alpha\beta}\det P'\right) \tag{7.9}$$

with

$$\phi_A^\alpha = \sum_{a=1}^m \sum_\beta \left[\mathcal{F}_A^{\alpha\beta}, \grave{A}_\beta\right]_{T_1 U} \tag{7.10}$$

$$J_{Ap}^\alpha = \frac{C_I}{8C_A}\frac{1}{i}\frac{1}{M_p} V_p^\alpha \sum_{a=1}^m \left\langle \psi_{0p}, [\psi_{0p}]\,[\theta_a]\right\rangle \overrightarrow{\theta}_a \tag{7.11}$$

So we have the same equation as in the first model and μ disappears. We have similarly :

$\forall\alpha = 0,...3 : \forall\alpha = 0,...3 : \phi_G^\alpha - J_{Gp}^\alpha = \frac{1}{2}\frac{1}{\det P'}\sum_\beta \frac{d}{d\xi^\beta}\left(\mathcal{F}_G^{\alpha\beta}\det P'\right)$

$\phi_G^\alpha = \sum_{\beta=0}^3 \left[\mathcal{F}_r^{\alpha\beta}, G_{r\beta}\right]_{T_1 Spin(3,1)}$

$J_G^\alpha = -\mu\frac{C_I M_p}{16 C_G} V^\alpha \mathbf{Ad}_\sigma \upsilon\,(k_0, 0)$

These equations holds only on the trajectories : $m = q_p(t)$. We will discuss the meaning of these equations in the following.

7.2.3 Tetrad equation

We have to compute the functional derivative on both integrals.

The functional derivative reads for the first :

$\frac{\delta}{\delta P_i^\alpha}\left(\int_\Omega \left(\sum_{\lambda\mu} C_G\left(\mathcal{F}_{r\lambda\mu}^t \mathcal{F}_r^{\lambda\mu} - \mathcal{F}_{w\lambda\mu}^t \mathcal{F}_w^{\lambda\mu}\right) + C_A \mathcal{F}_{A\lambda\mu}^t \mathcal{F}_A^{\lambda\mu}\right)\varpi_4\right)(\delta P_i^\alpha)$

$= \int_\Omega 4\sum_{\theta\lambda\mu}\sum_j \eta^{ij} g^{\mu\lambda} P_j^\theta \{C_G \sum_{a=1}^3 \mathcal{F}_{r\alpha\mu}^a \mathcal{F}_{r\theta\lambda}^a - \mathcal{F}_{w\alpha\mu}^a \mathcal{F}_{w\theta\lambda}^a + C_A \sum_{a=1}^m \mathcal{F}_{A\alpha\mu}^a \mathcal{F}_{A\theta\lambda}^a\}(\delta P_i^\alpha)\varpi_4$

$- \int_\Omega \left(\sum_{\lambda\mu} C_G\left(\mathcal{F}_{r\lambda\mu}^t \mathcal{F}_r^{\lambda\mu} - \mathcal{F}_{w\lambda\mu}^t \mathcal{F}_w^{\lambda\mu}\right) + C_A \mathcal{F}_{A\lambda\mu}^t \mathcal{F}_A^{\lambda\mu}\right)\left(P_\alpha'^i\right)(\delta P_i^\alpha)\varpi_4$

(to account for the derivative with respect to $\det P'$)

For the part related to the interactions V is defined by $V^\alpha = \sum_{i=0}^3 P_i^\alpha U^i$ and

$\frac{\delta}{\delta P_i^\alpha}\int_0^T \left(C_I \frac{1}{i}\frac{1}{M_p}\left\langle \psi_p, \nabla_{V_p}\psi_p\right\rangle\right)dt = C_I \frac{1}{i}\sum_{p=1}^N \frac{1}{M_p}\int_0^T U_p^i \left\langle \psi_p, \nabla_\alpha \psi_p\right\rangle dt$

Thus :

$\delta\mathcal{L} = \int_\Omega 4\sum_{\theta\lambda\mu}\sum_j \eta^{ij} g^{\mu\lambda} P_j^\theta \{C_G \sum_{a=1}^3 \mathcal{F}_{r\alpha\mu}^a \mathcal{F}_{r\theta\lambda}^a - \mathcal{F}_{w\alpha\mu}^a \mathcal{F}_{w\theta\lambda}^a + C_A \sum_{a=1}^m \mathcal{F}_{A\alpha\mu}^a \mathcal{F}_{A\theta\lambda}^a\}(\delta P_i^\alpha)\varpi_4$

$- \int_\Omega \left(\sum_{\lambda\mu} C_G\left(\mathcal{F}_{r\lambda\mu}^t \mathcal{F}_r^{\lambda\mu} - \mathcal{F}_{w\lambda\mu}^t \mathcal{F}_w^{\lambda\mu}\right) + C_A \mathcal{F}_{A\lambda\mu}^t \mathcal{F}_A^{\lambda\mu}\right)\left(P_\alpha'^i\right)(\delta P_i^\alpha)\varpi_4$

$+ \sum_{p=1}^N C_I \frac{1}{i}\frac{1}{M_p}\int_0^T U_p^i \left\langle \psi_p, \nabla_\alpha \psi_p\right\rangle (\delta P_i^\alpha)\,dt$

And the equation $\frac{\delta\mathcal{L}}{\delta P_i^\alpha}(\delta P_i^\alpha) = 0$ reads, for the solutions :

$\forall \delta P_i^\alpha$:

$\int_\Omega 4\sum_{\theta\lambda\mu}\sum_j \eta^{ij} g^{\mu\lambda} P_j^\theta \{C_G \sum_{a=1}^3 \mathcal{F}_{r\alpha\mu}^a \mathcal{F}_{r\theta\lambda}^a - \mathcal{F}_{w\alpha\mu}^a \mathcal{F}_{w\theta\lambda}^a + C_A \sum_{a=1}^m \mathcal{F}_{A\alpha\mu}^a \mathcal{F}_{A\theta\lambda}^a\}(\delta P_i^\alpha)\varpi_4$

$- \int_\Omega \left(\sum_{\lambda\mu} C_G\left(\mathcal{F}_{r\lambda\mu}^t \mathcal{F}_r^{\lambda\mu} - \mathcal{F}_{w\lambda\mu}^t \mathcal{F}_w^{\lambda\mu}\right) + C_A \mathcal{F}_{A\lambda\mu}^t \mathcal{F}_A^{\lambda\mu}\right)\left(P_\alpha'^i\right)(\delta P_i^\alpha)\varpi_4$

$+ \sum_{p=1}^N C_I \frac{1}{i}\frac{1}{M_p}\int_0^T U_p^i \left\langle \psi_p, \nabla_\alpha \psi_p\right\rangle (\delta P_i^\alpha)\,dt = 0$

With the same reasoning as above : for each particle along its trajectory :

$\int_0^T 4\sum_{\theta\lambda\mu}\sum_j \eta^{ij} g^{\mu\lambda} P_j^\theta \{C_G \sum_{a=1}^3 \mathcal{F}_{r\alpha\mu}^a \mathcal{F}_{r\theta\lambda}^a - \mathcal{F}_{w\alpha\mu}^a \mathcal{F}_{w\theta\lambda}^a + C_A \sum_{a=1}^m \mathcal{F}_{A\alpha\mu}^a \mathcal{F}_{A\theta\lambda}^a\}(\delta P_i^\alpha)\,dt$

$- \int_0^T \left(\sum_{\lambda\mu} C_G\left(\mathcal{F}_{r\lambda\mu}^t \mathcal{F}_r^{\lambda\mu} - \mathcal{F}_{w\lambda\mu}^t \mathcal{F}_w^{\lambda\mu}\right) + C_A \mathcal{F}_{A\lambda\mu}^t \mathcal{F}_A^{\lambda\mu}\right)\left(P_\alpha'^i\right)(\delta P_i^\alpha)\,dt$

$+ C_I \frac{1}{i}\frac{1}{M_p}\int_0^T U_p^i \left\langle \psi_p, \nabla_\alpha \psi_p\right\rangle (\delta P_i^\alpha)\,dt = 0$

$4\sum_{\theta\lambda\mu}\sum_j \eta^{ij} g^{\mu\lambda} P_j^{\theta}\{C_G \sum_{a=1}^{3} \mathcal{F}^a_{r\alpha\mu}\mathcal{F}^a_{r\theta\lambda} - \mathcal{F}^a_{w\alpha\mu}\mathcal{F}^a_{w\theta\lambda} + C_A \sum_{a=1}^{m} \mathcal{F}^a_{A\alpha\mu}\mathcal{F}^a_{A\theta\lambda}\}$
$-\left(\sum_{\lambda\mu} C_G \left(\mathcal{F}^t_{r\lambda\mu}\mathcal{F}^{\lambda\mu}_r - \mathcal{F}^t_{w\lambda\mu}\mathcal{F}^{\lambda\mu}_w\right) + C_A \mathcal{F}^t_{A\lambda\mu}\mathcal{F}^{\lambda\mu}_A\right)\left(P'^i_\alpha\right) + C_I \frac{1}{i}\frac{1}{M_p} U^i_p \left\langle \psi_p, \nabla_\alpha \psi_p \right\rangle = 0$
The equation, by product with P_i^{β} and summation on i gives :
$4\sum_{\gamma}\{C_G \sum_{a=1}^{3} \mathcal{F}^a_{r\alpha\gamma}\mathcal{F}^{a\beta\gamma}_r - \mathcal{F}^a_{w\alpha\gamma}\mathcal{F}^{a\beta\gamma}_w + C_A \sum_{a=1}^{m} \mathcal{F}^a_{A\alpha\gamma}\mathcal{F}^{a\beta\gamma}_A\} + C_I \frac{1}{i}\frac{1}{M_p} V^{\beta}_p \left\langle \psi_p, \nabla_\alpha \psi_p \right\rangle$
$= \delta^{\beta}_{\alpha}\left(\sum_{\lambda\mu}\left(C_G \sum_{a=1}^{3}\left(\mathcal{F}^a_{r\lambda\mu}\mathcal{F}^{a\lambda\mu}_r - \mathcal{F}^a_{w\lambda\mu}\mathcal{F}^{a\lambda\mu}_w\right) + C_A \sum_{a=1}^{m} \mathcal{F}^a_{A\lambda\mu}\mathcal{F}^{a\lambda\mu}_A\right)\right)$
$\forall\alpha,\beta = 0...3 : C_I \frac{1}{i}\frac{1}{M_p} V^{\beta}_p \left\langle \psi_p, \nabla_\alpha \psi_p \right\rangle + 4\sum_{\gamma}\{4C_G \left\langle \mathcal{F}_{G\alpha\gamma}, \mathcal{F}^{\beta\gamma}_G \right\rangle_{Cl} + C_A \left\langle \mathcal{F}_{A\alpha\gamma}, \mathcal{F}^{\beta\gamma}_A \right\rangle_{T_1U}\}$
$= \delta^{\alpha}_{\beta}\sum_{\lambda\mu}\left(4C_G \left\langle \mathcal{F}_{G\lambda\mu}, \mathcal{F}^{\lambda\mu}_G \right\rangle_{Cl} + C_A \left\langle \mathcal{F}_{A\lambda\mu}, \mathcal{F}^{\lambda\mu}_A \right\rangle_{T_1U}\right)$
With $\alpha = \beta$ and summing :

$$\left\langle \psi_p, \nabla_V \psi_p \right\rangle = 0 \tag{7.12}$$

$$\begin{gathered}\forall\alpha,\beta = 0...3 : \\ C_I \frac{1}{i}\frac{1}{M_p} V^{\beta}_p \left\langle \psi_p, \nabla_\alpha \psi_p \right\rangle + 4\sum_{\gamma}\{4C_G \left\langle \mathcal{F}_{G\alpha\gamma}, \mathcal{F}^{\beta\gamma}_G \right\rangle_{Cl} + C_A \left\langle \mathcal{F}_{A\alpha\gamma}, \mathcal{F}^{\beta\gamma}_A \right\rangle_{T_1U}\} \\ = \delta^{\beta}_{\alpha}\left(8C_G \left\langle \mathcal{F}_G, \mathcal{F}_G \right\rangle_G + 2C_A \left\langle \mathcal{F}_A, \mathcal{F}_A \right\rangle_U\right)\end{gathered} \tag{7.13}$$

7.3 MORE ON THE EQUATIONS FOR THE PARTICLES

The lagrangian used in the models above is based on the energy, and the conditions at equilibrium express its conservation of the whole system. In a continuous process, without collision and a constant number of particles, the conservation of energy implies for each particle : $\langle \psi, \nabla_V \psi \rangle = 0$: the kinetic energy is traded with the fields and the balance is even.

The Principle of Least Action is complementary to the Conservation of Energy. It brings two additional equations.

Accounting for the first equation $\sum_{\beta=0}^{3} V^\beta \langle \psi, \nabla_\beta \psi \rangle = 0$ they are :

$\left\langle \psi_0, \left[\gamma C \left(\left[\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a}, \nabla_V^G \sigma \right] \right) \right] [\psi_0] \right\rangle = \sum_{\beta,j=0}^{3} P_j^\beta \left[\frac{\partial \sigma}{\partial r_a} \cdot \sigma^{-1}, U \right]^j \langle \psi, \nabla_\beta \psi \rangle$

$\left\langle \psi_0, \left[\gamma C \left(\left[\sigma^{-1} \cdot \frac{\partial \sigma}{\partial w_a}, \nabla_V^G \sigma \right] \right) \right] [\psi_0] \right\rangle = \sum_{\beta,j=0}^{3} P_j^\beta \left[\frac{\partial \sigma}{\partial w_a} \cdot \sigma^{-1}, U \right]^j \langle \psi, \nabla_\beta \psi \rangle$

We have the same equations in both models. This calls for some comments.

i) The equations for the particles are computed with the same method : they hold at any point where there is a particle. In the first case the density does not appear in the result, and of course it is absent in the second. So, apparently, the equations do not depend on the physical location of the particles. Actually it does through the value of the potentials. Particles interact through the fields and, as we consider the particles independently from the fields in the second part of the action, the particles do not interact with each other : there is no collision, and the value of the field at their location is given. We retrieve the influence of the particles on the fields and the metric in the other equations. So the result is consistent with the principle of locality and the concept of field, without action at a distance. But it shows also that, in a given environment (when the potentials are defined), the motion of particles is given by sections of P_G, in particular their trajectories are organized along vector fields, even for individual particles. This is consistent with the idea of a continuous process, but gives a special significance to the representation of motion through spinors. And of course the result holds in SR and even Galilean Geometry.

ii) The equations are the first step : to get genuine PDE one needs to adjust the equations to the initial conditions which for the particles, include the location and the first derivatives. However the existence of vector fields for each kind of particles leads to definite patterns in the general location of the particles. This is more obvious with density, which depends on the vector field V. In a system in equilibrium, with continuous processes, the distribution of matter (and by extension of material bodies) follow distinctive patterns. Notably in Astronomy the arrangement of the orbits of planets in a star system follow typical patterns, which are usually understood as resulting from the interactions between planets, but cannot be computed directly (this is the, unsolved, N-body problem). And similarly for the shapes of galaxies. We have here a more general interpretation.

7.3.1 Solution of the equations

We choose the complex chart.

$\sigma = A + \sum_{a=1}^{3} Z^a \overrightarrow{\kappa}_a$

Z is a vector of $\mathbb{C}^3$ represented by a matrix $[Z]_{3\times 1}$, $r = \operatorname{Re} Z; w = \operatorname{Im} Z; A^2 = 1 - \frac{1}{4} Z^t Z$

$\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} = \sigma^{-1} \cdot \frac{\partial \sigma}{\partial \operatorname{Re} Z_a}; \sigma^{-1} \cdot \frac{\partial \sigma}{\partial w_a} = \sigma^{-1} \cdot \frac{\partial \sigma}{\partial \operatorname{Im} Z_a}$

and using the inertial and charge vectors :

$\left\langle \psi_0, \left[\gamma C \left(\left[\sigma^{-1} \cdot \frac{\partial \sigma}{\partial \operatorname{Re} Z_a}, \nabla_V^G \sigma \right] \right) \right] [\psi_0] \right\rangle = -i \frac{M^2}{2} k_0^t \operatorname{Re} \left[\sigma^{-1} \cdot \frac{\partial \sigma}{\partial \operatorname{Re} Z_a}, \nabla_V^G \sigma \right]$

$\langle \psi, \nabla_\beta \psi \rangle = -i \frac{1}{2} M^2 \left\{ k_0^t \operatorname{Re} \left(\nabla_\beta^G \sigma \right) + k_c^t \left[\grave{A}_\beta \right] \right\}$

$\nabla_V^G \sigma = \sum_{\beta,j=0}^{3} P_j^\beta U^j \nabla_\beta \sigma$ with $\nabla_\beta^G \sigma = \mathbf{Ad}_{\sigma^{-1}} \left(\partial_\beta \sigma \cdot \sigma^{-1} + G_\beta \right) = \sum_{b=1}^{3} X_\beta^b \overrightarrow{\kappa}_b$

the equations read :

$\sum_{\beta,j=0}^{3} P_j^\beta U^j \left\{ k_0^t \operatorname{Re} (X_\beta) + k_c^t \left[\grave{A}_\beta \right] \right\} = 0$

$k_0^t \operatorname{Re}\left[\sigma^{-1} \cdot \frac{\partial \sigma}{\partial \operatorname{Re} Z_a}, X_\beta\right] = \sum_{\beta,j=0}^3 P_j^\beta \left[\frac{\partial \sigma}{\partial \operatorname{Re} Z_a} \cdot \sigma^{-1}, U\right]^j \left\{k_0^t \operatorname{Re}(X_\beta) + k_c^t \left[\grave{A}_\beta\right]\right\}$

$k_0^t \operatorname{Re}\left[\sigma^{-1} \cdot \frac{\partial \sigma}{\partial \operatorname{Im} Z_a}, X_\beta\right] = \sum_{\beta,j=0}^3 P_j^\beta \left[\frac{\partial \sigma}{\partial \operatorname{Im} Z_a} \cdot \sigma^{-1}, U\right]^j \left\{k_0^t \operatorname{Re}(X_\beta) + k_c^t \left[\grave{A}_\beta\right]\right\}$

$\sigma^{-1} \cdot \frac{\partial \sigma}{\partial \operatorname{Re} Z_a} = D(-Z) \frac{\partial Z}{\partial Z_a} \frac{\partial Z_a}{\partial \operatorname{Re} Z_a} = D(-Z) \frac{\partial Z}{\partial Z_a} = \sum_{b=0}^3 [D(-Z)]_a^b \overrightarrow{\kappa}_b$

$\sigma^{-1} \cdot \frac{\partial \sigma}{\partial \operatorname{Im} Z_a} = D(-Z) \frac{\partial Z}{\partial Z_a} \frac{\partial Z_a}{\partial \operatorname{Im} Z_a} = iD(-Z) \frac{\partial Z}{\partial Z_a} = i\sum_{b=0}^3 [D(-Z)]_a^b \overrightarrow{\kappa}_b$

$\frac{\partial \sigma}{\partial \operatorname{Re} Z_a} \cdot \sigma^{-1} = D(Z) \frac{\partial Z}{\partial Z_a} \frac{\partial Z_a}{\partial \operatorname{Re} Z_a} = D(Z) \frac{\partial Z}{\partial Z_a} = \sum_{b=0}^3 [D(Z)]_a^b \overrightarrow{\kappa}_b$

$\frac{\partial \sigma}{\partial \operatorname{Im} Z_a} \cdot \sigma^{-1} = D(Z) \frac{\partial Z}{\partial Z_a} \frac{\partial Z_a}{\partial \operatorname{Im} Z_a} = iD(Z) \frac{\partial Z}{\partial Z_a} = i\sum_{b=0}^3 [D(Z)]_a^b \overrightarrow{\kappa}_b$

$k_0^t \operatorname{Re}\left[D(-Z) \frac{\partial Z}{\partial Z_a}, X_\beta\right] = \sum_{\beta,j=0}^3 P_j^\beta \left[D(Z) \frac{\partial Z}{\partial Z_a}, U\right]^j \left\{k_0^t \operatorname{Re}(X_\beta) + k_c^t \left[\grave{A}_\beta\right]\right\}$

$k_0^t \operatorname{Re}\left[iD(-Z) \frac{\partial Z}{\partial Z_a}, X_\beta\right] = \sum_{\beta,j=0}^3 P_j^\beta \left[iD(Z) \frac{\partial Z}{\partial Z_a} \cdot \sigma^{-1}, U\right]^j \left\{k_0^t \operatorname{Re}(X_\beta) + k_c^t \left[\grave{A}_\beta\right]\right\}$

$\Leftrightarrow -k_0^t \operatorname{Im}\left[D(-Z) \frac{\partial Z}{\partial Z_a}, X_\beta\right] = i\sum_{\beta,j=0}^3 P_j^\beta \left[D(Z) \frac{\partial Z}{\partial Z_a} \cdot \sigma^{-1}, U\right]^j \left\{k_0^t \operatorname{Re}(X_\beta) + k_c^t \left[\grave{A}_\beta\right]\right\}$

By product with i and summation we get the 2 complex equations :

i) $k_0^t \left[D(-Z) \frac{\partial Z}{\partial Z_a}, X_\beta\right] = 2\sum_{\beta,j=0}^3 P_j^\beta \left[D(Z) \frac{\partial Z}{\partial Z_a}, U\right]^j \left\{k_0^t \operatorname{Re}(X_\beta) + k_c^t \left[\grave{A}_\beta\right]\right\}$

ii) $\sum_{\beta,j=0}^3 P_j^\beta U^j k_0^t \left[D(-Z) \frac{\partial Z}{\partial Z_a}, X_\beta\right] = 2\sum_{\beta,j=0}^3 P_j^\beta \left[D(Z) \frac{\partial Z}{\partial Z_a}, U\right]^j \left\{k_0^t \operatorname{Re}(X_\beta) + k_c^t \left[\grave{A}_\beta\right]\right\}$

With :

$\left[D(-Z) \frac{\partial Z}{\partial Z_a}, X_\beta\right] = j(D(-Z)\varepsilon_a) X_\beta = -j(X_\beta) D(-Z)\varepsilon_a$

$U = c(\varepsilon_0 + \overrightarrow{u})$

$\left[D(Z) \frac{\partial Z}{\partial Z_a}, U\right] = c\operatorname{Im}(D(Z)\varepsilon_a) + \left\{\operatorname{Im}(D(Z)\varepsilon_a)^t cu\right\}\varepsilon_0 + j(\operatorname{Re}(D(Z)\varepsilon_a)) cu$

$\left[D(Z) \frac{\partial Z}{\partial Z_a}, U\right]^0 = \left\{\operatorname{Im}(D(Z)\varepsilon_a)^t cu\right\} = c\operatorname{Im}[u]^t [D(Z)] [\varepsilon_a]$

$j > 0$:

$\left[D(Z) \frac{\partial Z}{\partial Z_a}, U\right]^j = c\operatorname{Im}(D(Z)\varepsilon_a)^j - \{j(cu)\operatorname{Re}(D(Z)\varepsilon_a)\}^j = c[\varepsilon_j]^t (\operatorname{Im} D(Z) - j(u)\operatorname{Re} D(Z)) [\varepsilon_a]$

$U^j k_0^t \left[D(-Z) \frac{\partial Z}{\partial Z_a}, X_\beta\right] = -cu^j k_0^t j(X_\beta) D(-Z) [\varepsilon_a] = -c[\varepsilon_j]^t [u] k_0^t j(X_\beta) D(-Z) [\varepsilon_a]$

The equations read :

i) $\sum_{\beta,j=0}^3 P_j^\beta U^j \left\{k_0^t \operatorname{Re}(X_\beta) + k_c^t \left[\grave{A}_\beta\right]\right\} = 0$

ii) $a = 1,2,3$:

$\frac{1}{2}\sum_{\beta=0}^3 \left\{P_0^\beta c k_0^t (-j(X_\beta) D(-Z)\varepsilon_a) + \sum_{j=1}^3 P_j^\beta \left(-c[\varepsilon_j]^t [u] k_0^t j(X_\beta) D(-Z) [\varepsilon_a]\right)\right\} =$

$\sum_{\beta=0}^3 \left\{P_0^\beta c\operatorname{Im}[u]^t [D(Z)] [\varepsilon_a] + \sum_{j=1}^3 P_j^\beta \left(c[\varepsilon_j]^t (\operatorname{Im} D(Z) - j(u)\operatorname{Re} D(Z)) [\varepsilon_a]\right)\right\} \left\{k_0^t \operatorname{Re}(X_\beta) + k_c^t \left[\grave{A}_\beta\right]\right\}$

$\sum_{\beta=0}^3 P_0^\beta \left\{k_0^t [j(X_\beta)] [D(-Z)] [\varepsilon_a] + 2\operatorname{Im} u^t [D(Z)] [\varepsilon_a] \left\{k_0^t \operatorname{Re}(X_\beta) + k_c^t \left[\grave{A}_\beta\right]\right\}\right\}$

$+\sum_{\beta=0}^3 \sum_{j=1}^3$

$P_j^\beta \left\{[\varepsilon_j]^t [u] k_0^t j(X_\beta) D(-Z) [\varepsilon_a] + 2\left([\varepsilon_j]^t (\operatorname{Im} D(Z) - j(u)\operatorname{Re} D(Z)) [\varepsilon_a]\right) \left\{k_0^t \operatorname{Re}(X_\beta) + k_c^t \left[\grave{A}_\beta\right]\right\}\right\}$

$= 0$

thus we can eliminate the index a in the equation :

$\sum_{\beta=0}^3 P_0^\beta \left\{k_0^t [j(X_\beta)] [D(-Z)] + 2\operatorname{Im} u^t [D(Z)] \left\{k_0^t \operatorname{Re}(X_\beta) + k_c^t \left[\grave{A}_\beta\right]\right\}\right\} +$

$\sum_{\beta=0}^3 \sum_{j=1}^3 P_j^\beta \left\{[\varepsilon_j]^t [u] k_0^t j(X_\beta) D(-Z) + 2[\varepsilon_j]^t (\operatorname{Im} D(Z) - j(u)\operatorname{Re} D(Z)) \left\{k_0^t \operatorname{Re}(X_\beta) + k_c^t \left[\grave{A}_\beta\right]\right\}\right\}$

$= 0$

ii) $\sum_{\beta=0}^3 P_0^\beta \left\{k_0^t [j(X_\beta)] [D(-Z)] + 2\operatorname{Im} u^t [D(Z)] \left\{k_0^t \operatorname{Re}(X_\beta) + k_c^t \left[\grave{A}_\beta\right]\right\}\right\} +$

$\sum_{\beta=0}^3 \sum_{j=1}^3 P_j^\beta [\varepsilon_j]^t \left\{[u] [k_0]^t [j(X_\beta)] [D(-Z)] + 2(\operatorname{Im} D(Z) - j(u)\operatorname{Re} D(Z)) \left\{k_0^t \operatorname{Re}(X_\beta) + k_c^t \left[\grave{A}_\beta\right]\right\}\right\}$

$= 0$

In the standard chart :

$k_0^t [j(X_0)] [D(-Z)] + 2\operatorname{Im} u^t [D(Z)] \left\{k_0^t \operatorname{Re}(X_0) + k_c^t \left[\grave{A}_0\right]\right\} +$

$\sum_{\beta=1}^{3}[Q]^{\beta}\left\{[u][k_0]^t[j(X_\beta)][D(-Z)]+2(\operatorname{Im}D(Z)-j(u)\operatorname{Re}D(Z))\left\{k_0^t\operatorname{Re}(X_\beta)+k_c^t\left[\grave{A}_\beta\right]\right\}\right\}$
$=0$
Transpose, with $[D(-Z)]^t=[D(Z)]$
$-[D(Z)][j(X_0)]k_0+2\operatorname{Im}[D(-Z)]u\left\{k_0^t\operatorname{Re}(X_0)+k_c^t\left[\grave{A}_0\right]\right\}+$
$\sum_{\beta=1}^{3}\left\{-[D(Z)][j(X_\beta)][k_0][u]^t+2(\operatorname{Im}D(-Z)+\operatorname{Re}D(Z)j(u))\left\{k_0^t\operatorname{Re}(X_\beta)+k_c^t\left[\grave{A}_\beta\right]\right\}\right\}\left([Q]^{\beta}\right)^t$
$=0$
$[D(Z)][j(k_0)][X_0]+2\operatorname{Im}[D(-Z)]u\left\{k_0^t\operatorname{Re}(X_0)+k_c^t\left[\grave{A}_0\right]\right\}+$
$\sum_{\beta=1}^{3}\left\{[D(Z)][j(k_0)][X_\beta][u]^t+2(\operatorname{Im}D(-Z)+\operatorname{Re}D(-Z)j(u))\left\{k_0^t\operatorname{Re}(X_\beta)+k_c^t\left[\grave{A}_\beta\right]\right\}\right\}\left([Q]^{\beta}\right)^t$
$=0$
$[u]^t\left([Q]^{\beta}\right)^t=\left(([Q][u])^{\beta}\right)=\frac{1}{c}v^{\beta}$
$[D(Z)][j(k_0)][X_0]+2\operatorname{Im}[D(-Z)]u\left\{k_0^t\operatorname{Re}(X_0)+k_c^t\left[\grave{A}_0\right]\right\}+$
$[D(Z)][j(k_0)]\sum_{\beta=1}^{3}\frac{1}{c}v^{\beta}[X_\beta]+2(\operatorname{Im}D(-Z)+\operatorname{Re}D(-Z)j(u))\sum_{\beta=1}^{3}\left([Q]^{\beta}\right)^t\left\{k_0^t\operatorname{Re}(X_\beta)+k_c^t\left[\grave{A}_\beta\right]\right\}$
$=0$
$[D(Z)][j(k_0)]\left([X_0]+\frac{1}{c}\sum_{\beta=1}^{3}v^{\beta}[X_\beta]\right)$
$+2\operatorname{Im}[D(-Z)]u\left\{k_0^t\operatorname{Re}(X_0)+k_c^t\left[\grave{A}_0\right]\right\}$
$+2(\operatorname{Im}D(-Z)+\operatorname{Re}D(-Z)j(u))\sum_{\beta=1}^{3}\left([Q]^{\beta}\right)^t\left\{k_0^t\operatorname{Re}(X_\beta)+k_c^t\left[\grave{A}_\beta\right]\right\}=0$
The first equation gives :
$\sum_{\beta=0}^{3}v^{\beta}\left\{k_0^t\operatorname{Re}(X_\beta)+k_c^t\left[\grave{A}_\beta\right]\right\}=0$
$\Rightarrow$
$c\left\{k_0^t\operatorname{Re}(X_0)+k_c^t\left[\grave{A}_0\right]\right\}=-\sum_{\beta=1}^{3}v^{\beta}\left\{k_0^t\operatorname{Re}(X_\beta)+k_c^t\left[\grave{A}_\beta\right]\right\}$
$\left\{k_0^t\operatorname{Re}(X_0)+k_c^t\left[\grave{A}_0\right]\right\}=-\sum_{\beta=1}^{3}[u]^t\left([Q]^{\beta}\right)^t\left\{k_0^t\operatorname{Re}(X_\beta)+k_c^t\left[\grave{A}_\beta\right]\right\}$
$[D(Z)][j(k_0)]\left([X_0]+\frac{1}{c}\sum_{\beta=1}^{3}v^{\beta}[X_\beta]\right)$
$-2\operatorname{Im}[D(-Z)]u\sum_{\beta=1}^{3}[u]^t\left([Q]^{\beta}\right)^t\left\{k_0^t\operatorname{Re}(X_\beta)+k_c^t\left[\grave{A}_\beta\right]\right\}$
$+2(\operatorname{Im}D(-Z)+\operatorname{Re}D(-Z)j(u))\sum_{\beta=1}^{3}\left([Q]^{\beta}\right)^t\left\{k_0^t\operatorname{Re}(X_\beta)+k_c^t\left[\grave{A}_\beta\right]\right\}=0$
$[D(Z)][j(k_0)]\left([X_0]+\frac{1}{c}\sum_{\beta=1}^{3}v^{\beta}[X_\beta]\right)+$
$2\left\{-\operatorname{Im}[D(-Z)]u[u]^t+\operatorname{Im}D(-Z)+\operatorname{Re}D(-Z)j(u)\right\}\sum_{\beta=1}^{3}\left([Q]^{\beta}\right)^t\left\{k_0^t\operatorname{Re}(X_\beta)+k_c^t\left[\grave{A}_\beta\right]\right\}$
$=0$
$\frac{1}{c}[D(Z)][j(k_0)]\left(c[X_0]+\sum_{\beta=1}^{3}v^{\beta}[X_\beta]\right)$
$+2\left\{\operatorname{Im}D(-Z)\left(1-[u][u]^t\right)+\operatorname{Re}D(-Z)j(u)\right\}\sum_{\beta=1}^{3}\left([Q]^{\beta}\right)^t\left\{k_0^t\operatorname{Re}(X_\beta)+k_c^t\left[\grave{A}_\beta\right]\right\}=0$
$X_\beta=\nabla_\beta^G\sigma=\mathbf{Ad}_{\sigma^{-1}}\left(\partial_\beta\sigma\cdot\sigma^{-1}+G_\beta\right)=\sigma^{-1}\cdot\partial_\beta\sigma+\mathbf{Ad}_{\sigma^{-1}}G_\beta=[D(-Z)][\partial_\beta Z]+[Ad(-Z)][G_\beta]$
$[D(Z)][j(k_0)]\left(c[X_0]+\sum_{\beta=1}^{3}v^{\beta}[X_\beta]\right)$
$=[D(Z)][j(k_0)][D(-Z)]\left(c[\partial_0 Z]+\sum_{\beta=1}^{3}v^{\beta}[\partial_\beta Z]\right)$
$+[D(Z)][j(k_0)][Ad(-Z)]\left(c[G_0]+\sum_{\beta=1}^{3}v^{\beta}[G_\beta]\right)$
$=[D(Z)][j(k_0)]\left([D(-Z)]\left[\frac{dZ}{dt}\right]+[Ad(-Z)]\left[\widehat{G}\right]\right)$
$\operatorname{Im}D(-Z)\left(1-[u][u]^t\right)+\operatorname{Re}D(-Z)j(u)=\operatorname{Re}\left\{D(-Z)\left(j(u)-i\left(1-j(u)j(u)-[u]^t[u]\right)\right)\right\}$

The equations read :

1 $$k_0^t \operatorname{Re}\left(\left[D\left(-Z\right)\right]\left[\tfrac{dZ}{dt}\right]+\left[Ad\left(-Z\right)\right]\left[\widehat{G}\right]\right)+k_c^t\left[\widehat{A}\right]=0$$

2 $$-\tfrac{1}{2c}\left[D\left(Z\right)\right]\left[j\left(k_0\right)\right]\left\{\left[D\left(-Z\right)\right]\left[\tfrac{dZ}{dt}\right]+\left[Ad\left(-Z\right)\right]\left[\widehat{G}\right]\right\}=$$

$$\operatorname{Re}\left\{\left[D\left(-Z\right)\right]\left(i\left(\left[u\right]^t\left[u\right]-1\right)+j\left(u\right)+ij\left(u\right)j\left(u\right)\right)\right\}\sum_{\beta=1}^3\left(\left[Q\right]^\beta\right)^t\left\{k_0^t\operatorname{Re}\left(\nabla_\beta^G\sigma\right)+k_c^t\left[\grave{A}_\beta\right]\right\} \tag{7.14}$$

With :

$D\left(Z\right)=\frac{1}{A}+\frac{1}{2}j\left(Z\right)+\frac{1}{4A}j\left(Z\right)j\left(Z\right)=\frac{1}{4A}\left(A+\sqrt{A^2-4}+j\left(Z\right)\right)\left(A-\sqrt{A^2-4}+j\left(Z\right)\right)$

$\left[Ad\left(Z\right)\right]=1+Aj\left(Z\right)+\frac{1}{2}j\left(Z\right)j\left(Z\right)$

$\left[D\left(Z\right)\right]\left[j\left(k_0\right)\right]\left[D\left(-Z\right)\right]=j\left(k_0\right)+\frac{1}{8A}\left(k_0^tZ\right)j\left(Z\right)-\frac{1}{8A}\left(k_0^tZ\right)j\left(Z\right)j\left(Z\right)+\frac{1}{2A}\left(j\left(Z\right)j\left(k_0\right)-j\left(k_0\right)j\left(Z\right)\right)$

$\left[D\left(Z\right)\right]\left[j\left(k_0\right)\right]\left[Ad\left(-Z\right)\right]=Aj\left(k_0\right)+\frac{1}{4A}\left(k_0^tZ\right)j\left(Z\right)+\frac{1}{2}j\left(Z\right)j\left(k_0\right)-j\left(k_0\right)j\left(Z\right)+\frac{1}{4A}j\left(k_0\right)j\left(Z\right)j\left(Z\right)$

$u=-\frac{1}{A\overline{A}+\frac{1}{4}Z^t\overline{Z}}\operatorname{Im}\left\{\left(A+\frac{1}{4}j\left(Z\right)\right)\overline{Z}\right\}=-\frac{aw-br+\frac{1}{2}j(r)w}{a^2+b^2+\frac{1}{4}\left(r^tr+w^tw\right)}$

$\operatorname{Re}D\left(-Z\right)=\frac{1}{a^2+b^2}\left\{a-\frac{1}{2}\left(a^2+b^2\right)j\left(r\right)+\frac{a}{4}\left(j\left(r\right)j\left(r\right)-j\left(w\right)j\left(w\right)\right)+\frac{b}{4}\left(j\left(r\right)j\left(w\right)+j\left(w\right)j\left(r\right)\right)\right\}$

$\operatorname{Im}D\left(-Z\right)=\frac{1}{a^2+b^2}\left\{-b-\frac{1}{2}\left(a^2+b^2\right)j\left(w\right)-\frac{b}{4}\left(j\left(r\right)j\left(r\right)-j\left(w\right)j\left(w\right)\right)+\frac{a}{4}\left(j\left(r\right)j\left(w\right)+j\left(w\right)j\left(r\right)\right)\right\}$

$\operatorname{Re}\left[Ad\left(-Z\right)\right]=1-\frac{1}{a^2+b^2}\left(aj\left(r\right)+bj\left(w\right)\right)+\frac{1}{2}\left(j\left(r\right)j\left(r\right)-j\left(w\right)j\left(w\right)\right)$

$\operatorname{Im}\left[Ad\left(-Z\right)\right]=\frac{1}{a^2+b^2}\left(bj\left(r\right)-aj\left(w\right)\right)+\frac{1}{2}\left(j\left(r\right)j\left(w\right)+j\left(w\right)j\left(r\right)\right)$

The energy of the matter field is :

$\delta E=C_I\frac{1}{M}\frac{1}{i}\left\langle\psi,\nabla_V\psi\right\rangle=C_I\frac{1}{M}\frac{1}{i}\sum_{\beta=0}^3V^\beta\left\langle\psi,\nabla_\beta\psi\right\rangle$

$\delta E=-C_IMc\left(k_0^t\operatorname{Re}\left(\nabla_0^G\sigma\right)+k_c^t\left[\grave{A}_0\right]\right)-C_I\frac{M}{2}\sum_{\beta=1}^3\left[Q\right]^\beta\left[u\right]\left(k_0^t\operatorname{Re}\left(\nabla_\beta^G\sigma\right)+k_c^t\left[\grave{A}_\beta\right]\right)$

The variational derivative of the energy with respect to u is :

$\frac{\delta E}{\delta u_j}=-C_I\frac{M}{2}\sum_\beta^3Q_j^\beta\left(k_0^t\operatorname{Re}\left(\nabla_\beta^G\sigma\right)+k_c^t\left[\grave{A}_\beta\right]\right)$

Thus : $\sum_{\beta=1}^3\left[Q\right]_j^\beta\left\{k_0^t\operatorname{Re}\left(\nabla_\beta^G\sigma\right)+k_c^t\left[\grave{A}_\beta\right]\right\}=-\frac{2M}{C_I}\frac{\delta E}{\delta u_j}$

The right hand side of the second equation is real, thus :

$\operatorname{Im}\left\{\left[D\left(Z\right)\right]\left[j\left(k_0\right)\right]\left[D\left(-Z\right)\right]\left[\frac{dZ}{dt}\right]\right\}+\operatorname{Im}\left\{\left[D\left(Z\right)\right]\left[j\left(k_0\right)\right]\left[Ad\left(-Z\right)\right]\left[\widehat{G}\right]\right\}=0$

$$\operatorname{Re}\left(\left[D\left(Z\right)\right]\left[j\left(k_0\right)\right]\left[D\left(-Z\right)\right]\right)\frac{dw}{dt}+\operatorname{Im}\left(\left[D\left(Z\right)\right]\left[j\left(k_0\right)\right]\left[D\left(-Z\right)\right]\right)\frac{dr}{dt}=-\operatorname{Im}\left\{\left[D\left(Z\right)\right]\left[j\left(k_0\right)\right]\left[Ad\left(-Z\right)\right]\left[\widehat{G}\right]\right\} \tag{7.15}$$

The translational and rotational motions are related, through the gravitational field, independently on the other fields.

Moreover the first equation

$k_0^t\operatorname{Re}\left(\left[D\left(-Z\right)\right]\left[\frac{dZ}{dt}\right]+\left[Ad\left(-Z\right)\right]\left[\widehat{G}\right]\right)+k_c^t\left[\widehat{A}\right]=0$

reads :

$$k_0^t\operatorname{Re}\left[D\left(-Z\right)\right]\frac{dr}{dt}-k_0^t\operatorname{Im}\left[D\left(-Z\right)\right]\frac{dw}{dt}=-k_0^t\operatorname{Re}\left(\left[Ad\left(-Z\right)\right]\left[\widehat{G}\right]\right)-k_c^t\left[\widehat{A}\right] \tag{7.16}$$

We have seen (motion) that the spatial speed u^tu varies as :

$u^t\frac{du}{dt}=\left(1-\left(u^tu\right)\right)\left(u^t\operatorname{Im}D\left(-Z\right)\frac{dZ}{dt}\right)$

thus it is constant if $\operatorname{Im}D\left(-Z\right)\frac{dZ}{dt}=0$.

7.3.2 Special solutions

The second equation can be written in a more geometric way :

$[D(Z)][j(k_0)]\left([D(-Z)]\left[\frac{dZ}{dt}\right]+[Ad(-Z)]\left[\widehat{G}\right]\right)=[D(Z)][j(k_0)]\left(\nabla_V^G\sigma\right)$

$[D(Z)]^{-1}=A-\frac{1}{2}j(Z)$

$-\frac{1}{2c}[j(k_0)]\left(\nabla_V^G\sigma\right)=$

$\left(A-\frac{1}{2}j(Z)\right)\operatorname{Re}\left\{[D(-Z)]\left(i\left([u]^t[u]-1\right)+j(u)+ij(u)j(u)\right)\right\}$

$\times\sum_{\beta=1}^3\left([Q]^\beta\right)^t\left\{k_0^t\operatorname{Re}\left(\nabla_\beta^G\sigma\right)+k_c^t\left[\grave{A}_\beta\right]\right\}$

The matrix $A-\frac{1}{2}j(Z)$ is invertible, so

$[j(k_0)]\left(\nabla_V^G\sigma\right)=0\Leftrightarrow[X]=0$

with

$[X]=$

$\operatorname{Re}\left\{[D(-Z)]\left(i\left([u]^t[u]-1\right)+j(u)+ij(u)j(u)\right)\right\}\sum_{\beta=1}^3\left([Q]^\beta\right)^t\left\{k_0^t\operatorname{Re}\left(\nabla_\beta^G\sigma\right)+k_c^t\left[\grave{A}_\beta\right]\right\}$

However the real or the imaginary part of $[j(k_0)]\left(\nabla_V^G\sigma\right)$ can be null.

Denoting the real vector $[X]=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$ then :

$[j(k_0)]\left(\nabla_V^G\sigma\right)=\begin{bmatrix}ax_1-\frac{1}{2}r_2x_3+\frac{1}{2}r_3x_2\\ax_2+\frac{1}{2}r_1x_3-\frac{1}{2}r_3x_1\\ax_3-\frac{1}{2}r_1x_2+\frac{1}{2}r_2x_1\end{bmatrix}+i\begin{bmatrix}bx_1-\frac{1}{2}w_2x_3+\frac{1}{2}w_3x_2\\bx_2+\frac{1}{2}w_1x_3-\frac{1}{2}x_1w_3\\bx_3-\frac{1}{2}w_1x_2+\frac{1}{2}w_2x_1\end{bmatrix}$

We have 2 cases of interest :

i) $b=0,\operatorname{Im}Z=0\Rightarrow[j(k_0)]\operatorname{Im}\left(\nabla_V^G\sigma\right)=0$. This is the bonded particle case.

i) $a=0,\operatorname{Re}Z=0\Rightarrow[j(k_0)]\operatorname{Re}\left(\nabla_V^G\sigma\right)=0$. This is the non rotating particle case.

Bonded particles

With the choice of the chart we can always take $w=0$. Then :

$V(t)=c\sum_{\alpha=0}^3P_0^\alpha(q(t))\partial\xi_\alpha=c\varepsilon_0\Rightarrow u=0$

$\widehat{G}=G_0;\widehat{\grave{A}}=\grave{A}_0$

$\operatorname{Im}Z=0;\operatorname{Re}Z=r;A=a_r=\sqrt{1-\frac{1}{4}r^tr}$

$D(Z)=\frac{1}{a_r}+\frac{1}{2}j(r)+\frac{1}{4a_r}j(r)j(r)$

$Ad(Z)=1-a_rj(r)+\frac{1}{2}j(r)j(r)$

The first equation reads :

$k_0^t\left(D(-r)\frac{dr}{dt}+[Ad(-r)]G_{r0}\right)=-k_c^t\left[\grave{A}_0\right]$

The second equation reads :

$D(-r)\left[\frac{dr}{dt}\right]+[Ad(-r)][G_{r0}]=\lambda k_0$

$\lambda k_0^tk_0=\lambda=-k_c^t\left[\grave{A}_0\right]$

$\left[\frac{dr}{dt}\right]=[D(-r)]^{-1}(\lambda k_0-[Ad(-r)][G_{r0}])=[D(-r)]^{-1}(\lambda k_0-[G_{r0}])$

$\left[\frac{dr}{dt}\right]=\left(a_r+\frac{1}{2}j(r)\right)\left(-\left(k_c^t\left[\grave{A}_0\right]\right)k_0-[G_{r0}]\right)$

and with the EM field : $k_c^t\left[\grave{A}_0\right]=-2q\grave{A}_0$

$$\frac{dr}{dt}=\left(a_r+\frac{1}{2}j(r)\right)\left(2q\grave{A}_0k_0-G_{r0}\right)\tag{7.17}$$

This is the gyromagnetic equation, showing that the inertial vector acts as a magnetic moment. It provides a way to measure k_0 and G_{r0}.

If the particle is considered in his own field, without any exterior field, then the field is the one which has for origin the particle itself and is propagated. We have seen that $\delta G_0^a(\tau) = \theta_G(\tau)\delta G_0^a(O), \delta \grave{A}_0^a(\tau) = \theta_A(\tau)\delta \grave{A}_0^a(O)$. And the particle has a rotating motion with a constant speed :

$r = Ct,$

$r = \lambda\left(\left(2q\grave{A}_0 k_0 - G_{r0}\right)\right),$

$a_r = 0 \Rightarrow \lambda^2\left(2q\grave{A}_0 k_0 - G_{r0}\right)^t\left(2q\grave{A}_0 k_0 - G_{r0}\right) = 4$

This is the usual model of an atom.. or a star.

Non rotating particle

If $r = 0, \operatorname{Im} Z = w; \operatorname{Re} Z = 0; A = a = \sqrt{1 + \frac{1}{4}w^t w} \Leftrightarrow 4\left(A^2 - 1\right) = w^t w, b = 0$

$D(Z) = \frac{1}{A} + i\frac{1}{2}j(w) - \frac{1}{4A}j(w)j(w)$

$[Ad(Z)] = 1 + iAj(w) - \frac{1}{2}j(w)j(w)$

$[D(Z)][j(k_0)][D(-Z)] = j(k_0) - \frac{1}{8A}(k_0^t w)j(w) + i\left\{\frac{1}{2A}(j(w)j(k_0) - j(k_0)j(w)) - \frac{1}{8A}(k_0^t w)j(w)j(w)\right\}$

$[D(Z)][j(k_0)][Ad(-Z)] = Aj(k_0) - \frac{1}{4A}(k_0^t w)j(w) - \frac{1}{4A}j(k_0)j(w)j(w) + i\left\{\frac{1}{2}j(w)j(k_0) - j(k_0)j(w)\right\}$

$u = \frac{\sqrt{1+\frac{1}{4}w^t w}}{1+\frac{1}{2}w^t w}w = \frac{A}{2A^2-1}w$

Then the equations read :

i) $k_0^t\left\{\frac{1}{2}j(w)\frac{dw}{dt} + \left(1 - \frac{1}{2}j(w)j(w)\right)\widehat{G}_r + \frac{1}{A}j(w)\widehat{G}_w\right\} + k_c^t\left[\widehat{\grave{A}}\right] = 0$

ii) $\left\{j(k_0) - \frac{1}{8A}(k_0^t w)j(w)\right\}\frac{dw}{dt}$

$= -\left\{Aj(k_0) - \frac{1}{4A}(k_0^t w)j(w) - \frac{1}{4A}j(k_0)j(w)j(w)\right\}\left[\widehat{G}_w\right] - \left\{\frac{1}{2}j(w)j(k_0) - j(k_0)j(w)\right\}\left[\widehat{G}_r\right]$

iii) $-\frac{1}{c}\left\{\frac{1}{2A}(j(w)j(k_0) - j(k_0)j(w)) - \frac{1}{8A}(k_0^t w)j(w)j(w)\right\}\frac{dw}{dt}$

$+\frac{1}{c}\left\{Aj(k_0) - \frac{1}{4A}(k_0^t w)j(w) - \frac{1}{4A}j(k_0)j(w)j(w)\right\}\left[\widehat{G}_r\right] - \frac{1}{c}\left\{\frac{1}{2}j(w)j(k_0) - j(k_0)j(w)\right\}\left[\widehat{G}_w\right]$

$= -\left(\frac{1}{2A^2-1}\right)j(w)\sum_{\beta=1}^3\left([Q]^\beta\right)^t\left\{k_0^t\operatorname{Re}\left(\nabla_\beta^G\sigma\right) + k_c^t\left[\grave{A}_\beta\right]\right\}$

The general solution is $w = \lambda(t)k_0$ and :

$V = c\varepsilon_0 + \frac{A}{2A^2-1}\sum_{j=1}^3[Q']_j^\beta w_j\partial\xi_\beta$

i) $\sum_{\beta=0}^3 V^\beta\left(k_0^t G_{r\beta} + k_c^t\grave{A}_\beta\right) = 0$

ii) $\sum_{\beta=0}^3 V^\beta\left(AG_{w\beta} + \frac{1}{2}\lambda j(k_0)[G_{r\beta}]\right) = \mu k_0$

iii) $\sum_{\beta=1}^3\left([Q]^\beta\right)^t\left\{k_0^t\operatorname{Re}\left(\nabla_\beta^G\sigma\right) + k_c^t\left[\grave{A}_\beta\right]\right\} = \nu k_0$

with complex scalars functions $\mu(t), \nu(t)$.

Geodesics

Geodesics for the gravitational fields are vector fields such that $\nabla_V^G\sigma \in Spin(3)$ which, in complex formalism, is equivalent to $\operatorname{Im}\nabla_V^G\sigma = 0$.

We must have, with the notations above :

$$\begin{bmatrix} bx_1 - \frac{1}{2}w_2x_3 + \frac{1}{2}w_3x_2 \\ bx_2 + \frac{1}{2}w_1x_3 - \frac{1}{2}x_1w_3 \\ bx_3 - \frac{1}{2}w_1x_2 + \frac{1}{2}w_2x_1 \end{bmatrix} = 0$$

so, either $x_i = 0$ or $\det\begin{bmatrix} b & +\frac{1}{2}w_3 & -\frac{1}{2}w_2 \\ -\frac{1}{2}w_3 & b & \frac{1}{2}w_1 \\ +\frac{1}{2}w_2 & -\frac{1}{2}w_1 & b \end{bmatrix} = b\left(b^2 + \frac{1}{4}w^t w\right) = 0 \Leftrightarrow b = 0$

Which, besides the case of bonded particle, implies :

$$\operatorname{Re}\left\{[D(-Z)]\left(i\left([u]^t[u]-1\right)+j(u)+ij(u)j(u)\right)\right\}\sum_{\beta=1}^{3}\left([Q]^\beta\right)^t\left\{k_0^t\operatorname{Re}\left(\nabla_\beta^G\sigma\right)+k_c^t\left[\grave{A}_\beta\right]\right\}=0 \tag{7.18}$$

This is the generic expression for fields of geodesics.

Estimates

In the equation :

$-\frac{1}{2c}[D(Z)][j(k_0)]\left\{[D(-Z)]\left[\frac{dZ}{dt}\right]+[Ad(-Z)]\left[\widehat{G}\right]\right\}=$

$\operatorname{Re}\left\{[D(-Z)]\left(i\left([u]^t[u]-1\right)+j(u)+ij(u)j(u)\right)\right\}\sum_{\beta=1}^{3}\left([Q]^\beta\right)^t\left\{k_0^t\operatorname{Re}\left(\nabla_\beta^G\sigma\right)+k_c^t\left[\grave{A}_\beta\right]\right\}$

the left hand side is usually small so a good estimate of the solution is given by :

$\operatorname{Re}\left\{[D(-Z)]\left(i\left([u]^t[u]-1\right)+j(u)+ij(u)j(u)\right)\right\}\sum_{\beta=1}^{3}\left([Q]^\beta\right)^t\left\{k_0^t\operatorname{Re}\left(\nabla_\beta^G\sigma\right)+k_c^t\left[\grave{A}_\beta\right]\right\}$
$=0$

that is by geodesics.

Theorem 101 *At non relativist speeds particles follow geodesics of the gravitational field.*

This result is important, theoretically and practically. It justifies the common assumption, validated by experience of geodesics as the "normal" trajectories for particles (it includes here the rotational motion) in many models. It helps to specify the choice of a section in a continuous model : particles of the same type have the same behavior and can be represented by a section $\psi \in \mathfrak{X}(Q[E\otimes F,\vartheta])$, but we can also add that, at non relativist speed, the motion can be represented by a geodesic. Then the equations for the particles are satisfied.

Periodic solutions

Important cases are periodic solutions :

- with respect to the time : atoms, molecules or planets in a star system. The solutions are similar to those seen in the previous chapter.

- with respect to space : bonded particles in a regular environment (such as and crystals). 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.

Of particular interest are periodic solutions with respect to the time, as they can be seen as stable states.

We have seen how to model a section $\sigma \in P_G$ such that its integral curves correspond to periodic motions. It can be transposed in the complex chart :

R is a periodic map : $R:\mathbb{R}\to\mathbb{R}^3 :: R(t+T_G)=R(t)$

$\theta:\mathbb{R}^3\to\mathbb{R} :: \theta(\xi)$ gives the period for the points $\varphi_o(ct,\xi)$

$\rho:\mathbb{R}^4\to\mathbb{R}^3 :: \rho(c(t+\theta(\xi)),\xi)=\rho(t,\xi)$

The section is $\sigma(\varphi_o(ct,\xi))=(a_w+\upsilon(0,R(t)))\cdot(a_r+\upsilon(\rho(t,\xi),0))=A+Z(\varphi_o(ct,\xi))$

with

$a_w^2=1+\frac{1}{4}R^tR, a_r^2=1-\frac{1}{4}\rho^t\rho, A^2=1-\frac{1}{4}Z^tZ$

$A=a_wa_r-i\frac{1}{4}\left(R^t\rho\right)=\sqrt{1+\frac{1}{4}R^tR}\sqrt{1-\frac{1}{4}\rho^t\rho}-i\frac{1}{4}\left(R^t\rho\right)$

$Z=a_w\rho+i\left(\frac{1}{2}j(R)\rho+a_rR\right)=\sqrt{1+\frac{1}{4}R^tR}\rho+i\left(\frac{1}{2}j(R)\rho+R\sqrt{1-\frac{1}{4}\rho^t\rho}\right)$

$u=\sum_{a=1}^{3}\frac{a_w}{2a_w^2-1}R$

7.3.3 Deformable solids

A deformable solid can be represented by a spinor field. If the external fields are given, then r, w, μ are deduced from the equations, with the parameters $\widehat{G}, \widehat{\hat{A}}$, 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 ($\widehat{G}, \widehat{\hat{A}}$ do not depend on x) then the solutions r, w depend only on t and the initial conditions : we have usually a rigid solid. The model can be used the other way around. It is built on the assumption that the particles constituting the material body are represented by a matter field, so that its cohesion is kept. The solutions can be seen as the sum of internal fields and external (known) fields. The internal fields are those necessary to keep its cohesion : they counterbalance the external fields.

Of course the model can be useful in Astro-Physics, for instance for the modelling of galaxies. Except perhaps in the core collisions are rare. The rotational inertia of celestial bodies is significant. Star systems follow trajectories which can be represented by a section of P_G, and a model of the first type with a density should be adequate.

7.4 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.

7.4.1 Definition

In writing the field equations we have introduced new quantities, ϕ, J for the particles and for the fields : they are the currents. They both come from derivatives of the lagrangien.

Currents associated to the fields

The quantities ϕ_G, ϕ_A are defined everywhere. Up to a constant, they are derivatives $\frac{\partial L}{\partial \grave{A}^a_\alpha}, \frac{\partial L}{\partial G^a_\alpha}$ of the lagrangian and, as such, are vectors (see covariance of lagrangians). They are valued in the Lie algebras. For instance in the first model, with $L_G = C_G \langle \mathcal{F}_G, \mathcal{F}_G \rangle$ and using the complex notation :

$\frac{\partial L_G}{\partial \operatorname{Re} G^a_\alpha} = -8C_G \sum_{\lambda=0}^3 \left[j \left(\mathcal{F}_r^{\alpha\lambda} \right) G_{r\lambda} - j \left(\mathcal{F}_w^{\alpha\lambda} \right) G_{w\lambda} \right]^a$

$\frac{\partial L_G}{\partial \operatorname{Im} G^a_\alpha} = 8C_G \sum_{\lambda=0}^3 \left[j \left(\mathcal{F}_r^{\alpha\lambda} \right) G_{w\lambda} + j \left(\mathcal{F}_w^{\alpha\lambda} \right) G_{r\lambda} \right]^a$

$\frac{\partial L_G}{\partial G^a_\alpha} = \frac{\partial L_G}{\partial \operatorname{Re} G^a_\alpha} + \frac{1}{i} \frac{\partial L_G}{\partial \operatorname{Im} G^a_\alpha} = -8C_G \sum_{\lambda=0}^3 \left(j \left(\mathcal{F}_G^{\alpha\lambda} \right) G_\lambda \right)^a = -8C_G \sum_{\lambda=0}^3 \left[\mathcal{F}_G^{\alpha\lambda}, G_\lambda \right]^a$

So we can define, at any point $\varphi_o(t,x)$, the tensors :

$\phi_G = \sum_{\beta=0}^3 \left[\mathcal{F}_G^{\alpha\beta}, G_\beta \right] \otimes \partial\xi_\alpha = -\frac{1}{8C_G} \sum_{\alpha=0}^3 \frac{\partial L_G}{\partial G^a_\alpha} \overrightarrow{\kappa}_a \otimes \partial\xi_\alpha$

$\phi_A = \sum_{\beta=0}^3 \left[\mathcal{F}_A^{\alpha\beta}, \grave{A}_\beta \right] \otimes \partial\xi_\alpha = \frac{1}{8C_A} \sum_{\alpha=0}^3 \frac{\partial L_A}{\partial \grave{A}^a_\alpha} \overrightarrow{\theta}_a \otimes \partial\xi_\alpha$

$$\begin{aligned} \phi_G &= \sum_{\beta=0}^3 \left[\mathcal{F}_G^{\alpha\beta}, G_\beta \right]_{T_1 Spin(3,1)} \otimes \partial\xi_\alpha \in T_1 Spin(3,1) \otimes TM \\ \phi_A &= \sum_{\beta=0}^3 \left[\mathcal{F}_A^{\alpha\beta}, \grave{A}_\beta \right]_{T_1 U} \otimes \partial\xi_\alpha \in T_1 U \otimes TM \end{aligned} \tag{7.19}$$

For the EM field the bracket is null on $T_1 U(1)$ thus $\phi_{EM} = 0$.

They are the **currents** associated to the fields.

It is convenient to write the components in matrix format like vectors :

$[\phi_G]_{a=1...6}^{\beta=0..3} : \phi_G = \sum_{\beta=0}^3 [\phi_G]_a^\beta \overrightarrow{\kappa}_a \otimes \partial\xi_\beta$

$[\phi_A]_{a=1...m}^{\beta=0..3} : \phi_A = \sum_{\beta=0}^3 [\phi_A]_a^\beta \overrightarrow{\theta}_a \otimes \partial\xi_\beta$

Expression of the currents

Using the matrix notation $\left[\mathcal{F}^{b*} \right] = \left[\mathcal{F}^{b,\alpha\beta} \right]$ and the structure coefficients :

$[\phi_A]_a^\beta = \sum_{\gamma=0}^3 \left[\mathcal{F}^{\beta\gamma}, \grave{A}_\gamma \right]^a = \sum_{b,c=1}^m C_{bc}^a \left\{ \left[\mathcal{F}^{b*} \right] \left[\grave{A}^c \right]^t \right\}_\beta$ with $\left[\grave{A} \right]_{\beta=0..3}^{a=1...m}$

The matrix $\left[\mathcal{F}^{c*} \right] = [g]^{-1} \left[\mathcal{F}^c \right] [g]^{-1}$

$$\phi_A = \sum_{\beta=0}^3 \sum_{a,b,c=1}^m C_{bc}^a \left\{ [g]^{-1} \left[\mathcal{F}_A^b \right] [g]^{-1} \left[\grave{A}^c \right]^t \right\}^\beta \partial\xi_\beta \otimes \overrightarrow{\theta}_a \tag{7.20}$$

and with the Hodge dual :

$[*\mathcal{F}] = -[g]^{-1} [\mathcal{F}] [g]^{-1} \det P'$

$\phi_A = (\det P) \{ \sum_{\beta=1}^3 \left[\grave{A}_\beta, [*\mathcal{F}_A^w]_\beta \right] \otimes \partial\xi_0$

$+\left(-\left[\grave{A}_0,[*\mathcal{F}_A^w]_1\right]-\left[\grave{A}_2,[*\mathcal{F}_A^r]_3\right]+\left[\grave{A}_3,[*\mathcal{F}_A^r]_2\right]\right)\otimes\partial\xi_1$
$+\left(-\left[\grave{A}_0,[*\mathcal{F}_A^w]_2\right]+\left[\grave{A}_1,[*\mathcal{F}_A^r]_3\right]-\left[\grave{A}_3,[*\mathcal{F}_A^r]_1\right]\right)\otimes\partial\xi_2$
$+\left(-\left[\grave{A}_0,[*\mathcal{F}_A^w]_3\right]-\left[\grave{A}_1,[*\mathcal{F}_A^r]_2\right]+\left[\grave{A}_2,[*\mathcal{F}_A^r]_1\right]\right)\otimes\partial\xi_3\}$

Currents for the gravitational field

We have similarly with $C_{bc}^a=\epsilon(a,b,c)$

$$\phi_G=\sum_{\beta=0}^{3}\sum_{a,b,c=1}^{6}\epsilon(a,b,c)\left\{[g]^{-1}\left[\mathcal{F}_G^b\right][g]^{-1}\left[G^c\right]^t\right\}^\beta\partial\xi_\beta\otimes\overrightarrow{\kappa}_a \tag{7.21}$$

$\phi_G=\sum_{\beta=0}^{3}\sum_{a,b,c=1}^{6}\epsilon(a,b,c)\left\{\left[*\mathcal{F}_G^b\right]\left[G^c\right]^t\right\}^\beta(\det P)\,\partial\xi_\beta\otimes\overrightarrow{\kappa}_a$

and we get with the complex notation :

$[G]_{3\times3}=\begin{bmatrix}G_1^1+iG_1^4 & G_2^1+iG_2^4 & G_3^1+iG_3^4\\ G_1^2+iG_1^5 & G_2^2+iG_2^5 & G_3^2+iG_3^5\\ G_1^3+iG_1^6 & G_2^3+iG_2^6 & G_3^3+iG_3^6\end{bmatrix}$

$[G_0]_{3\times1}=\begin{bmatrix}G_0^1+iG_0^4\\ G_0^2+iG_0^5\\ G_0^3+iG_0^6\end{bmatrix}$

$[*\mathcal{F}^r]=[*\mathcal{F}_r^r]+i\,[*\mathcal{F}_w^r]$

$[*\mathcal{F}^w]=[*\mathcal{F}_r^w]+i\,[*\mathcal{F}_w^w]$

$\phi_G^0=\sum_{a=1}^{3}\sum_{\beta=1}^{3}\left(j(G_\beta)[*\mathcal{F}^r]_\beta\right)^a(\det P)\,\overrightarrow{\kappa}_a$

$\phi_G^1=\sum_{a=1}^{3}\left(-j(G_0)[*\mathcal{F}^r]_1-j(G_2)[*\mathcal{F}^w]_3+j(G_3)[*\mathcal{F}^w]_2\right)^a(\det P)\,\overrightarrow{\kappa}_a$

$\phi_G^2=\sum_{a=1}^{3}\left(-j(G_0)[*\mathcal{F}^r]_2+j(G_1)[*\mathcal{F}^w]_3-j(G_3)[*\mathcal{F}^w]_1\right)^a(\det P)\,\overrightarrow{\kappa}_a$

$\phi_G^3=\sum_{a=1}^{3}\left(-j(G_0)[*\mathcal{F}^r]_3-j(G_1)[*\mathcal{F}^w]_2+j(G_2)[*\mathcal{F}^w]_1\right)^a(\det P)\,\overrightarrow{\kappa}_a$

Currents associated to a signal

The interaction of a particle with a field creates a signal $[\delta\mathcal{F}^a(O)],\left[\delta\grave{A}(O)\right]$ which propagates, along a curve $q(\tau),q(0)=O,\frac{dq}{d\tau}=V$ as :

$[\delta\mathcal{F}^a(\tau)]=\theta(\tau)[K(\tau)]^t[\delta\mathcal{F}^a(O)][K(\tau)]$

$\left[\delta\grave{A}(\tau)\right]=\theta(\tau)\left[\delta\grave{A}(O)\right][K(\tau)]$

The associated current is :

$\delta\phi_A=\sum_{\beta=0}^{3}\sum_{a,b,c=1}^{m}C_{bc}^a\left\{[g]^{-1}[\delta\mathcal{F}_A^c][g]^{-1}\left[\delta\grave{A}^c\right]^t\right\}^\beta\partial\xi_\beta\otimes\overrightarrow{\theta}_a$

$[\delta\phi_A(q(\tau))]_a^\beta=(\theta(\tau))^2\sum_{b,c=1}^{m}C_{bc}^a\left\{[g(q(\tau))]^{-1}[K(\tau)]^t\left[\delta\mathcal{F}^b(O)\right][K(\tau)][g(q(\tau))]^{-1}[K(\tau)]^t\left[\delta\grave{A}^c(O)\right]^t\right\}^\beta$

Along the curve : $[\delta\phi_A(q(\tau))]=(\theta(\tau))^2\left[\widetilde{\delta\phi_A}(O)\right]$ where $\widetilde{\delta\phi_A}(O)$ is the current at O transported along the propagation curve by the flow of V.

Currents associated to the particles

The quantities J_G,J_A are similarly tensors.

Using the charge vector : $a=1...m:k_c^a=-2\epsilon\frac{1}{M_p^2}\frac{1}{i}\langle\psi_0,[\psi_0][\theta_a]\rangle$ and $k_c=-2q$ for the EM field for a matter field :

$$\begin{array}{c} J_G = -\frac{C_I}{16C_G}\mu M\mathbf{Ad}_\sigma \upsilon\left(k_0,0\right)\otimes V \in T_1 Spin\left(3,1\right)\otimes TM \\ J_A = -\frac{C_I}{16C_A}\mu M\sum_{a=1}^m k_c^a \overrightarrow{\theta}_a \otimes V \in T_1 U \otimes TM \\ J_{EM} = \frac{C_I}{8C_{EM}}\mu q M V \in TM \end{array} \quad (7.22)$$

and for individual particles :

$$\begin{array}{c} J_{Gp} = -\frac{C_I}{16C_G} M_p \mathbf{Ad}_\sigma \upsilon\left(k_{0p},0\right)\otimes V_p \in T_1 Spin\left(3,1\right)\otimes TM \\ J_{Ap} = -\frac{C_I}{16C_A} M_p\sum_{a=1}^m k_{cp}^a \overrightarrow{\theta}_a \otimes V_p \in T_1 U \otimes TM \\ J_{EM} = \frac{C_I}{8C_{EM}} q M_p V \in TM \end{array} \quad (7.23)$$

They are the **currents** associated to the particles.

In the first model the currents are defined by the derivatives :

$J_G = -\frac{1}{8C_G}\frac{\partial}{\partial G_\alpha} C_I \mu \frac{1}{i}\frac{1}{M}\left\langle \psi, \nabla_V \psi\right\rangle$

$J_A = -\frac{1}{8C_A}\frac{\partial}{\partial \grave{A}_\alpha} C_I \mu \frac{1}{i}\frac{1}{M}\left\langle \psi, \nabla_V \psi\right\rangle$

and in the second model by the variational derivatives

$J_G = -\frac{1}{8C_G}\frac{\delta}{\delta G_\alpha} C_I \frac{1}{i}\frac{1}{M}\left\langle \psi, \nabla_V \psi\right\rangle$

$J_A = -\frac{1}{8C_A}\frac{\delta}{\delta \grave{A}_\alpha} C_I \frac{1}{i}\frac{1}{M}\left\langle \psi, \nabla_V \psi\right\rangle$

In both cases the computation is done through $\widehat{G}, \widehat{\grave{A}}$.

A matter field, or a section of P_G with a fixed ψ_0, defines a vector field of trajectories, thus $\widehat{G}, \widehat{\grave{A}}$ at any point, and this holds also for a collection of individual particles whose trajectories do not cross. The currents are associated to vector fields on the tensorial bundles :

$P_G\left[T_1 Spin\left(3;1\right), \mathbf{Ad}\right]\otimes P_G\left[\mathbb{R}^4, \mathbf{Ad}\right] \sim P_G\left[T_1 Spin\left(3;1\right)\otimes \mathbb{R}^4, \mathbf{Ad}\right]$

$J_G\left(p\left(m\right)\right) = \left(\mathbf{p}\left(m\right), -\frac{C_I}{16C_G} M_p \mathbf{Ad}_{\sigma(m)}\upsilon\left(k_0,0\right)\otimes U\left(m\right)\right)$

$P_U\left[T_1 U, \varrho\right]\otimes P_G\left[\mathbb{R}^4, \mathbf{Ad}\right]$

$J_A\left(p\left(m\right)\right) = \left(\mathbf{q}\left(m\right), -\frac{C_I}{16C_A} M_p k_c \otimes U\left(m\right)\right)$

with $U\left(m\right) = -\frac{c}{\left\langle \mathbf{Ad}_{\sigma(m)}\varepsilon_0, \varepsilon_0\right\rangle_{Cl}}\mathbf{Ad}_{\sigma(m)}\varepsilon_0$

To each type of particle one can associate a section $\sigma \in \mathfrak{X}\left(P_G\right)$, which gives their trajectories and motion, and it gives also the currents. So, mathematically, the currents for the particles are defined everywhere. Then the density μ or the maps $t \to \psi\left(q\left(t\right)\right)$ precise the support of the currents.

The components $M_p \mathbf{Ad}_{\sigma(m)}\upsilon\left(k_0,0\right), M_p k_c$ depend on m, and ψ_0. 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 so we have opposite currents.

It will be convenient to express the coordinates in matrix form as vectors :

$\left[J_G\right]_{a=1..6}^{\beta=0..3} : J_G = \sum_{a=1}^6 \sum_{\beta=0}^{3.} \left[J_G\right]_a^\beta \partial\xi^\beta \otimes \overrightarrow{\kappa}_a$

$\left[J_A\right]_{a=1..m}^{\beta=0..3} : J_A = \sum_{a=1}^6 \sum_{\beta=0}^{3.} \left[J_A\right]_a^\beta \partial\xi^\beta \otimes \overrightarrow{\theta}_a$

Computation of the gravitational currents associated to particles

$J_G = -\frac{C_I}{16C_G} M\mathbf{Ad}_\sigma \upsilon\left(k_0,0\right)\otimes V$

With the coordinates : $\sigma = \left(a_w + \upsilon\left(0,w\right)\right)\cdot\left(a_r + \upsilon\left(r,0\right)\right)$

$\left[\mathbf{Ad}_{\sigma_r}\right] = \begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix}$

with $\left[C\left(r\right)\right] = \left[1 + a_r j\left(r\right) + \frac{1}{2} j\left(r\right) j\left(r\right)\right]$

$\left[\mathbf{Ad}_{\sigma_w}\right] = \begin{bmatrix} A & -B \\ B & A \end{bmatrix}$

with :
$[A(w)] = \left[1 - \frac{1}{2} j(w) j(w)\right]$
$[B(w)] = a_w [j(w)]$
$[\mathbf{Ad}_\sigma] = \begin{bmatrix} A & -B \\ B & A \end{bmatrix} \begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix} = \begin{bmatrix} AC & -BC \\ BC & AC \end{bmatrix}$
$\mathbf{Ad}_\sigma \upsilon(k_0, 0) = \upsilon([A(w)][C(r)]k_0, -[B(w)][C(r)]k_0)$

$$J_G = \frac{C_I}{16C_G} M\upsilon(-[A(w)][C(r)]k_0, [B(w)][C(r)]k_0) \otimes V \tag{7.24}$$

With
$w \simeq \left(1 + \frac{3}{8}\frac{\|\overrightarrow{v}\|^2}{c^2}\right)\frac{\overrightarrow{v}}{c}$
$a_w \simeq 1 + \frac{1}{8}\frac{\|\overrightarrow{v}\|^2}{c^2}$
$J_G \simeq \frac{C_I}{16C_G} M\upsilon\left(-[C(r)]k_0, \left(1 + \frac{1}{2}\frac{\|\overrightarrow{v}\|^2}{c^2}\right) j\left(\frac{\overrightarrow{v}}{c}\right)[C(r)]k_0\right) \otimes V$

so that usually the current : $J_G \simeq \frac{C_I}{16C_G} M\upsilon(-[C(r)]k_0, 0) \otimes V$ and for $r = 0$: $J_G \simeq -\frac{C_I}{16C_G} M\upsilon(k_0, 0) \otimes V$. The translational motion (represented by w) has an effect, which is usually very weak.

The intensity of the coupling between the gravitational field (represented by the potential) and the particle, can be estimated through the scalar product $\langle J_G, J_G\rangle$, which can be computed with the scalar product on $T_1 Spin(3,1)$.

$\langle J_G, J_G\rangle = \left(\frac{C_I}{16C_G} M\right)^2 \langle V, V\rangle_{TM}$
$\times \langle \upsilon([A(w)][C(r)]k_0, -[B(w)][C(r)]k_0), \upsilon([A(w)][C(r)]k_0, -[B(w)][C(r)]k_0)\rangle_{Cl}$
$\langle \upsilon([A(w)][C(r)]k_0, -[B(w)][C(r)]k_0), \upsilon([A(w)][C(r)]k_0, -[B(w)][C(r)]k_0)\rangle$
$= \frac{1}{4}(k_0^t C^t A^t A C k_0 - k_0^t C^t B^t B C k_0)$
$= \frac{1}{4} k_0^t C^t (A^2 + B^2) C k_0$
$= \frac{1}{4} k_0^t C^t C k_0 = \frac{1}{4} k_0^t k_0 = \frac{1}{4}$
using the identities :
$A = A^t, B^t = -B$
$A^2 + B^2 = I; AB = BA$
$CC^t = C^t C = I_3$
$\langle V, V\rangle = -c^2\left(1 - \frac{\|\overrightarrow{v}\|^2}{c^2}\right)$

$$\langle J_G, J_G\rangle = -\frac{1}{1024}\left(c\frac{C_I}{C_G} M\right)^2 \left(1 - \frac{\|\overrightarrow{v}\|^2}{c^2}\right) \tag{7.25}$$

Expression of the lagrangian with the currents

The meaning of the currents is more obvious by rewriting the lagrangian with them. The interaction term in the lagrangian reads, distinguishing the EM field :
$C_I \frac{1}{i}\frac{1}{M}\langle \psi, \nabla_V \psi\rangle$
$= C_I \frac{1}{i}\frac{1}{M}\left\langle \psi_0, \gamma C\left(\sigma^{-1} \cdot \frac{d\sigma}{dt}\right)\psi_0\right\rangle + C_I \frac{1}{i}\frac{1}{M}\left\langle \psi_0, \gamma C\left(\mathbf{Ad}_{\sigma^{-1}} \sum_{\alpha=0}^3 V^\alpha G_\alpha\right)\psi_0\right\rangle$
$+ C_I \frac{1}{i}\frac{1}{M}\left\langle \psi_0, \sum_{\alpha=0}^3 V^\alpha [\psi_0]\left[\grave{A}_\alpha\right]\right\rangle + C_I \frac{1}{i}\frac{1}{M}\left\langle \psi_0, iq \sum_{\alpha=0}^3 V^\alpha \psi_0 \grave{A}_\alpha\right\rangle$
$C_I \frac{1}{i}\frac{1}{M}\left\langle \psi_0, \gamma C\left(\sigma^{-1} \cdot \frac{d\sigma}{dt}\right)\psi_0\right\rangle = -\epsilon\frac{M^2}{2}\frac{1}{M} C_I k_0^t \operatorname{Re}\left(\sigma^{-1} \cdot \frac{d\sigma}{dt}\right) = \frac{dK}{dt}\frac{2}{\epsilon M_p^2}\epsilon\frac{M^2}{2} C_I = C_I \frac{dK}{dt}$
$C_I \frac{1}{i}\frac{1}{M}\left\langle \psi_0, \gamma C\left(\mathbf{Ad}_{\sigma^{-1}} \sum_{\alpha=0}^3 V^\alpha G_\alpha\right)\psi_0\right\rangle$
$= -\frac{M}{2} C_I k_0^t \operatorname{Re}\left(\mathbf{Ad}_{\sigma^{-1}} \sum_{\alpha=0}^3 V^\alpha G_\alpha\right)$

$= -\frac{M}{2} C_I \sum_{\alpha=0}^{3} V^\alpha \operatorname{Re}\left(k_0^t \left(1 - Aj(Z) + \frac{1}{2} j(Z) j(Z)\right) G_\alpha\right)$
$= -\frac{M}{2} C_I \sum_{\alpha=0}^{3} V^\alpha \operatorname{Re}\left(G_\alpha^t \left(1 + Aj(Z) + \frac{1}{2} j(Z) j(Z)\right) k_0\right)$
$= -\frac{M}{2} C_I \sum_{\alpha=0}^{3} V^\alpha \operatorname{Re}\left(G_\alpha^t \mathbf{Ad}_\sigma k_0\right)$
$= -\frac{M}{2} C_I \sum_{\alpha=0}^{3} V^\alpha \left(\operatorname{Re} G_\alpha^t \operatorname{Re} \mathbf{Ad}_\sigma k_0 - \operatorname{Im} G_\alpha^t \operatorname{Im} \mathbf{Ad}_\sigma k_0\right)$
$= -\frac{M}{2} C_I \sum_{\alpha=0}^{3} V^\alpha 4 \left\langle G_\alpha, \mathbf{Ad}_\sigma \upsilon(k_0, 0)\right\rangle_{Cl}$
$= \frac{M}{2} C_I 4 \frac{16 C_G}{C_I M} \sum_{\alpha=0}^{3} \left\langle G_\alpha, J_G^\alpha\right\rangle_{Cl}$
$= 32 C_G \sum_{\alpha=0}^{3} \left\langle G_\alpha, J_G^\alpha\right\rangle_{Cl}$
$= 32 C_G \mathbf{G}(J_G)$
$C_I \frac{1}{i} \frac{1}{M} \left\langle \psi_0, \sum_{\alpha=0}^{3} V^\alpha [\psi_0] \left[\grave{A}_\alpha\right]\right\rangle$
$= C_I \frac{1}{i} \frac{1}{M} \sum_{\alpha=0}^{3} V^\alpha \sum_{a=1}^{m} \left\langle \psi_0, \grave{A}_\alpha^a [\psi_0] [\theta_a]\right\rangle$
$= C_I \frac{1}{i} \frac{1}{M} \sum_{\alpha=0}^{3} V^\alpha \left\langle \grave{A}_\alpha, \langle \psi_0, [\psi_0] \sum_{a=1}^{m} [\theta_a]\rangle \overrightarrow{\theta_a}\right\rangle_{T_1 U}$
$= 8 C_A \sum_{\alpha=0}^{3} \left\langle \grave{A}_\alpha, J_A^\alpha\right\rangle_{T_1 U}$
$= 8 C_A \grave{\mathbf{A}}(J_A)$
$C_I \frac{1}{i} \frac{1}{M} \left\langle \psi_0, iq \sum_{\alpha=0}^{3} V^\alpha \psi_0 \grave{A}_\alpha\right\rangle = C_I \frac{1}{M} \sum_{\alpha=0}^{3} V^\alpha \grave{A}_\alpha q \langle \psi_0, \psi_0\rangle$
$= C_I \epsilon M q \sum_{\alpha=0}^{3} V^\alpha \grave{A}_\alpha = 8 C_{EM} \sum_{\alpha=0}^{3} J_{EM}^\alpha \grave{A}_\alpha = 8 C_{EM} \grave{\mathbf{A}}_{EM}(J_{EM})$

$$C_I \frac{1}{i} \frac{1}{M} \langle \psi, \nabla_V \psi\rangle = C_I \frac{dK}{dt} + 8\left(4 C_G \mathbf{G}(J_G) + C_A \grave{\mathbf{A}}(J_A) + C_{EM} \grave{\mathbf{A}}_{EM}(J_{EM})\right) \tag{7.26}$$

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. What is significant is that the same occurs with the gravitational field. The expression for the EM field is identical to that of the "other fields".

At equilibrium : $\langle \psi, \nabla_V \psi\rangle = 0$, then, if only the EM field is present :
$\frac{dK}{dt} = -8\left(4 \frac{C_G}{C_I} \mathbf{G}(J_G) + \frac{C_{EM}}{C_I} \frac{C_I}{8 C_{EM}} \mu q M \sum_{\alpha=0}^{3} V^\alpha \grave{A}_\alpha\right) = -32 \frac{C_G}{C_I} \mathbf{G}(J_G) - \mathcal{I} \sum_{\alpha=0}^{3} V^\alpha \grave{A}_\alpha$
with the EM current : $\mathcal{I} = \mu q M$

7.4.2 Main theorem

The fields equations are not computed the same way in the two models, even if the results are similar, and it is important to understand the differences.

Let us start with the first model, and a continuous distribution of particles - a matter field.

The equations express equalities between components of tensors, at each point :
$\forall a, \forall \beta = 0, ...3 : \phi_{Aa}^\beta - J_{Aa}^\beta = \frac{1}{2} \frac{1}{\det P'} \sum_\gamma \frac{d}{d\xi^\gamma}\left(\mathcal{F}_A^{\beta\gamma} \det P'\right)$

The exterior differential $d(*\mathcal{F})$ of $\mathcal{F} \in \Lambda_2(M; \mathbb{R})$ is a 3 form, which reads :

$$d(*\mathcal{F}) = \sum_{\alpha=0}^{3} (-1)^\alpha \left(\sum_{\beta=0}^{3} \partial_\beta \left(\mathcal{F}^{\alpha\beta} \det P'\right)\right) d\xi^0 \wedge ...\widehat{d\xi^\alpha}... \wedge d\xi^3 \tag{7.27}$$

where $\widehat{\ }$ means that the vector is skipped [1].

Thus $\frac{d}{d\xi^\beta}\left(\mathcal{F}_A^{\alpha\beta} \det P'\right) = (-1)^\alpha (d * \mathcal{F}_A)_{0..\widehat{\alpha}..3}$

By product with $(\det P')(-1)^\alpha d\xi^0 \wedge ...\widehat{d\xi^\alpha}... \wedge d\xi^3$ and summing the equation reads :

[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.

$\sum_{\alpha=0}^{3}(-1)^{\alpha}\left(\sum_{\beta}\left[\mathcal{F}_{A}^{\alpha\beta},\grave{A}_{\beta}\right]^{a}\right)(\det P')\,d\xi^{0}\wedge...\widehat{d\xi^{\alpha}}...\wedge d\xi^{3}=\varpi_{4}(\phi_{Aa})=i_{\phi_{Aa}}\varpi_{4}$

$\sum_{\beta=0}^{3}(-1)^{\beta}J_{Aa}^{\beta}(\det P')\,d\xi^{0}\wedge...\widehat{d\xi^{\beta}}...\wedge d\xi^{3}=\varpi_{4}(J_{Aa})=i_{J_{A}a}\varpi_{4}$

$i_{\phi_A}\varpi_4, i_{J_A}\varpi_4$ can be interpreted as the densities of the currents and the equation reads :

$\forall a=1...m: i_{\phi_{Aa}}\varpi_4 - i_{J_{Aa}}\varpi_4=\frac{1}{2}d*\mathcal{F}_A^a$

To the vectors J_{Aa},ϕ_{Aa} we can associate the one forms using the metric :

$J_A=\sum_\beta J_A^\beta\partial\xi_\beta\rightarrow J_A^*=\sum_{\lambda\beta}g_{\beta\lambda}J_A^\lambda d\xi^\beta=\sum_\beta J_{A\beta}^*d\xi^\beta$

$\phi_A=\sum_\beta \phi_A^\beta\partial\xi_\beta\rightarrow \phi_A^*=\sum_{\lambda\beta}g_{\beta\lambda}\phi_A^\lambda d\xi^\beta=\sum_\beta \phi_{A\beta}^*d\xi^\beta$

The Hodge dual of J_{Aa}^* is a 3-form :

$J_a^*\rightarrow *J^*=\sum_{\alpha,\beta=0}^{3}(-1)^{\alpha}g^{\alpha\beta}J_{A\beta}^{*}(\det P')\,d\xi^0\wedge...\widehat{d\xi^\alpha}...\wedge d\xi^3$

$=\sum_{\alpha,\beta=0}^{3}(-1)^{\alpha}g^{\alpha\beta}g_{\beta\lambda}J_{A}^{\lambda}(\det P')\,d\xi^0\wedge...\widehat{d\xi^\alpha}...\wedge d\xi^3$

$=\sum_{\alpha=0}^{3}(-1)^{\alpha}J_{A}^{\alpha}(\det P')\,d\xi^0\wedge...\widehat{d\xi^\alpha}...\wedge d\xi^3$

$i_{J_G}\varpi_4=*J^*$

and similarly for ϕ_A, thus the equation reads :

$*\phi_{Aa}^*-*J_{Aa}^*=\frac{1}{2}d*\mathcal{F}_A^a$

The Hodge duality has the property that : $**\lambda_r=-(-1)^{r(4-r)}\lambda_r$. Thus, taking the Hodge dual :

$**\phi_{Aa}^*-**J_{Aa}^*=\frac{1}{2}*d*\mathcal{F}_A^a$

$**J_A^*=J_A^*, **\phi_A^*=\phi_A^*$

$*d*\mathcal{F}_A^a=\delta\mathcal{F}_A^a$ is the codifferential, the operator δ acting on scalar r forms (Maths.32.2) :

$\delta:\Lambda_r(M;\mathbb{R})\rightarrow\Lambda_{r-1}(M;\mathbb{R})::\delta\lambda_r=*d\circ*\lambda_r$

The equation is :

$\phi_{Aa}^*-J_{Aa}^*=\frac{1}{2}\delta\mathcal{F}_A^a$

and we have similarly for the gravitational field : $\phi_{Ga}^*-J_{Ga}^*=\frac{1}{2}\delta\mathcal{F}_G^a$

For the EM field $\phi_{EM}^*=0:-J_{EM}^*=\frac{1}{2}\delta\mathcal{F}_{EM}$ is the geometric formulation of the 2nd Maxwell's law 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.2491) : for any 1-form λ on TM :

$\forall\lambda\in\Lambda_1(M;\mathbb{R}),\mathcal{F}\in\Lambda_2(M;\mathbb{R}):G_1(\lambda,\delta\mathcal{F})=G_2(d\lambda,\mathcal{F})$

Thus :

$\phi_{Aa}^*-J_{Aa}^*=\frac{1}{2}\delta\mathcal{F}_A^a\Rightarrow\forall\lambda\in\Lambda_1(M;\mathbb{R}):G_1(\lambda,\phi_{Aa}^*-J_{Aa}^*)=\frac{1}{2}G_1(\lambda,\delta\mathcal{F}_A^a)=\frac{1}{2}G_2(d\lambda,\mathcal{F}_A^a)$

So, for any closed form : $G_1(\lambda,\phi_{Aa}^*-J_{Aa}^*)=0$

Take $\lambda=df$ with any function : $G_1(df,\phi_{Aa}^*-J_{Aa}^*)=0=\sum_{\alpha,\beta=0}^{3}g^{\alpha\beta}(\partial_\alpha f)(\phi_{Aa}^*-J_{Aa}^*)_\beta=\sum_{\beta=0}^{3}(\partial_\alpha f)(\phi_{Aa}-J_{Aa})^\beta$

Take $f(m)=\xi^\alpha$ with $\alpha=0,...3$

$(\phi_{Aa}-J_{Aa})^\beta=0\Rightarrow(\phi_{Aa}-J_{Aa})=0\Rightarrow\delta\mathcal{F}_A=*d*\mathcal{F}_A=0\Rightarrow d*\mathcal{F}_A=0$

The result holds for the gravitational field : $\phi_{Ga}-J_{Ga}=0,\delta\mathcal{F}_G=0$

For the EM field : $-J_{EM}=0,\delta\mathcal{F}_{EM}=0$ and indeed we know that the currents are null inside a conductive medium.

In the second model we have written similarly :

$\forall a,\forall\alpha=0,...3:\phi_{Aa}^\alpha-J_{Aa}^\alpha=\frac{1}{2}\frac{1}{\det P'}\sum_\beta\frac{d}{d\xi^\beta}\left(\mathcal{F}_G^{\alpha\beta}\det P'\right)$

however this equality must be understood "in the meaning of distributions" : the computation is based upon the use of variational derivatives, and of "test-functions" $\delta\grave{A}_\alpha^a$ which are smooth, compactly supported one form. Actually the equation is :

$\forall X\in\mathfrak{X}_{\infty,c}(\Lambda_1(M;T_1U)),\forall a:$

$C_A\int_\Omega 8\sum_{\beta,\gamma}\left[\grave{A}_\gamma,\mathcal{F}_A^{\beta\gamma}\right]^a\left(X_\beta^a\right)\varpi_4+4\frac{1}{\det P'}C_A\int_\Omega\sum_{\beta\gamma}\frac{d}{d\xi^\gamma}\left(\mathcal{F}_A^{a\beta\gamma}\det P'\right)\left(X_\beta^a\right)\varpi_4$

$+\int_0^T C_I \frac{1}{i}\frac{1}{M_p}\sum_\beta \left(V_p^\beta X_\beta^a(q)\right)\langle\psi_0,[\psi_0][\theta_a]\rangle dt = 0$
and in the last integral the value of X is involved at the location of each particle.
$d(*\mathcal{F}_A) = \left(\sum_{\beta=0}^3 \partial_\beta\left(\mathcal{F}^{0\beta}\det P'\right)\right) d\xi^1\wedge d\xi^2\wedge d\xi^3 - \left(\sum_{\beta=0}^3 \partial_\beta\left(\mathcal{F}^{1\beta}\det P'\right)\right) d\xi^0\wedge d\xi^2\wedge d\xi^3$
$+\left(\sum_{\beta=0}^3 \partial_\beta\left(\mathcal{F}^{2\beta}\det P'\right)\right) d\xi^0\wedge d\xi^1\wedge d\xi^3 - \left(\sum_{\beta=0}^3 \partial_\beta\left(\mathcal{F}^{3\beta}\det P'\right)\right) d\xi^0\wedge d\xi^1\wedge d\xi^2$
$(d*\mathcal{F}_A)\wedge\delta\grave{A}_0^a = -\delta\grave{A}_0^a\left(\sum_{\beta=0}^3 \partial_\beta\left(\mathcal{F}^{0\beta}\det P'\right)\right) d\xi^0\wedge d\xi^1\wedge d\xi^2\wedge d\xi^3$
$(d*\mathcal{F}_A)\wedge\delta\grave{A}_\gamma^a = -\delta\grave{A}_\gamma^a\left(\sum_{\beta=0}^3 \partial_\beta\left(\mathcal{F}^{\gamma\beta}\det P'\right)\right) d\xi^0\wedge d\xi^1\wedge d\xi^2\wedge d\xi^3$
With :
$\frac{1}{\det P'}\sum_{\beta,\gamma}\partial_\gamma\left(\mathcal{F}_A^{a\beta\gamma}\det P'\right)\left(X_\beta^a\right)\varpi_4 = -\left(d*\mathcal{F}_A^a\right)\wedge\left(\sum_\beta X_\beta^a d\xi^\beta\right)$
$C_A\int_\Omega 4\frac{1}{\det P'}\sum_\gamma\frac{d}{d\xi^\gamma}\left(\mathcal{F}_A^{a\beta\gamma}\det P'\right)\left(\sum_\beta X_\beta^a\right)\varpi_4$
$= -4C_A\int_\Omega d(*\mathcal{F}^a)\wedge\left(\sum_\beta X_\beta^a d\xi^\alpha\right)$
and the currents, the equation reads :
$\forall a: -8C_A\int_\Omega\sum_\beta\phi_{Aa}^\beta X_\beta^a\varpi_4 - 4C_A\int_\Omega d(*\mathcal{F}^a)\wedge\left(\sum_\beta X_\beta^a d\xi^\beta\right) + 8C_A\int_0^T\sum_\beta X_{\beta\alpha}^a J_{Aa}^\beta dt = 0$
$\frac{1}{2}\int_\Omega d(*\mathcal{F}^a)\wedge\left(\sum_\beta X_\beta^a d\xi^\beta\right) = \int_0^T\sum_\beta X_\beta^a J_{Aa}^\beta dt - \int_\Omega\sum_\beta\left(X_\beta^a\right)\phi_{Aa}^\beta\varpi_4$
or : $\forall X\in\mathfrak{X}_{\infty,c}(\Lambda_1(M;\mathbb{R}))$
$\frac{1}{2}\int_\Omega d(*\mathcal{F}_A^a)\wedge X + \int_\Omega X(\phi_{Aa})\varpi_4 = \int_0^T X(J_{Aa})dt$
and we have similarly :
$\frac{1}{2}\int_\Omega d(*\mathcal{F}_G^a)\wedge X + \int_\Omega X(\phi_{Ga})\varpi_4) = \int_0^T X(J_{Ga})dt$
$\frac{1}{2}\int_\Omega d(*\mathcal{F}_{EM})\wedge X = \int_0^T X(J_{EM})dt$
where the currents for the particles are understood as the sum of the current of each particle :
$\int_0^T X(J)dt = \sum_{p=1}^N\int_0^T X(J_p)dt$
If we take $X = *Y_3, Y_3\in\mathfrak{X}_{\infty,c}(\Lambda_3(M;\mathbb{R}))$:
$d(*\mathcal{F}_A^a)\wedge *Y_3 = (-1)^{3\times 1} * Y_3\wedge d(*\mathcal{F}_A^a) = -G_3(Y_3, d(*\mathcal{F}_A^a))\varpi_4$
and with the isomorphism (Maths.1605) : $*:\mathfrak{X}(\Lambda_r(M;\mathbb{R}))\to\mathfrak{X}(\Lambda_{4-r}(M;\mathbb{R}))$:
$G_3(Y_3, d(*\mathcal{F}_A^a)) = G_1(*Y_3, *d(*\mathcal{F}_A^a)) = G_1(X, *d(*\mathcal{F}_A^a)) = G_1(X, \delta\mathcal{F}_A^a)$
$d(*\mathcal{F}_A^a)\wedge *Y_3 = -G_1(X,\delta\mathcal{F}_A^a)\varpi_4 = -\sum_{\lambda,\mu=0}^3 g^{\lambda\mu}X_\lambda(\delta\mathcal{F}^a)_\mu\varpi_4$
Using $\phi_{Aa}^*\in\mathfrak{X}(\Lambda_1(M;\mathbb{R}))$ and $J_{Aa}^* = \sum_{\lambda\alpha}g_{\alpha\lambda}J_{Aa}^\lambda d\xi^\alpha$ as above :
$X(\phi_{Aa}) = \sum_\beta\phi_{Aa}^\beta X_\beta = \sum_{\lambda\beta}g^{\beta\lambda}\phi_{Aa\lambda}^* X_\beta = G_1(X,\phi_{Aa}^*)$
$X(J_{Ga}) = \sum_{\lambda\beta}g^{\beta\lambda}J_{Aa\lambda}^* X_\beta = G_1(X, J_{Aa}^*)$
the equation reads :
$\int_\Omega G_1(X,\phi_{Aa}^*)\varpi_4 - \frac{1}{2}\int_\Omega G_1(X,\delta\mathcal{F}_A^a)\varpi_4 = \int_0^T G_1(X, J_{Aa}^*)dt$
$G_1(X,\delta\mathcal{F}^a) = G_2(dX,\mathcal{F}^a)$
$\int_\Omega G_1(X,\phi_{Aa}^*)\varpi_4 - \frac{1}{2}\int_\Omega G_2(dX,\mathcal{F}_A^a)\varpi_4 = \int_0^T G_1(X, J_{Aa}^*)dt$
For any compactly supported closed form X : $G_1(X,\delta\mathcal{F}_A^a) = G_2(dX,\mathcal{F}^a) = 0$.
$\int_\Omega G_1(X,\phi_{Aa}^*)\varpi_4 = \int_0^T G_1(X, J_{Aa}^*)dt$
Let us take $\lambda = \chi(\Omega_0)d\xi^\beta$ with $\beta = 0,...3$ and Ω_0 any compact :
$G_1(X,\phi_{Aa}^*) = \phi_{Aa\beta}^*$
$G_1(X, J_{Aa}^*) = J_{Aa\beta}^*$
$\int_{\Omega_0}\phi_{A\alpha}^*\varpi_4 = \int_0^T J_{A\alpha}^*\chi(\Omega_0\cap p(t))dt$
Thus the value of ϕ_{Aa} does not depend on $\delta\mathcal{F}_A^a$, which is continuous and defined over Ω, so the only solution is $\delta\mathcal{F}^a = 0$.

Similarly for the gravitational field : $\delta\mathcal{F}_G = 0$, but obviously the last proof fails for the EM field.

Then the equations can be written :
$\forall X\in\mathfrak{X}_{\infty,c}(\Lambda_1(M;\mathbb{R}))$:
$\int_\Omega G_1(X,\phi_{Ga}^*)\varpi_4 = \int_0^T G_1(X, J_{Ga}^*)dt$

$\int_{\Omega} G_1 (X, \phi_a^*) \varpi_4 = \int_0^T G_1 (X, J_{Aa}^*) dt$
$\delta \mathcal{F}_A^a = 0; \delta \mathcal{F}_G^a = 0$

They can still be written : $J_A = \phi_A, J_G = \phi_G$ but it must be understood "in the meaning of distributions" : the quantities on both sides have not the same support, but their measure (through any X) gives the same result.

Take $X = V^* = \sum_{\lambda=0}^3 V^\lambda g_{\lambda\mu} d\xi_\mu$ with a compactly supported smooth vector field V :
$G_1 (V^*, \phi_a^*) = \sum_{\alpha,\beta=0}^3 g^{\alpha\beta} \left(V^\lambda g_{\beta\lambda}\right) \phi_{a\alpha}^* = \phi_a^* (V)$
$G_1 (V^*, J_a^*) = J_a^* (V)$

The equation reads : $\int_{\Omega} \phi_a^* (V) \varpi_4 = \int_0^T J_\alpha^* (V) dt$ and it means, physically, that the measure of the currents ϕ_a, J_a along any vector field brings the same result.

In the second model it is explicitly assumed that the trajectory belongs to some matter field : $\psi (t) = \psi (p (t))$ with a section ψ. In the field equations we do not vary the trajectory and we can associate a current J to the matter field :
$J_G (p (m)) = \left(\mathbf{p} (m) , -\frac{C_I}{16 C_G} M_p \mathbf{Ad}_{\sigma(m)} \upsilon (k_0, 0) \otimes U (m)\right)$

So we can still write : $\phi_a (m) = J_\alpha (m)$ as in the first model with this interpretation.

For the EM field the equation reads :
$-\frac{1}{2} \int_{\Omega} G_1 (X, \delta \mathcal{F}_{EM}) \varpi_4 = \int_0^T G_1 (X, J_{EM}^*) dt.$
and we can will still write : $J_{EM}^* = -\frac{1}{2} \delta \mathcal{F}_{EM}$ with the same meaning.
$\delta \mathcal{F}_{EM} = \sum_{\alpha=0}^3 g_{\alpha\alpha} \left(\sum_{\beta=0}^3 \partial_\beta \left(\mathcal{F}_{EM}^{\alpha\beta} \det P'\right)\right) (\det P) d\xi^\alpha$ (see Annex)

However there is a classic formulation using the laplacian. The Laplacian is the differential operator : $\Delta = -(d\delta + \delta d)$, which does not change the order of a form. Thus :
$\Delta \mathcal{F}_G = -(d\delta + \delta d) \mathcal{F}_G = -\delta d \mathcal{F}_G$
$\Delta \mathcal{F}_A = -(d\delta + \delta d) \mathcal{F}_A = -\delta d \mathcal{F}_A$
For the EM field, with $\mathcal{F}_{EM} = d\grave{A}$:
$\Delta \grave{A} = -(d\delta + \delta d) \grave{A} = -d\delta \grave{A} - \delta \mathcal{F}_{EM}$
$\delta \mathcal{F}_{EM} = -d\delta \grave{A} - \Delta \grave{A}$
The equation reads :
$\frac{1}{2} \int_{\Omega} G_1 \left(X, d\delta \grave{A}\right) \varpi_4 + \frac{1}{2} \int_{\Omega} G_1 \left(X, \Delta \grave{A}\right) \varpi_4 = \int_0^T G_1 (X, J_{EM}^*) dt$
$\frac{1}{2} \int_{\Omega} G_1 \left(X, d\delta \grave{A}\right) \varpi_4 = \frac{1}{2} \int_{\Omega} G_1 \left(\delta X, \delta \grave{A}\right) \varpi_4$
The codifferential of a one form is a function (see Annex) :
$\delta \left(\sum_{\alpha=0}^3 \grave{A}_\alpha d\xi^\alpha\right) = (\det P) \sum_{\alpha,\beta=0}^3 \partial_\alpha \left(g^{\alpha\beta} \grave{A}_\beta (\det P')\right)$
$\delta \left(\sum_{\alpha=0}^3 X_\alpha d\xi^\alpha\right) = (\det P) \sum_{\alpha,\beta=0}^3 \partial_\alpha \left(g^{\alpha\beta} X_\beta (\det P')\right)$
Take $X_\beta = Ct$ on a compact : $\delta X = 0$
$\frac{1}{2} \int_{\Omega} G_1 \left(X, \Delta \grave{A}\right) \varpi_4 = \int_0^T G_1 (X, J_{EM}^*) dt$

and we can write, "in the meaning of distributions" : $\frac{1}{2} \Delta \grave{A} = J_{EM}^*$ or the same equation, using the matter field associated to the particle. This is an alternate formulation of the 2nd Maxwell's law, but the expression of the laplacian is simple only in SR, in GR the codifferential is more convenient.

Theorem 102 *For the EM field*

$$\begin{gathered} \phi_{EM} = 0 \\ \Delta \grave{A}_{EM} = -\delta \mathcal{F}_{EM} = 2 J_{EM}^* \end{gathered} \tag{7.28}$$

For the other fields :

$$\begin{matrix} J_A = \phi_A & J_G = \phi_G \\ d (*\mathcal{F}_A) = 0 & d (*\mathcal{F}_G) = 0 \end{matrix} \tag{7.29}$$

These equations come from the variation of the field, the state of particles being constant. Usually they are called "equation of motion" but this name is inaccurate : the field is a free variable in their proof, and we have seen previously the equations for the particles. They give a special importance to the Hodge dual $*\mathcal{F}$: the metric is part in the equilibrium of the fields.

We see now what can be deduced from these results.

7.4.3 Codifferential Equation

The codifferential equation : $\delta\mathcal{F}_A = 0$; $\delta\mathcal{F}_G = 0$ holds at any point and in the vacuum (or a conductive medium) for the EM field in the first model.

PDE

For any 2 form $\mathcal{K} = \mathcal{K}^r + \mathcal{K}^w$:

$d(\mathcal{K}) = \sum_{\alpha=0}^{3} (-1)^\alpha \left(\sum_{\beta=0}^{3} \partial_\beta \mathcal{K}\right) d\xi^0 \wedge ...\widehat{d\xi^\alpha}... \wedge d\xi^3$

$d(\mathcal{K}) = -\left(\sum_{\beta=1}^{3} \partial_\beta [\mathcal{K}^r]_\beta\right) d\xi^1 \wedge d\xi^2 \wedge d\xi^3$

$+(-\partial_0 [\mathcal{K}^r]_3 + \partial_2 [\mathcal{K}^w]_1 - \partial_1 [\mathcal{K}^w]_2) d\xi^0 \wedge d\xi^1 \wedge d\xi^2$

$+(\partial_0 [\mathcal{K}^r]_2 + \partial_3 [\mathcal{K}^w]_1 - \partial_1 [\mathcal{K}^w]_3) d\xi^0 \wedge d\xi^1 \wedge d\xi^3$

$+(-\partial_0 [\mathcal{K}^r]_1 + \partial_3 [\mathcal{K}^w]_2 - \partial_2 [\mathcal{K}^w]_3) d\xi^0 \wedge d\xi^2 \wedge d\xi^3$

Thus $d(*\mathcal{F}_G) = 0 \Leftrightarrow$

$\sum_{\beta=1}^{3} \partial_\beta [*\mathcal{F}_G^r]_\beta = 0$

$\partial_0 [*\mathcal{F}_G^r]_1 - \partial_3 [*\mathcal{F}_G^w]_2 + \partial_2 [*\mathcal{F}_G^w]_3 = 0$

$\partial_0 [*\mathcal{F}_G^r]_2 + \partial_3 [*\mathcal{F}_G^w]_1 - \partial_1 [*\mathcal{F}_G^w]_3 = 0$

$\partial_0 [*\mathcal{F}_G^r]_3 - \partial_2 [*\mathcal{F}_G^w]_1 + \partial_1 [*\mathcal{F}_G^w]_2 = 0$

$$\begin{gathered} \sum_{\beta=1}^{3} \partial_\beta [*\mathcal{F}_G^r]_\beta = 0 \\ \partial_0 [*\mathcal{F}_G^r] = \sum_{\beta=1}^{3} (\partial_\beta [*\mathcal{F}_G^w]) [j(\varepsilon_\beta)] \end{gathered} \tag{7.30}$$

We have similar equations for the other fields, and the EM field in the first model.

As we have 24 components for $*\mathcal{F}_G$, depending on 4 arguments, these 24 equations do not suffice to determine the 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 $\mathcal{K} \in *_1(M; T_1 Spin(3,1))$ (not unique) such that : $*\mathcal{F}_G = d\mathcal{K}$ (Maths.7.6.4)

Thus in matrix notation with $[d\mathcal{K}_r^r]_{3\times3}, [d\mathcal{K}_w^r]_{3\times3}, [d\mathcal{K}_r^w]_{3\times3}, [d\mathcal{K}_w^w]_{3\times3}$:

$[*\mathcal{F}_r^r] = [d\mathcal{K}_r^r]; [*\mathcal{F}_r^w] = [d\mathcal{K}_r^w]$

$[*\mathcal{F}_w^r] = [d\mathcal{K}_w^r]; [*\mathcal{F}_w^w] = [d\mathcal{K}_w^w]$

The Hodge dual is defined by :

$[*\mathcal{F}^r] = \left([\mathcal{F}^r] j(H_0)[h] - [\mathcal{F}^w]\left(H_0^0 [h] - [H_0][H_0]^t\right)\right) \det P'$

$[*\mathcal{F}^w] = -\left([\mathcal{F}^r][h]^{-1} \det h + [\mathcal{F}^w][h] j(H_0)\right) \det P'$

So, using $*\mathcal{F}_G = d\mathcal{K}$ and inversing the relation one gets :

$[\mathcal{F}^r] = -\left([d\mathcal{K}^r] j(H)[h] + [d\mathcal{K}^w]\left([H][H]^t - [H]_0^0 [h]\right)\right)(\det P')$

$[\mathcal{F}^w] = \left([d\mathcal{K}^r][h]^{-1} \det h + [d\mathcal{K}^w][h] j(H)\right)(\det P')$

Chern Identity

The Chern identity reads for the gravitational field in complex notation : $\operatorname{Re} Tr\left([\mathcal{F}^r]^t [\mathcal{F}^w]\right) = 0$

A straightforward computation gives :

$Tr\left([\mathcal{F}^r]^t [\mathcal{F}^w]\right) = (\det P')^2 \times$

$Tr\{-[h]\, j(H)\, [d\mathcal{K}^r]^t [d\mathcal{K}^r]\, [h]^{-1} \det h + \left([H][H]^t - [H]_0^0 [h]\right) [d\mathcal{K}^w]^t [d\mathcal{K}^r]\, [h]^{-1} \det h$

$-[h]\, j(H)\, [d\mathcal{K}^r]^t [d\mathcal{K}^w]\, [h]\, j(H) + \left([H][H]^t - [H]_0^0 [h]\right) [d\mathcal{K}^w]^t [d\mathcal{K}^w]\, [h]\, j(H)\}$

$= (\det P')^2 \left\{ Tr\, [d\mathcal{K}^w]^t [d\mathcal{K}^r] (\det P)^2 + \left(H^t [h]^{-1} H\right) Tr\, [d\mathcal{K}^w]\, j\left([h]^{-1} H\right) [d\mathcal{K}^w]^t \det [h]^{-1} \right\}$

$Tr\left([\mathcal{F}^r]^t [\mathcal{F}^w]\right) = Tr\, [d\mathcal{K}^w]^t [d\mathcal{K}^r] = Tr\, [d\mathcal{K}^r]^t [d\mathcal{K}^w] = Tr\, [*\mathcal{F}^r]^t [*\mathcal{F}^w]$

with

$\left([H][H]^t - [H]_0^0 [h]\right) = \left([h]\, j(H)\, [h]\, j(H)\, [h] + (\det P)^2 [h]\right) \det [h]^{-1}$

$Tr\, [M]\, j(r)\, [M]^t = 0$

For any differential of a one form : $\sum_{p=1}^{3} \left([d\mathcal{K}^w]_p [d\mathcal{K}^r]_p\right) = 0$

$(\partial_0 k_1 - \partial_1 k_0)(\partial_3 k_2 - \partial_2 k_3) + (\partial_0 k_2 - \partial_2 k_0)(\partial_1 k_3 - \partial_3 k_1) + (\partial_0 k_3 - \partial_3 k_0)(\partial_2 k_1 - \partial_1 k_2) = 0$

Thus : $Tr\left([\mathcal{F}^r]^t [\mathcal{F}^w]\right) = Tr\left([d\mathcal{K}^r]^t [d\mathcal{K}^w]\right) = 0$

and the Chern-Weil identity is always met, which comforts the codifferential equation. A similar computation can be done for the fields other than the EM field.

Identification of the one form

The question which arises is : what is this one form $\mathcal{K}$? It should be a one form on M valued in the Lie algebra, which is equivariant in a change of gauge (thus this precludes the potential).

The current J for particles can be defined everywhere, through a section $\psi \in \mathfrak{X}(Q[E \otimes F, \vartheta])$. Its dual J^* is then a one form valued in the Lie algebra. If there is only one type of particle, that is a unique fundamental state ψ_0, the natural answer for the choice of the one form $\mathcal{K}$ is then : $*\mathcal{F} = dJ^*$, for each type of field, except the EM field, which implies, with the previous result : $*\mathcal{F}(m) = d\phi^*(m)$ and the field follows a PDE, involving the metric, which replaces the codifferential equation.

This assumption is sensible in a model of the first type, the density is then part of the current J. If there are different types of particles it does not stand, but then a continuous model is probably not realistic.

So I let open this assumption. If it checks it would provide a direct path to the computation of the fields.

7.4.4 Currents equations

Gravitational field

The goal is to give a more convenient expression to the current equations $\phi_G = J_G$, using the Hodge dual $*\mathcal{F}_G$.

$[\phi_G], [J_G]$ are 4×6 matrices.

$J_G = -\frac{C_I}{16 C_G} M \mathbf{Ad}_\sigma \upsilon(k_0, 0) \otimes V$

$[J_G]_a^\beta = -\frac{C_I}{16 C_G} M \{\mathbf{Ad}_\sigma \upsilon(k_0, 0)\}^a V^\beta$

In complex format let us denote in the basis of $T_1 Spin(3,1)$

$[k_G]_{1\times 3} = -\frac{C_I}{16 C_G} M \{\mathbf{Ad}_\sigma \upsilon(k_0, 0)\}$

then $[J_G] = [V][k_G]$

In complex format the equations read :

$\phi^0_{Ga} = \sum_{\beta=1}^3 \left(j\left(G_\beta\right) \left[*\mathcal{F}^r\right]_\beta \right)^a (\det P) = \left[V^0\right] \left[k_G\right] = c\left[k_G\right]_a$
$\phi^1_{Ga} = \left(-j\left(G_0\right)\left[*\mathcal{F}^r\right]_1 - j\left(G_2\right)\left[*\mathcal{F}^w\right]_3 + j\left(G_3\right)\left[*\mathcal{F}^w\right]_2\right)^a (\det P) = v^1 \left[k_G\right]_a$
$\phi^2_{Ga} = \left(-j\left(G_0\right)\left[*\mathcal{F}^r\right]_2 + j\left(G_1\right)\left[*\mathcal{F}^w\right]_3 - j\left(G_3\right)\left[*\mathcal{F}^w\right]_1\right)^a (\det P) = v^2 \left[k_G\right]_a$
$\phi^3_{Ga} = \left(-j\left(G_0\right)\left[*\mathcal{F}^r\right]_3 - j\left(G_1\right)\left[*\mathcal{F}^w\right]_2 + j\left(G_2\right)\left[*\mathcal{F}^w\right]_1\right)^a (\det P) = v^3 \left[k_G\right]_a$
We denote :
$[G]_{3\times 3} = \left[G^a_\beta\right]^{a=1..3}_{\beta=1..3}, [G_0]_{3\times 1} = \left[G^a_0\right]^{a=1..3}$

The first equation
$(\det P) \sum_{\beta=1}^3 j\left(G_\beta\right)\left[*\mathcal{F}^r\right]_\beta = c\left[k_G\right]$
is equivalent to :
$\left[*\mathcal{F}^r\right][G]^t - \left(\left[*\mathcal{F}^r\right][G]^t\right)^t = c\left[j\left(k_G\right)\right](\det P')$
with $[M][N] - ([M][N])^t = j\left(\sum_{\beta=1}^3 j\left([N]^\beta\right)[M]_\beta\right)$
The other equations read :

$$-\begin{bmatrix} j\left(G_0\right)\left[*\mathcal{F}^r\right]_1 \\ j\left(G_0\right)\left[*\mathcal{F}^r\right]_2 \\ j\left(G_0\right)\left[*\mathcal{F}^r\right]_3 \end{bmatrix}_{9\times 1} + \begin{bmatrix} 0 & j\left(G_3\right) & -j\left(G_2\right) \\ -j\left(G_3\right) & 0 & j\left(G_1\right) \\ j\left(G_2\right) & -j\left(G_1\right) & 0 \end{bmatrix}_{9\times 9} \begin{bmatrix} \left[*\mathcal{F}^w\right]_1 \\ \left[*\mathcal{F}^w\right]_2 \\ \left[*\mathcal{F}^w\right]_3 \end{bmatrix}_{9\times 1}$$
$$= \begin{bmatrix} v^1\left[k_G\right]^t \\ v^2\left[k_G\right]^t \\ v^3\left[k_G\right]^t \end{bmatrix}_{9\times 1} \det P'$$

The second 9×9 matrix is invertible if $\det G \neq 0$:

$$\begin{bmatrix} 0 & j\left(G_3\right) & -j\left(G_2\right) \\ -j\left(G_3\right) & 0 & j\left(G_1\right) \\ j\left(G_2\right) & -j\left(G_1\right) & 0 \end{bmatrix}^{-1} = \begin{bmatrix} N_{11} & N_{12} & N_{13} \\ N_{21} & N_{22} & N_{23} \\ N_{31} & N_{32} & N_{33} \end{bmatrix}$$

with :
$\left[N_{pq}\right]_{3\times 3} = \left(2\left[G_q\right]_{3\times 1}\left[G_p\right]^t_{1\times 3} - \left[G_p\right]_{3\times 1}\left[G_q\right]^t_{1\times 3}\right)_{3\times 3} \frac{1}{\det[G]}$
$\left[N_{pq}\right]^a_b = \left(2G^a_q G^b_p - G^a_p G^b_q\right)\frac{1}{\det[G]}$
Then these equations give $\left[*\mathcal{F}^w\right]_p, p = 1,2,3$ with respect to $J_G, \left[*\mathcal{F}^r\right], G$.

$$\begin{bmatrix} \left[*\mathcal{F}^w\right]_1 \\ \left[*\mathcal{F}^w\right]_2 \\ \left[*\mathcal{F}^w\right]_3 \end{bmatrix} = \begin{bmatrix} N_{11} & N_{12} & N_{13} \\ N_{21} & N_{22} & N_{23} \\ N_{31} & N_{32} & N_{33} \end{bmatrix} \left(\begin{bmatrix} v^1\left[k_G\right]^t \\ v^2\left[k_G\right]^t \\ v^3\left[k_G\right]^t \end{bmatrix} \det P' + \begin{bmatrix} j\left(G_0\right)\left[*\mathcal{F}^r\right]_1 \\ j\left(G_0\right)\left[*\mathcal{F}^r\right]_2 \\ j\left(G_0\right)\left[*\mathcal{F}^r\right]_3 \end{bmatrix} \right)$$

$a, p = 1,2,3$:
$\left[*\mathcal{F}^w\right]^a_p = \sum_{q,b=1}^3 \left[N_{pq}\right]^a_b v^q\left[k_G\right]_b + \left[N_{pq}\right]^a_b \left(j\left(G_0\right)\left[*\mathcal{F}^r\right]_q\right)^b$
$= \frac{1}{\det[G]}\sum_{q,b=1}^3 \left(2G^a_q G^b_p - G^a_p G^b_q\right) v^q\left[k_G\right]_b \det P' + \left(2G^a_q G^b_p - G^a_p G^b_q\right)\left(j\left(G_0\right)\left[*\mathcal{F}^r\right]_q\right)^b$
$= \frac{1}{\det[G]}\sum_{q,b=1}^3 \left(2G^a_q\left(\left[k_G\right][G]\right)_p v^q - G^a_p\left([G][v]\right)^b\left[k_G\right]_b\right)\det P'$
$+ \sum_{b=1}^3 \left(2[G]^b_p\left(j\left(G_0\right)\left[*\mathcal{F}^r\right][G]^t\right)^b_a - [G]^a_p\left(j\left(G_0\right)\left[*\mathcal{F}^r\right]\right)^b_q\left([G]^t\right)^q_b\right)\}$
$= \frac{1}{\det[G]}\{\left(2\left([G][v]\right)^a\left(\left[k_G\right][G]\right)_p - \left(\left[k_G\right][G][v]\right)[G]^a_p\right)\det P'$
$+ \left(-2\left([G]\left[*\mathcal{F}^r\right]^t j\left(G_0\right)[G]\right)^a_p - [G]^a_p Tr\left(j\left(G_0\right)\left[*\mathcal{F}^r\right][G]^t\right)\right)\}$
$\left[*\mathcal{F}^w\right] = \frac{1}{\det[G]}\left(2\left([G][v]\right)\left(\left[k_G\right][G]\right) - \left(\left[k_G\right][G][v]\right)[G]\right)\det P'$
$+ \left(-2[G]\left[*\mathcal{F}^r\right]^t j\left(G_0\right)[G] - [G]\,Tr\left(j\left(G_0\right)\left[*\mathcal{F}^r\right][G]^t\right)\right)\}$
$= 2\frac{\det P'}{\det[G]}\left([G][v]\right)\left(\left[k_G\right][G]\right) - \frac{\det P'}{\det[G]}\left(\left[k_G\right][G][v]\right)[G]$
$-2\frac{1}{\det[G]}[G]\left[*\mathcal{F}^r\right]^t j\left(G_0\right)[G] - [G]\frac{1}{\det[G]}Tr\left(j\left(G_0\right)\left[*\mathcal{F}^r\right][G]^t\right)$

From the first equation :
$[G]\left[*\mathcal{F}^r\right]^t = \left[*\mathcal{F}^r\right][G]^t - c\left[j\left(k_G\right)\right]\left(\det P'\right)$
$[G]\left[*\mathcal{F}^r\right]^t j\left(G_0\right)[G] = \left(\left[*\mathcal{F}^r\right][G]^t - c\left[j\left(k_G\right)\right]\left(\det P'\right)\right) j\left(G_0\right)[G]$
$= \left[*\mathcal{F}^r\right][G]^t j\left(G_0\right)[G] - c\left[j\left(k_G\right)\right]\left[j\left(G_0\right)\right][G]\left(\det P'\right)$
$= \left[*\mathcal{F}^r\right] j\left([G]^{-1} G_0\right)\det[G] - c\left[j\left(k_G\right)\right]\left[j\left(G_0\right)\right][G]\left(\det P'\right)$
with $[G]^t j\left(G_0\right)[G] = j\left([G]^{-1} G_0\right)\det[G]$
$Tr\left(j\left(G_0\right)\left[*\mathcal{F}^r\right][G]^t\right) = -Tr\left(\left(\left[*\mathcal{F}^r\right][G]^t\right)^t j\left(G_0\right)\right)$
$= -Tr\left(\left(\left[*\mathcal{F}^r\right][G]^t - c\left[j\left(k_G\right)\right]\left(\det P'\right)\right) j\left(G_0\right)\right)$
$= -Tr\left(\left[*\mathcal{F}^r\right][G]^t j\left(G_0\right)\right) + cTr\left(\left(\left[j\left(k_G\right)\right]\left(\det P'\right)\right) j\left(G_0\right)\right)$
$= -Tr\left(j\left(G_0\right)\left[*\mathcal{F}^r\right][G]^t\right) + cTr\left(\left(\left[j\left(k_G\right)\right]\left(\det P'\right)\right) j\left(G_0\right)\right)$
$Tr\left(j\left(G_0\right)\left[*\mathcal{F}^r\right][G]^t\right) = \frac{1}{2}cTr\left(\left[j\left(k_G\right)\right]\left[j\left(G_0\right)\right]\right)\left(\det P'\right) = -c\left[k_G\right]\left[G_0\right]\left(\det P'\right)$
$\left[*\mathcal{F}^w\right] = 2\frac{\det P'}{\det[G]}\left([G][v]\right)\left(\left[k_G\right][G]\right) - \frac{\det P'}{\det[G]}\left(\left[k_G\right][G][v]\right)[G]$
$-2\frac{1}{\det[G]}\left(\left[*\mathcal{F}^r\right] j\left([G]^{-1} G_0\right)\det[G] - c\left[j\left(k_G\right)\right]\left[j\left(G_0\right)\right][G]\left(\det P'\right)\right) + [G]\frac{1}{\det[G]}c\left[k_G\right]\left[G_0\right]\left(\det P'\right)$
$= -2\left[*\mathcal{F}^r\right] j\left([G]^{-1} G_0\right) + \frac{\det P'}{\det[G]}\left\{2\left([G][v]\right)\left[k_G\right] - \left(\left[k_G\right][G][v]\right) + 2c\left[j\left(k_G\right)\right]\left[j\left(G_0\right)\right] + c\left[k_G\right]\left[G_0\right]\right\}[G]$
$\left[*\mathcal{F}^w\right] + 2\left[*\mathcal{F}^r\right] j\left([G]^{-1} G_0\right) = \frac{\det P'}{\det[G]}\left\{2\left(\left([G][v]\right)\left[k_G\right] + c\left[j\left(k_G\right)\right]\left[j\left(G_0\right)\right]\right) + \left[k_G\right]\left(c\left[G_0\right] - [G][v]\right)\right\}[G]$
$= \frac{\det P'}{\det[G]}\left\{2\left(j\left(k_G\right) j\left([G][v]\right) + \left[k_G\right][G][v] + c\left[j\left(k_G\right)\right]\left[j\left(G_0\right)\right]\right) + \left[k_G\right]\left(c\left[G_0\right] - [G][v]\right)\right\}[G]$
$= \frac{\det P'}{\det[G]}\left\{2j\left(k_G\right)\left(j\left([G][v]\right) + c\left[j\left(G_0\right)\right]\right) + \left[k_G\right]\left(c\left[G_0\right] + [G][v]\right)\right\}[G]$
$\left[*\mathcal{F}^w\right] + 2\left[*\mathcal{F}^r\right] j\left([G]^{-1} G_0\right) = \frac{\det P'}{\det[G]}\left\{2j\left(k_G\right) j\left(cG_0 + [G][v]\right) + \left[k_G\right]\left(c\left[G_0\right] + [G][v]\right)\right\}[G]$
So the currents equations for the gravitational field are equivalent to :

$$\begin{gathered}\left[*\mathcal{F}^r\right][G]^t - \left(\left[*\mathcal{F}^r\right][G]^t\right)^t = c\left[j\left(k_G\right)\right]\left(\det P'\right) \\ \left[*\mathcal{F}^w\right] + 2\left[*\mathcal{F}^r\right] j\left([G]^{-1} G_0\right) = \\ \frac{\det P'}{\det G}\left\{2j\left(k_G\right) j\left(cG_0 + [G][v]\right) + \left[k_G\right]\left(c\left[G_0\right] + [G][v]\right)\right\}[G]\end{gathered} \tag{7.31}$$

These equations give 18 first order PDE on the 12 variables for the potentials, with parameters the metric and the currents for the particles.

EM field

For the EM field there is a single equation, which reads :
$-\frac{1}{2}\delta\mathcal{F}_{EM}(m) = \frac{1}{2}\Delta\grave{A}_{EM}(m) = \left[J^*_{EM}(m)\right]$ with
As above we can compute PDE expressed with the Hodge dual $*\mathcal{F}_{EM}$ or the potential. In SR the use of the Laplacian and the potential is more usual, but in RG actually the use of $\delta\mathcal{F}_{EM}$ is simpler.
$\left[J_{EM}(m)\right]_{4\times 1} = \mu(m)\frac{C_I}{8C_{EM}}qM\left[V(m)\right],\ \mu = 1$ in the second model.
Denote $k_E = \mu\frac{C_I}{4C_{EM}}qM$
$\left[J^*_{EM}\right]_{1\times 4} = \frac{1}{2}k_E[V]^t[g]$
The equation reads : $\left[\delta\mathcal{F}_{EM}\right]_{1\times 4} = -k_E[V]^t[g]$
with (see Annex) :
$\delta\mathcal{F}_{EM} = \sum_{\alpha=0}^3 g_{\alpha\alpha}\left(\sum_{\beta=0}^3\partial_\beta\left(\mathcal{F}^{\alpha\beta}_{EM}\det P'\right)\right)\left(\det P\right)d\xi^\alpha$
$= g_{00}\left(\sum_{\beta=0}^3\partial_\beta\left(\mathcal{F}^{0\beta}_{EM}\det P'\right)\right)\left(\det P\right)d\xi^0 + \sum_{\alpha=1}^3 g_{\alpha\alpha}\partial_0\left(\mathcal{F}^{\alpha 0}_{EM}\det P'\right)\left(\det P\right)d\xi^\alpha$

$+g_{11}\left(\partial_2\left(\mathcal{F}_{EM}^{12}\det P'\right)+\partial_3\left(\mathcal{F}_{EM}^{13}\det P'\right)\right)(\det P)\,d\xi^1$
$+g_{22}\left(\partial_1\left(\mathcal{F}_{EM}^{21}\det P'\right)+\partial_3\left(\mathcal{F}_{EM}^{23}\det P'\right)\right)(\det P)\,d\xi^2$
$+g_{33}\left(\partial_1\left(\mathcal{F}_{EM}^{31}\det P'\right)+\partial_2\left(\mathcal{F}_{EM}^{32}\det P'\right)\right)(\det P)\,d\xi^3$
$=\{g_{00}\left(\sum_{\beta=0}^{3}-\partial_\beta\left[*\mathcal{F}\right]_0^\beta\right)d\xi^0$
$+g_{11}\left(-\partial_0\left[*\mathcal{F}\right]_1^0-\partial_2\left[*\mathcal{F}\right]_1^2-\partial_3\left[*\mathcal{F}\right]_1^3\right)d\xi^1$
$+g_{22}\left(-\partial_0\left[*\mathcal{F}\right]_2^0-\partial_1\left[*\mathcal{F}\right]_2^1-\partial_3\left[*\mathcal{F}\right]_2^3\right)d\xi^2$
$+g_{33}\left(-\partial_0\left[*\mathcal{F}\right]_3^0-\partial_1\left[*\mathcal{F}\right]_3^1-\partial_2\left[*\mathcal{F}\right]_3^2\right)d\xi^3\}\det P$
$=-\sum_{\beta,\gamma=0}^{3}\{g_{\gamma\gamma}\partial_\beta\left(\left[*\mathcal{F}_{EM}\right]_\gamma^\beta\right)d\xi^\gamma\}\det P$
$=-k_E\sum_{\beta,\gamma=0}^{3}\left\{[V]^t[g]\right\}_\beta d\xi^\beta$

And we get the 4 scalar linear first order PDE in E,B :

$$\gamma=0..3:g_{\gamma\gamma}\sum_{\beta=0}^{3}\{\partial_\beta\left(\left[*\mathcal{F}_{EM}\right]_\gamma^\beta\right)\}\det P=-k_E\left\{[V]^t[g]\right\}_\gamma \tag{7.32}$$

7.5 ENERGY AND MOMENTUM OF THE SYSTEM

7.5.1 Energy of the system

The lagrangian is the balance of energy between the components of the system.

Energy of the fields

The energy density is, for the gravitational field :

$\langle \mathcal{F}_G, \mathcal{F}_G \rangle_G$

$= \frac{1}{4} \sum_{a=1}^{3} \sum_{\{\alpha\beta\}} \mathcal{F}_r^{a\alpha\beta} \mathcal{F}_{r\alpha\beta}^a - \mathcal{F}_w^{a\alpha\beta} \mathcal{F}_{w\alpha\beta}^a$

$= \frac{1}{4} \sum_{a=1}^{3} G_2 \left(\mathcal{F}_r^a, \mathcal{F}_r^a \right) - G_2 \left(\mathcal{F}_w^a \mathcal{F}_w^a \right)$

$= \frac{1}{4} \sum_{a=1}^{3} G_2 \left(\mathcal{F}_r^a, dG_r^a \right) + 2G_2 \left(\mathcal{F}_r^a, \sum_{\{\alpha,\beta\}=0}^{3} \left(j \left(G_{r\alpha} \right) G_{r\beta} - j \left(G_{w\alpha} \right) G_{w\beta} \right)^a d\xi^\alpha \wedge d\xi^\beta \right)$

$-G_2 \left(\mathcal{F}_w^a dG_w^a \right) - 2G_2 \left(\mathcal{F}_w^a, \sum_{\{\alpha,\beta\}=0}^{3} \left(j \left(G_{w\alpha} \right) G_{r\beta} + j \left(G_{r\alpha} \right) G_{w\beta} \right)^a d\xi^\alpha \wedge d\xi^\beta \right)$

The codifferential is the adjoint of the exterior differential :

$G_2 \left(\mathcal{F}_r^a, dG_r^a \right) = G_1 \left(\delta \mathcal{F}_r^a, G_r^a \right)$

$G_2 \left(\mathcal{F}_w^a, dG_w^a \right) = G_1 \left(\delta \mathcal{F}_w^a, G_w^a \right)$

and on shell [2] : $\delta \mathcal{F}_G = 0$

$\langle \mathcal{F}_G, \mathcal{F}_G \rangle_G = \frac{1}{2} \sum_{a=1}^{3} \{ G_2 \left(\mathcal{F}_r^a, \sum_{\{\alpha,\beta\}=0}^{3} \left(j \left(G_{r\alpha} \right) G_{r\beta} - j \left(G_{w\alpha} \right) G_{w\beta} \right)^a d\xi^\alpha \wedge d\xi^\beta \right)$

$-G_2 \left(\mathcal{F}_w^a, \sum_{\{\alpha,\beta\}=0}^{3} \left(j \left(G_{w\alpha} \right) G_{r\beta} + j \left(G_{r\alpha} \right) G_{w\beta} \right)^a d\xi^\alpha \wedge d\xi^\beta \right) \}$

$= 2 \left\langle \mathcal{F}_G, \sum_{a=1}^{6} \sum_{\{\alpha,\beta\}=0}^{3} \left[G_\alpha, G_\beta \right]^a d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\kappa}_a \right\rangle_G$

We have proven earlier that :

$\langle X, [Y, Z] \rangle_G = \sum_{\{\alpha\beta\}} \left\langle X^{\alpha\beta}, [Y_\alpha, Z_\beta] \right\rangle_{Cl} = \sum_{\{\alpha\beta\}} \left\langle \left[X^{\alpha\beta}, Y_\alpha \right], Z_\beta \right\rangle_{Cl}$

$\left\langle \mathcal{F}_G, \sum_{a=1}^{6} \sum_{\{\alpha,\beta\}=0}^{3} \left[G_\alpha, G_\beta \right]^a d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\kappa}_a \right\rangle_G$

$= - \left\langle \mathcal{F}_G, \sum_{a=1}^{6} \sum_{\{\alpha,\beta\}=0}^{3} \left[G_\beta, G_\alpha \right]^a d\xi^\alpha \wedge d\xi^\beta \otimes \overrightarrow{\kappa}_a \right\rangle_G$

$= - \sum_{\{\alpha\beta\}} \left\langle \left[\mathcal{F}_G^{\alpha\beta}, G_\beta \right], G_\alpha \right\rangle_{Cl} = -\frac{1}{2} \sum_{\alpha\beta} \left\langle \left[\mathcal{F}_G^{\alpha\beta}, G_\beta \right], G_\alpha \right\rangle_{Cl}$

$= -\frac{1}{2} \sum_\alpha \left\langle \sum_\beta \left[\mathcal{F}_G^{\alpha\beta}, G_\beta \right], G_\alpha \right\rangle_{Cl}$

$= -\frac{1}{2} \sum_\alpha \left\langle \phi_G^\alpha, G_\alpha \right\rangle_{Cl} = -\frac{1}{2} \mathbf{G} \left(\phi \right)$

On shell :

$$\langle \mathcal{F}_G, \mathcal{F}_G \rangle_G = -\mathbf{G} \left(\phi_G \right) \tag{7.33}$$

And we have similarly for the fields other than EM :

$\langle \mathcal{F}_A, \mathcal{F}_A \rangle$

$= \sum_{a=1}^{m} \sum_{\{\alpha\beta\}} \mathcal{F}_A^{a\alpha\beta} \mathcal{F}_{A\alpha\beta}^a$

$= \sum_{a=1}^{m} G_2 \left(\mathcal{F}_A^a, \mathcal{F}_A^a \right)$

$= \sum_{a=1}^{m} G_2 \left(\mathcal{F}_A^a, d\grave{A}^a \right) + 2G_2 \left(\mathcal{F}_A^a, \sum_{\{\alpha,\beta\}=0}^{3} \left[\grave{A}_\alpha, \grave{A}_\beta \right]^a d\xi^\alpha \wedge d\xi^\beta \right)$

$= \sum_{a=1}^{m} G_1 \left(\delta \mathcal{F}_A^a, \grave{A}^a \right) + 2G_2 \left(\mathcal{F}_A^a, \sum_{\{\alpha,\beta\}=0}^{3} \left[\grave{A}_\alpha, \grave{A}_\beta \right]^a d\xi^\alpha \wedge d\xi^\beta \right)$

$= 2 \sum_{a=1}^{m} G_2 \left(\mathcal{F}_A^a, \sum_{\{\alpha,\beta\}=0}^{3} \left[\grave{A}_\alpha, \grave{A}_\beta \right]^a d\xi^\alpha \wedge d\xi^\beta \right)$

$= 2 \sum_{a=1}^{m} \sum_{\{\alpha\beta\}} \mathcal{F}_A^{a\alpha\beta} \left[\grave{A}_\alpha, \grave{A}_\beta \right]^a$

$= 2 \sum_{\{\alpha\beta\}} \left\langle \mathcal{F}_A^{\alpha\beta}, \left[\grave{A}_\alpha, \grave{A}_\beta \right] \right\rangle_{T_1 U}$

[2] In the usual jargon "on shell" means "when the equations for equilibrium are met".

$= 2\sum_{\{\alpha\beta\}} \left\langle \left[\mathcal{F}_A^{\alpha\beta}, \grave{A}_\alpha \right], \grave{A}_\beta \right\rangle_{T_1U}$
$= -2\sum_{\{\alpha\beta\}} \left\langle \left[\mathcal{F}_A^{\beta\alpha}, \grave{A}_\alpha \right], \grave{A}_\beta \right\rangle_{T_1U}$
$= -\sum_{\alpha\beta} \left\langle \left[\mathcal{F}_A^{\beta\alpha}, \grave{A}_\alpha \right], \grave{A}_\beta \right\rangle_{T_1U}$
$= -\sum_{\beta} \left\langle \sum_\alpha \left[\mathcal{F}_A^{\beta\alpha}, \grave{A}_\alpha \right], \grave{A}_\beta \right\rangle_{T_1U}$
$= -\sum_{\beta} \left\langle \phi_{A\beta}, \grave{A}_\beta \right\rangle_{T_1U}$
$= -\grave{\mathbf{A}} (\phi_A)$

$$\langle \mathcal{F}_A, \mathcal{F}_A \rangle_U = -\grave{\mathbf{A}} (\phi_A) \tag{7.34}$$

For the EM field the currents are null.
$\langle \mathcal{F}_{EM}, \mathcal{F}_{EM} \rangle = \sum_{\{\alpha\beta\}} \mathcal{F}_{EM}^{\alpha\beta} \mathcal{F}_{EM\alpha\beta} = G_2 (\mathcal{F}_{EM}, \mathcal{F}_{EM}) = G_2 \left(\mathcal{F}_{EM}, d\grave{A} \right)$
$= G_1 \left(\delta\mathcal{F}_{EM}, \grave{A} \right) = \grave{\mathbf{A}}_{EM} (\delta\mathcal{F}_{EM})$

$$\langle \mathcal{F}_{EM}, \mathcal{F}_{EM} \rangle_{EM} = \grave{\mathbf{A}}_{EM} (\delta\mathcal{F}_{EM}) \tag{7.35}$$

So that on shell :
$L_{Fields} =$
$\sum_{\alpha\beta} \left\{ C_G \left(\sum_{a=1}^3 \mathcal{F}_{G\alpha\beta}^a \mathcal{F}_G^{a\alpha\beta} - \sum_{a=4}^6 \mathcal{F}_{G\alpha\beta}^a \mathcal{F}_G^{a\alpha\beta} \right) + C_A \sum_{a=1}^m \mathcal{F}_{A\alpha\beta}^a \mathcal{F}_A^{a\alpha\beta} + C_{EM} \mathcal{F}_{EM\alpha\beta} \mathcal{F}_{EM}^{\alpha\beta} \right\}$
$= 8C_G \langle \mathcal{F}_G, \mathcal{F}_G \rangle_G + 2C_A \langle \mathcal{F}_A, \mathcal{F}_A \rangle_U + C_{EM} \left\langle \mathcal{F}_{EM\alpha\beta}, \mathcal{F}_{EM}^{\alpha\beta} \right\rangle$
$= -8C_G \mathbf{G} (\phi_G) - 2C_A \grave{\mathbf{A}} (\phi_A) + C_{EM} \grave{\mathbf{A}}_{EM} (\delta\mathcal{F}_{EM})$

Energy of the system on shell

On shell :
$J_G = \phi_G; J_A = \phi_A; \frac{1}{2}\delta\mathcal{F}_{EM} = -J_{EM}$
$L_{Fields} = -8C_G \mathbf{G} (J_G) - 2C_A \grave{\mathbf{A}} (J_A) - 2C_{EM} \grave{\mathbf{A}}_{EM} (J_{EM})$
We have seen previously that the energy of the particles can be expressed as :
$L_{Particles} = C_I \frac{1}{i} \frac{1}{M_p} \langle \psi, \nabla_V \psi \rangle = C_I \frac{dK}{dt} + 32C_G \mathbf{G} (J_G) + 8C_A \grave{\mathbf{A}} (J_A) + 8C_{EM} \grave{\mathbf{A}}_{EM} (J_{EM})$
And on shell : $\langle \psi, \nabla_V \psi \rangle = 0$
Then, on shell :

$$L_{System} = L_{Fields} = -8C_G \mathbf{G} (J_G) - 2C_A \grave{\mathbf{A}} (J_A) - 2C_{EM} \grave{\mathbf{A}}_{EM} (J_{EM}) = -\frac{1}{4} C_I \frac{dK}{dt} \tag{7.36}$$

This holds at each point for a continuous distribution of particles or for individual particles. This is equivalent to say that the variation of kinetic energy of the particles is equal to the variation of energy of the fields. In the vacuum, at equilibrium, $L_{System} = L_{Fields} = 0$.

For a perfect gas the internal energy is proportional to the kinetic energy of the molecules, and to the temperature (in °Kelvin). So we can say that L_{System} is proportional to the variation of the temperature, and that at equilibrium the temperature is constant.

7.5.2 Energy-momentum tensor

Energy momentum tensor with the perturbative lagrangian

$T = \sum_{\alpha\beta} \{ \sum_{ij} \frac{\partial L}{\partial \nabla_\alpha \psi^{ij}} \partial_\beta \psi^{ij} + \sum_{a,\gamma} \frac{\partial L}{\partial \partial_\alpha \grave{A}_\gamma^a} \partial_\beta \grave{A}_\gamma^a + \sum_{a\gamma} \frac{\partial L}{\partial \partial_\alpha G_\gamma^a} \partial_\beta G_\gamma^a \} \partial\xi_\alpha \otimes d\xi^\beta$
With the perturbative lagrangian in a model of the first type :

Particles :

$\frac{\partial L}{\partial \partial_\alpha r_a} = C_I \frac{1}{i} \mu \frac{1}{M_p} V^\alpha \left\langle \psi_0, \left[\gamma C \left(\left(\sigma^{-1} \cdot \frac{\partial \sigma}{\partial r_a} \right) \right) \right] [\psi_0] \right\rangle$

$\frac{\partial L}{\partial \partial_\alpha w_a} = C_I \frac{1}{i} \mu \frac{1}{M_p} V^\alpha \left\langle \psi_0, \left[\gamma C \left(\left(\sigma^{-1} \cdot \frac{\partial \sigma}{\partial w_a} \right) \right) \right] [\psi_0] \right\rangle$

$\sum_a \frac{\partial L}{\partial \partial_\alpha r_a} \partial_\beta r_a + \frac{\partial L}{\partial \partial_\alpha w_a} \partial_\beta w_a$

$= C_I \frac{1}{i} \mu \frac{1}{M} V^\alpha \left\langle \psi_0, \left[\gamma C \left(\left(\sigma^{-1} \cdot \sum_a \left(\frac{\partial \sigma}{\partial r_a} \partial_\beta r_a + \frac{\partial \sigma}{\partial w_a} \partial_\beta w_a \right) \right) \right) \right] [\psi_0] \right\rangle$

$= C_I \frac{1}{i} \mu \frac{1}{M} V^\alpha \left\langle \psi_0, \left[\gamma C \left(\left(\sigma^{-1} \cdot \partial_\beta \sigma \right) \right) \right] [\psi_0] \right\rangle$

$= C_I \frac{1}{i} \mu \frac{1}{M} V^\alpha \left(-i \frac{M^2}{2} k_0^t \operatorname{Re} D(-Z) \partial_\beta Z \right)$

$= -C_I \mu \frac{M}{2} V^\alpha k_0^t \operatorname{Re} D(-Z) \partial_\beta Z$

Fields:

$\frac{\partial L}{\partial \partial_\alpha G^a_{r\gamma}} = -4 C_G \mathcal{F}_r^{a\gamma\alpha}$

$\frac{dL}{d \partial_\alpha G^a_{w\gamma}} = 4 C_G \mathcal{F}_w^{a\gamma\alpha}$

$\frac{dL}{d \partial_\alpha \grave{A}^a_\gamma} = -4 C_A \mathcal{F}_A^{a\gamma\alpha}$

$\sum_{\gamma=0}^3 \sum_{a=1}^3 \frac{\partial L}{\partial \partial_\alpha G^a_{r\gamma}} \partial_\beta G^a_{r\gamma} + \frac{dL}{d \partial_\alpha G^a_{w\gamma}} \partial_\beta G^a_{w\gamma} = 4 \sum_{\gamma=0}^3 \sum_{a=1}^3 C_G \left(-\mathcal{F}_r^{a\gamma\alpha} \partial_\beta G^a_{r\gamma} + \mathcal{F}_w^{a\gamma\alpha} \partial_\beta G^a_{w\gamma} \right)$

$= -16 C_G \sum_{\gamma=0}^3 \left\langle \mathcal{F}_G^{\gamma\alpha}, \partial_\beta G_\gamma \right\rangle_{Cl}$

$\sum_{\gamma=0}^3 \sum_{a=1}^m \frac{dL}{d \partial_\alpha \grave{A}^a_\gamma} = -4 C_A \mathcal{F}_A^{a\gamma\alpha} \partial_\beta \grave{A}^a_\gamma = -4 C_A \sum_{\gamma=0}^3 \left\langle \mathcal{F}_A^{\gamma\alpha}, \partial_\beta \grave{A}_\gamma \right\rangle_{T_1 U}$

Total :

$T = -\sum_{\alpha\beta} \{ C_I \mu \frac{M^2}{2} V^\alpha k_0^t \operatorname{Re} D(-Z) \partial_\beta Z$

$+4 \sum_{\gamma=0}^3 \left(4 C_G \left\langle \mathcal{F}_G^{\gamma\alpha}, \partial_\beta G_\gamma \right\rangle_{Cl} + C_A \left\langle \mathcal{F}_A^{\gamma\alpha}, \partial_\beta \grave{A}_\gamma \right\rangle_{T_1 U} \right) \} \partial \xi_\alpha \otimes d\xi^\beta$

which can be written :

$$T = -\frac{1}{2} C_I \mu M V \otimes k_0^t \operatorname{Re} D(-Z) \, dZ + 4 \sum_{\alpha\beta\gamma} \left(4 C_G \left\langle \mathcal{F}_G^{\alpha\gamma}, \partial_\beta G_\gamma \right\rangle_{Cl} + C_A \left\langle \mathcal{F}_A^{\alpha\gamma}, \partial_\beta \grave{A}_\gamma \right\rangle_{T_1 U} \right) \partial \xi_\alpha \otimes d\xi^\beta \tag{7.37}$$

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.

Trace

The trace of the tensor is :

$Tr(T) = -\frac{1}{2} C_I \mu M k_0^t \operatorname{Re} D(-Z) \frac{dZ}{dt} + 4 \sum_{\alpha,\gamma=0}^3 \left(4 C_G \left\langle \mathcal{F}_G^{\alpha\gamma}, \partial_\alpha G_\gamma \right\rangle_{Cl} + C_A \left\langle \mathcal{F}_A^{\alpha\gamma}, \partial_\alpha \grave{A}_\gamma \right\rangle_{T_1 U} \right)$

$-\frac{1}{2} C_I \mu M k_0^t \operatorname{Re} D(-Z) \frac{dZ}{dt} = C_I \frac{dK}{dt}$

$\sum_{\alpha,\gamma=0}^3 \left\langle \mathcal{F}_G^{\alpha\gamma}, \partial_\alpha G_\gamma \right\rangle_{Cl} = \sum_{\{\alpha,\gamma\}=0}^3 \left\langle \mathcal{F}_G^{\alpha\gamma}, \partial_\alpha G_\gamma - \partial_\gamma G_\alpha \right\rangle_{Cl} = \sum_{\{\alpha,\gamma\}=0}^3 \left\langle \mathcal{F}_G^{\alpha\gamma}, \mathcal{F}_{G\alpha\gamma} - 2 [G_\alpha, G_\gamma] \right\rangle_{Cl}$

$= \sum_{\{\alpha,\gamma\}=0}^3 \left\langle \mathcal{F}_G^{\alpha\gamma}, \mathcal{F}_{G\alpha\gamma} \right\rangle_{Cl} + 2 \sum_{\{\alpha,\gamma\}=0}^3 \left\langle \mathcal{F}_G^{\alpha\gamma}, [G_\gamma, G_\alpha] \right\rangle_{Cl}$

$= \sum_{\{\alpha,\gamma\}=0}^3 \left\langle \mathcal{F}_G^{\alpha\gamma}, \mathcal{F}_{G\alpha\gamma} \right\rangle_{Cl} + 2 \sum_{\{\alpha,\gamma\}=0}^3 \left\langle \left[\mathcal{F}_G^{\alpha\gamma}, G_\gamma \right], G_\alpha \right\rangle_{Cl}$

$= \left\langle \mathcal{F}_G, \mathcal{F}_G \right\rangle_G + 2 \sum_{\alpha=0}^3 \left\langle \phi_\alpha, G_\alpha \right\rangle_{Cl}$

$= \left\langle \mathcal{F}_G, \mathcal{F}_G \right\rangle_G + 2 G(\phi_G)$

$Tr(T) = C_I \frac{dK}{dt} + 16 C_G \left\langle \mathcal{F}_G, \mathcal{F}_G \right\rangle_G + 4 C_A \left\langle \mathcal{F}_A, \mathcal{F}_A \right\rangle_A + 32 G(\phi_G) + 8 \grave{A}(\phi_A)$

$C_I \frac{1}{i} \frac{1}{M_p} \left\langle \psi, \nabla_V \psi \right\rangle = C_I \frac{dK}{dt} + 8 \left(4 C_G \mathbf{G}(J_G) + C_A \grave{\mathbf{A}}(J_A) \right)$

$Tr(T) = C_I \frac{1}{i} \frac{1}{M_p} \left\langle \psi, \nabla_V \psi \right\rangle + 16 C_G \left\langle \mathcal{F}_G, \mathcal{F}_G \right\rangle_G + 4 C_A \left\langle \mathcal{F}_A, \mathcal{F}_A \right\rangle_A$

$Tr(T) = L_{Particles} + 4L_{Fields}$
So it represents the energy exchanged in a transformation of the system.

Momentum of the fields

The Energy Momentum tensor related to the fields is :
$T_F = 4\sum_{\alpha\beta}\sum_{\gamma=0}^{3}\left(4C_G\left\langle\mathcal{F}_G^{\alpha\gamma},\partial_\beta G_\gamma\right\rangle_{Cl}+C_A\left\langle\mathcal{F}_A^{\alpha\gamma},\partial_\beta \grave{A}_\gamma\right\rangle_{T_1U}\right)\partial\xi_\alpha\otimes d\xi^\beta=\sum_{\alpha\beta}T_{F\beta}^{\alpha}\partial\xi_\alpha\otimes d\xi^\beta$
which can be written :
$T_F =$
$Tr(T_F)\sum_{\alpha=0}^{3}\partial\xi_\alpha\otimes d\xi^\alpha+\sum_{\alpha\beta}\left(T_{F\beta}^{\alpha}+T_{F\alpha}^{\beta}-\delta_\alpha^\beta Tr(T_F)\right)\partial\xi_\alpha\otimes d\xi^\beta+\sum_{\alpha\beta}\left(T_{F\beta}^{\alpha}-T_{F\alpha}^{\beta}\right)\partial\xi_\alpha\otimes d\xi^\beta$
In the vacuum :
$Tr(T_F) = 4L_{Fields}$
In a motion along the direction $\partial\xi_\beta$ the value of the potential change as : $\delta\grave{A}_\gamma=\partial_\beta\grave{A}_\gamma$ and the system reacts by a change of energy by $Tr(T)$, a transversal force $\sum_{\alpha=0}^{3}\left(T_\beta^\alpha+T_\alpha^\beta-\delta_\alpha^\beta Tr(T)\right)\partial\xi_\alpha$ and a torque : $\sum_{\alpha\beta}\left(T_\beta^\alpha-T_\alpha^\beta\right)\partial\xi_\alpha$.

These forces are present everywhere, and in presence of particles the equilibrium is reached by an adjustment of the momentum of the particle $-\frac{1}{2}C_I\mu\epsilon M_p V\otimes k_0^t\operatorname{Re}D(-Z)\,dZ$ such that $\delta T=0$.

This mechanism, which is responsible for the "radiation wind", imparts a momentum to the fields.

7.5.3 Tetrad equation

The tetrad equation reads :
$\forall\alpha,\beta=0...3: C_I\frac{1}{i}\frac{1}{M}V^\alpha\langle\psi,\nabla_\beta\psi\rangle=4[X]_\beta^\alpha-2\delta_\beta^\alpha Tr[X]$
with : $[X]_\beta^\alpha=-\sum_{\gamma=0}^{3}\{4C_G\left\langle\mathcal{F}_G^{\alpha\gamma},\mathcal{F}_{G\beta\gamma}\right\rangle_{Cl}+C_A\left\langle\mathcal{F}_A^{\alpha\gamma},\mathcal{F}_{A\beta\gamma}\right\rangle_{T_1U}\}$
It expresses the balance of energy in the system, between all the elements of the system.

Terms related to the particles

The equation is computed by varying the tetrad P_i^α, and this is equivalent to a variation $\delta\xi_\alpha$ with respect to the fixed tetrad of the observer : $\delta\xi_\alpha=\sum_{j=0}^{3}P_\alpha^{\prime j}\left(\delta P_i^\alpha\right)\varepsilon_j$

In the equation the terms related to the particles are :
$\alpha,\beta=0...3: C_I\frac{1}{i}\frac{1}{M}V^\alpha\langle\psi,\nabla_\beta\psi\rangle=V^\alpha\frac{\delta\mathcal{E}}{\delta\xi_\beta}$
with the variational derivative of the energy of the particles in the direction $\partial\xi_\beta$ of the chart :
$\frac{\delta\mathcal{E}}{\delta\xi_\beta}=C_I\frac{1}{i}\frac{1}{M}\langle\psi,\nabla_\beta\psi\rangle=-\frac{1}{2}C_I M\left\{k_0^t\operatorname{Re}\left(\nabla_\beta^G\sigma\right)+k_c^t\left[\grave{A}_\beta\right]\right\}$
So it implies the variation of the trajectories of the particles. The equation must be understood "in the meaning of distributions" and for particles we must consider the matter fields to which they belong.

The quantity
$\sum_{\alpha,\beta=0}^{3}V^\alpha\frac{\delta\mathcal{E}}{\delta\xi_\beta}\partial\xi_\alpha\otimes d\xi^\beta=\left(\sum_{\alpha=0}^{3}V^\alpha\partial\xi_\alpha\right)\otimes\left(\sum_{\beta=0}^{3}\frac{\delta\mathcal{E}}{\delta\xi_\beta}d\xi^\beta\right)\in TM\otimes TM^*$
is a tensor. Its trace is null at equilibrium : $\langle\psi,\nabla_V\psi\rangle=\sum_{\beta=0}^{3}V^\beta\langle\psi,\nabla_\beta\psi\rangle=0$
We will denote, with $\xi^0=ct$:

$$[\delta\mathcal{E}]=C_I\frac{1}{i}\frac{1}{M}\mu\left[\langle\psi,\nabla_\beta\psi\rangle\right]_{\beta=0..3}=\left[\begin{array}{cccc}\frac{\delta\mathcal{E}}{c\delta t}&\frac{\delta\mathcal{E}}{\delta\xi_1}&\frac{\delta\mathcal{E}}{\delta\xi_2}&\frac{\delta\mathcal{E}}{\delta\xi_3}\end{array}\right]=\left[\begin{array}{cc}\frac{1}{c}\frac{\delta\mathcal{E}}{\delta t}&\frac{\delta\mathcal{E}}{\delta x}\end{array}\right] \quad (7.38)$$

and as usual the velocity of the particle : $V=\left[V^\beta\right]_{4\times1}=\left[\begin{array}{c}c\\v\end{array}\right]$

And : $\langle \psi, \nabla_V \psi \rangle = 0 \Rightarrow$

$$\frac{\delta \mathcal{E}}{\delta t} + \left[\frac{\delta \mathcal{E}}{\delta x} \right] [v] = 0 \tag{7.39}$$

The tetrad equation reads :
$\forall \alpha, \beta = 0...3 : C_I \frac{1}{i} \frac{1}{M} V^\alpha \langle \psi, \nabla_\beta \psi \rangle = 4 [X]^\alpha_\beta - 2\delta^\alpha_\beta Tr [X] = V^\alpha [\delta \mathcal{E}]_\beta$
$[X]^\alpha_\beta - \frac{1}{2}\delta^\alpha_\beta Tr [X] = \frac{1}{4} V^\alpha [\delta \mathcal{E}]_\beta$

Terms related to the fields

The terms $\sum_\gamma \langle \mathcal{F}_G^{\alpha\gamma}, \mathcal{F}_{G\beta\gamma} \rangle_{Cl}, \sum_\gamma \langle \mathcal{F}_A^{\alpha\gamma}, \mathcal{F}_{A\beta\gamma} \rangle_{T_1 U}$ can be expressed in a form more appropriate to the computations.

$\sum_{\gamma=0}^3 \langle \mathcal{F}_G^{\alpha\gamma}, \mathcal{F}_{G\beta\gamma} \rangle_{Cl} = \frac{1}{4} \sum_{\gamma=0}^3 \sum_{a=1}^3 \mathcal{F}_G^{a,\alpha\gamma} \mathcal{F}^a_{G\beta\gamma} - \mathcal{F}_G^{a+3,\alpha\gamma} \mathcal{F}^{a+3}_{G\beta\gamma}$
$= \frac{1}{4} \sum_{\gamma=0}^3 \sum_{a=1}^3 [\mathcal{F}_G^{*a}]^\alpha_\gamma [\mathcal{F}_G^a]^\beta_\gamma - [\mathcal{F}_G^{*a+3}]^\alpha_\gamma [\mathcal{F}_G^{a+3}]^\beta_\gamma$
$= -\frac{1}{4} \sum_{\gamma=0}^3 \sum_{a=1}^3 [\mathcal{F}_G^{*a}]^\alpha_\gamma [\mathcal{F}_G^a]^\gamma_\beta - [\mathcal{F}_G^{*a+3}]^\alpha_\gamma [\mathcal{F}_G^{a+3}]^\gamma_\beta$
$= -\frac{1}{4} \sum_{a=1}^3 \left\{ [\mathcal{F}_G^{*a}] [\mathcal{F}_G^a] - [\mathcal{F}_G^{*a+3}] [\mathcal{F}_G^{a+3}] \right\}^\alpha_\beta$
With the complex format :
$\sum_{\gamma=0}^3 \langle \mathcal{F}_G^{\alpha\gamma}, \mathcal{F}_{G\beta\gamma} \rangle_{Cl} = -\frac{1}{4} \operatorname{Re} \sum_{a=1}^3 \{ [\mathcal{F}_G^{*a}] [\mathcal{F}_G^a] \}^\alpha_\beta$
Similarly :
$\sum_{\gamma=0}^3 \langle \mathcal{F}_A^{\alpha\gamma}, \mathcal{F}_{A\beta\gamma} \rangle_{T_1 U} = \sum_{\gamma=0}^3 \sum_{a=1}^m \mathcal{F}_A^{a,\alpha\gamma} \mathcal{F}^a_{A\beta\gamma} = \sum_{\gamma=0}^3 \sum_{a=1}^m [\mathcal{F}_A^{*a}]^\alpha_\gamma [\mathcal{F}_A^a]^\beta_\gamma$
$= -\sum_{a=1}^m \{ [\mathcal{F}_A^{*a}] [\mathcal{F}_A^a] \}^\alpha_\beta$
$[X]^\alpha_\beta = -\sum_{\gamma=0}^3 \{ 4C_G \langle \mathcal{F}_G^{\alpha\gamma}, \mathcal{F}_{G\beta\gamma} \rangle_{Cl} + C_A \langle \mathcal{F}_A^{\alpha\gamma}, \mathcal{F}_{A\beta\gamma} \rangle_{T_1 U} \}$
$= -\{ -4C_G \frac{1}{4} \operatorname{Re} \sum_{a=1}^3 \{ [\mathcal{F}_G^{*a}] [\mathcal{F}_G^a] \}^\alpha_\beta - C_A \sum_{a=1}^m \{ [\mathcal{F}_A^{*a}] [\mathcal{F}_A^a] \}^\alpha_\beta \}$
$= C_G \operatorname{Re} \sum_{a=1}^3 \{ [\mathcal{F}_G^{*a}] [\mathcal{F}_G^a] \}^\alpha_\beta + C_A \sum_{a=1}^m \{ [\mathcal{F}_A^{*a}] [\mathcal{F}_A^a] \}^\alpha_\beta \}$
$[X] = C_G \operatorname{Re} \sum_{a=1}^3 [\mathcal{F}_G^{*a}] [\mathcal{F}_G^a] + C_A \sum_{a=1}^m [\mathcal{F}_A^{*a}] [\mathcal{F}_A^a]$
To compute the summations with respect to the index a it is useful to go through the Hodge dual.
$[\mathcal{F}^*] = -[*\mathcal{F}] \det P$
$[X] = -\left\{ C_G \operatorname{Re} \sum_{a=1}^3 [*\mathcal{F}_G^a] [\mathcal{F}_G^a] + C_A \sum_{a=1}^m [*\mathcal{F}_A^a] [\mathcal{F}_A^a] \right\} \det P$
$[Y^a] = [*\mathcal{F}^a] [\mathcal{F}^a]$ is a 4×4 matrix :
$[*\mathcal{F}^a] [\mathcal{F}^a] = \begin{bmatrix} 0 & [*\mathcal{F}^r]^a \\ -([*\mathcal{F}^r]^a)^t & j([*\mathcal{F}^w]^a) \end{bmatrix} \begin{bmatrix} 0 & [\mathcal{F}^w]^a \\ -([\mathcal{F}^w]^a)^t & [j([\mathcal{F}^r]^a)] \end{bmatrix}$
$[Y] = \sum_a [Y^a] = \sum_a [\mathcal{F}^a] [*\mathcal{F}^a]$
$\alpha, \beta = 1, 2, 3$:
$[Y]^0_0 = -\sum_a [*\mathcal{F}^r]^a ([\mathcal{F}^w]^a)^t$
$[Y]^0_\beta = \{ \sum_a [*\mathcal{F}^r]^a [j([\mathcal{F}^r]^a)] \}_\beta$
$[Y]^\alpha_0 = -\left\{ \sum_a j([*\mathcal{F}^w]^a) ([\mathcal{F}^w]^a)^t \right\}^\alpha$
$[Y]^\alpha_\beta = \sum_a \left\{ -([*\mathcal{F}^r]^a)^t [\mathcal{F}^w]^a + ([\mathcal{F}^r]^a)^t ([*\mathcal{F}^w]^a) \right\}^\alpha_\beta - \delta^\alpha_\beta ([\mathcal{F}^r]^a) ([*\mathcal{F}^w]^a)^t$
where the index a runs from $a = 1, 2, 3$ for the gravitational field, and $a = 1...m$ for the other fields.
Then the computation of each element of $[Y]$ is straightforward.
$[Y]^0_0 = -Tr \left([*\mathcal{F}^r] [\mathcal{F}^w]^t \right)$
$[Y]^0_1 = \left\{ [\mathcal{F}^r]^t [*\mathcal{F}^r] \right\}^3_2 - \left\{ [\mathcal{F}^r]^t [*\mathcal{F}^r] \right\}^2_3$
$[Y]^0_2 = -\left\{ [\mathcal{F}^r]^t [*\mathcal{F}^r] \right\}^3_1 + \left\{ [\mathcal{F}^r]^t [*\mathcal{F}^r] \right\}^1_3$

$[Y]^0_3 = \left\{[\mathcal{F}^r]^t [*\mathcal{F}^r]\right\}^2_1 - \left\{[\mathcal{F}^r]^t [*\mathcal{F}^r]\right\}^1_2$
$[Y]^1_0 = \left\{[\mathcal{F}^w]^t [*\mathcal{F}^w]\right\}^2_3 - \left\{[\mathcal{F}^w]^t [*\mathcal{F}^w]\right\}^3_2$
$[Y]^2_0 = -\left\{[\mathcal{F}^w]^t [*\mathcal{F}^w]\right\}^1_3 + \left\{[\mathcal{F}^w]^t [*\mathcal{F}^w]\right\}^3_1$
$[Y]^3_0 = \left\{[\mathcal{F}^w]^t [*\mathcal{F}^w]\right\}^1_2 - \left\{[\mathcal{F}^w]^t [*\mathcal{F}^w]\right\}^2_1$
$\alpha, \beta = 1,2,3 : [Y]^\alpha_\beta = -\left\{[*\mathcal{F}^r]^t [\mathcal{F}^w]\right\}^\alpha_\beta + \left\{[\mathcal{F}^r]^t [*\mathcal{F}^w]\right\}^\alpha_\beta - \delta^\alpha_\beta Tr\left([\mathcal{F}^r]^t [*\mathcal{F}^w]\right)$
and using
$[*\mathcal{F}^r] = [\mathcal{F}^w][g_3]^{-1} \det Q'$
$[*\mathcal{F}^w] = -[\mathcal{F}^r][g_3] \det Q$
$[Y]^0_0 = -Tr\left([\mathcal{F}^w][g_3]^{-1}[\mathcal{F}^w]^t\right) \det Q'$
$[Y]^0_1 = \left\{\left\{[\mathcal{F}^r]^t [\mathcal{F}^w][g_3]^{-1}\right\}^3_2 - \left\{[\mathcal{F}^r]^t [\mathcal{F}^w][g_3]^{-1}\right\}^2_3\right\} \det Q'$
$[Y]^0_2 = \left\{-\left\{[\mathcal{F}^r]^t [\mathcal{F}^w][g_3]^{-1}\right\}^3_1 + \left\{[\mathcal{F}^r]^t [\mathcal{F}^w][g_3]^{-1}\right\}^1_3\right\} \det Q'$
$[Y]^0_3 = \left\{\left\{[\mathcal{F}^r]^t [\mathcal{F}^w][g_3]^{-1}\right\}^2_1 - \left\{[\mathcal{F}^r]^t [\mathcal{F}^w][g_3]^{-1}\right\}^1_2\right\} \det Q'$
$[Y]^1_0 = \left\{-\left\{[\mathcal{F}^w]^t [\mathcal{F}^r][g_3]\right\}^2_3 + \left\{[\mathcal{F}^w]^t [\mathcal{F}^r][g_3]\right\}^3_2\right\} \det Q$
$[Y]^2_0 = \left\{\left\{[\mathcal{F}^w]^t [\mathcal{F}^r][g_3]\right\}^1_3 - \left\{[\mathcal{F}^w]^t [\mathcal{F}^r][g_3]\right\}^3_1\right\} \det Q$
$[Y]^3_0 = \left\{-\left\{[\mathcal{F}^w]^t [\mathcal{F}^r][g_3]\right\}^1_2 + \left\{[\mathcal{F}^w]^t [\mathcal{F}^r][g_3]\right\}^2_1\right\} \det Q$
$\alpha, \beta = 1,2,3 :$
$[Y]^\alpha_\beta = -\left\{[g_3]^{-1}[\mathcal{F}^w]^t [\mathcal{F}^w] \det Q'\right\}^\alpha_\beta - \left\{[\mathcal{F}^r]^t [\mathcal{F}^r][g_3] \det Q\right\}^\alpha_\beta + \delta^\alpha_\beta Tr\left([\mathcal{F}^r]^t [\mathcal{F}^r][g_3] \det Q\right)$
The results are conveniently expressed with :

$$\begin{aligned} [Y_{RR}] &= C_G \operatorname{Re}[\mathcal{F}^r_G]^t [\mathcal{F}^r_G] + C_A [\mathcal{F}^r_A]^t [\mathcal{F}^r_A] \\ [Y_{WW}] &= C_G \operatorname{Re}\left([\mathcal{F}^w_G]^t [\mathcal{F}^w_G]\right) + C_A [\mathcal{F}^w_A]^t [\mathcal{F}^w_A] \\ [Y_{RW}] &= C_G \operatorname{Re}[\mathcal{F}^r_G]^t [\mathcal{F}^w_G] + C_A [\mathcal{F}^r_A]^t [\mathcal{F}^w_A] \end{aligned} \tag{7.40}$$

$[Y_{WR}] = C_G \operatorname{Re}\left([\mathcal{F}^w_G]^t [\mathcal{F}^r_G]\right) + C_A [\mathcal{F}^w_A]^t [\mathcal{F}^r_A] = [Y_{RW}]^t$
For the EM field :
$[\mathcal{F}^r_A]^t [\mathcal{F}^r_A] = [Q'][B][B]^t [Q']^t \det [g_3]^{-1}$
$[\mathcal{F}^w_A]^t [\mathcal{F}^w_A] = [Q]^t [E][E]^t [Q]$
$[\mathcal{F}^r_A]^t [\mathcal{F}^w_A] = [Q'][B][E]^t [Q] \det Q$
$[X]^0_0 = Tr\left([Y_{WW}][g_3]^{-1}\right)$
$[X]^0_1 = -\left\{[Y_{RW}][g_3]^{-1}\right\}^3_2 + \left\{[Y_{RW}][g_3]^{-1}\right\}^2_3$
$[X]^0_2 = \left\{[Y_{RW}][g_3]^{-1}\right\}^3_1 - \left\{[Y_{RW}][g_3]^{-1}\right\}^1_3$
$[X]^0_3 = -\left\{[Y_{RW}][g_3]^{-1}\right\}^2_1 + \left\{[Y_{RW}][g_3]^{-1}\right\}^1_2$
$[X]^1_0 = \left\{\left\{[Y_{RW}]^t [g_3]\right\}^2_3 - \left\{[Y_{RW}]^t [g_3]\right\}^3_2\right\} \det [g_3]^{-1}$
$[X]^2_0 = \left\{-\left\{[Y_{RW}]^t [g_3]\right\}^1_3 + \left\{[Y_{RW}]^t [g_3]\right\}^3_1\right\} \det [g_3]^{-1}$

$[X]_0^3 = \left\{ \left\{ [Y_{RW}]^t [g_3] \right\}_2^1 - \left\{ [Y_{RW}]^t [g_3] \right\}_1^2 \right\} \det [g_3]^{-1}$
$\alpha, \beta = 1,2,3 :$
$[X]_\beta^\alpha = \left\{ [g_3]^{-1} [Y_{WW}] \right\}_\beta^\alpha + \{[Y_{RR}] [g_3]\}_\beta^\alpha \det [g_3]^{-1} - \delta_\beta^\alpha Tr \left([Y_{RR}] [g_3] \det [g_3]^{-1} \right)$

Equations

The tetrad equation reads :
$[X]_\beta^\alpha - \frac{1}{2} \delta_\beta^\alpha Tr [X] = \frac{1}{4} V^\alpha [\delta \mathcal{E}]_\beta$
$Tr [X] - 2Tr [X] = \frac{1}{4} \sum_{\beta=0}^3 V^\beta [\delta \mathcal{E}]_\beta = 0 \Rightarrow Tr [X] = 0$
$Tr [X] = Tr \left([Y_{WW}] [g_3]^{-1} \right) + \sum_{\beta=1}^3 \left\{ [g_3]^{-1} [Y_{WW}] \right\}_\beta^\beta + \{[Y_{RR}] [g_3]\}_\beta^\beta \det [g_3]^{-1} - Tr \left([Y_{RR}] [g_3] \det [g_3]^{-1} \right)$
$= Tr \left([Y_{WW}] [g_3]^{-1} \right) + Tr \left\{ [g_3]^{-1} [Y_{WW}] \right\} + Tr \{[Y_{RR}] [g_3]\} \det [g_3]^{-1} - 3Tr ([Y_{RR}] [g_3]) \det [g_3]^{-1}$
$= 2Tr \left([Y_{WW}] [g_3]^{-1} \right) - 2Tr \{[Y_{RR}] [g_3]\} \det [g_3]^{-1}$
$Tr [X] = 0 \Leftrightarrow Tr \left([Y_{WW}] [g_3]^{-1} \right) = Tr \{[Y_{RR}] [g_3]\} \det [g_3]^{-1}$
$[X]_\beta^\alpha = \frac{1}{4} V^\alpha [\delta \mathcal{E}]_\beta$
$Tr \{[Y_{RR}] [g_3]\} \det [g_3]^{-1} = \frac{1}{4} c \frac{\delta \mathcal{E}}{c \delta t} = -\frac{1}{4} \left[\frac{\delta \mathcal{E}}{\delta x} \right] [v] = Tr \left([Y_{WW}] [g_3]^{-1} \right)$
$\alpha, \beta = 1,2,3 :$
$\left\{ [g_3]^{-1} [Y_{WW}] \right\}_\beta^\alpha + \{[Y_{RR}] [g_3]\}_\beta^\alpha \det [g_3]^{-1} - \delta_\beta^\alpha Tr \left([Y_{RR}] [g_3] \det [g_3]^{-1} \right) = \frac{1}{4} V^\alpha [\delta \mathcal{E}]_\beta$
$[g_3]^{-1} [Y_{WW}] + [Y_{RR}] [g_3] \det [g_3]^{-1} - I_3 Tr \left([Y_{RR}] [g_3] \det [g_3]^{-1} \right) = \frac{1}{4} [v] \left[\frac{\delta \mathcal{E}}{\delta x} \right]$
$[g_3]^{-1} [Y_{WW}] + [Y_{RR}] [g_3] \det [g_3]^{-1} = \frac{1}{4} [v] \left[\frac{\delta \mathcal{E}}{\delta x} \right] - I_3 \frac{1}{4} \left[\frac{\delta \mathcal{E}}{\delta x} \right] [v] = \frac{1}{4} j \left(\frac{\delta \mathcal{E}}{\delta x} \right) j (v)$
$-\left\{ [Y_{RW}] [g_3]^{-1} \right\}_2^3 + \left\{ [Y_{RW}] [g_3]^{-1} \right\}_3^2 = \frac{1}{4} c \frac{\delta \mathcal{E}}{\delta \xi_1}$
$\left\{ [Y_{RW}] [g_3]^{-1} \right\}_1^3 - \left\{ [Y_{RW}] [g_3]^{-1} \right\}_3^1 = \frac{1}{4} c \frac{\delta \mathcal{E}}{\delta \xi_2}$
$-\left\{ [Y_{RW}] [g_3]^{-1} \right\}_1^2 + \left\{ [Y_{RW}] [g_3]^{-1} \right\}_2^1 = \frac{1}{4} c \frac{\delta \mathcal{E}}{\delta \xi_3}$
$\left\{ \left\{ [Y_{RW}]^t [g_3] \right\}_3^2 - \left\{ [Y_{RW}]^t [g_3] \right\}_2^3 \right\} \det [g_3]^{-1} = \frac{1}{4} v^1 \frac{\delta \mathcal{E}}{c \delta t}$
$\left\{ -\left\{ [Y_{RW}]^t [g_3] \right\}_3^1 + \left\{ [Y_{RW}]^t [g_3] \right\}_1^3 \right\} \det [g_3]^{-1} = \frac{1}{4} v^2 \frac{\delta \mathcal{E}}{c \delta t}$
$\left\{ \left\{ [Y_{RW}]^t [g_3] \right\}_2^1 - \left\{ [Y_{RW}]^t [g_3] \right\}_1^2 \right\} \det [g_3]^{-1} = \frac{1}{4} v^3 \frac{\delta \mathcal{E}}{c \delta t}$
The last equations can also be written :
$[Y_{RW}] [g_3]^{-1} - [g_3]^{-1} [Y_{RW}]^t = -\frac{1}{4} c j \left(\frac{\delta \mathcal{E}}{\delta x} \right)$
$\Leftrightarrow [g_3] [Y_{RW}] - [Y_{RW}]^t [g_3] = -\frac{1}{4} c [g_3] j \left(\frac{\delta \mathcal{E}}{\delta x} \right) [g_3] = -\frac{1}{4} c j \left(\left[\frac{\delta \mathcal{E}}{\delta x} \right] [g_3]^{-1} \right) \det [g_3]$
$[Y_{RW}]^t [g_3] - [g_3] [Y_{RW}] = -\frac{1}{4} \frac{\delta \mathcal{E}}{c \delta t} j (v)$
$[g_3] [Y_{RW}] - [Y_{RW}]^t [g_3] = -\frac{1}{4} c j \left(\left[\frac{\delta \mathcal{E}}{\delta x} \right] [g_3]^{-1} \right) \det [g_3] = \frac{1}{4} \frac{\delta \mathcal{E}}{c \delta t} j (v)$
So we have :
$Tr \{[Y_{RR}] [g_3]\} \det [g_3]^{-1} = Tr \left([Y_{WW}] [g_3]^{-1} \right) = \frac{1}{4} c \frac{\delta \mathcal{E}}{c \delta t} = -\frac{1}{4} \left[\frac{\delta \mathcal{E}}{\delta x} \right] [v]$
$[g_3] [Y_{RW}] - [Y_{RW}]^t [g_3] = -\frac{1}{4} c j \left(\left[\frac{\delta \mathcal{E}}{\delta x} \right] [g_3]^{-1} \right) \det [g_3] = \frac{1}{4} \frac{\delta \mathcal{E}}{c \delta t} j (v)$
$[g_3]^{-1} [Y_{WW}] + [Y_{RR}] [g_3] \det [g_3]^{-1} = \frac{1}{4} [v] \left[\frac{\delta \mathcal{E}}{\delta x} \right] - I_3 \frac{1}{4} \left[\frac{\delta \mathcal{E}}{\delta x} \right] [v] = \frac{1}{4} j \left(\frac{\delta \mathcal{E}}{\delta x} \right) j (v)$
But :
$-\frac{1}{4} c j \left(\left[\frac{\delta \mathcal{E}}{\delta x} \right] [g_3]^{-1} \right) \det [g_3] = \frac{1}{4} \frac{\delta \mathcal{E}}{c \delta t} j (v)$

$\Leftrightarrow \left[\frac{\delta\mathcal{E}}{\delta x}\right] = -\frac{\delta\mathcal{E}}{\delta t}\frac{1}{c^2}[v]^t[g_3]\det[g_3]^{-1}$
$\Rightarrow \left[\frac{\delta\mathcal{E}}{\delta x}\right][v] = -\frac{\delta\mathcal{E}}{\delta t}\frac{1}{c^2}[v]^t[g_3][v]\det[g_3]^{-1} = -\frac{\delta\mathcal{E}}{\delta t} = -\frac{\delta\mathcal{E}}{\delta t}\frac{\|v\|^2}{c^2}\det[g_3]^{-1}$
The only solution is $\frac{\delta\mathcal{E}}{\delta t} = 0 \Rightarrow \left[\frac{\delta\mathcal{E}}{\delta x}\right] = 0$

The tetrad equation is about the energy. The quantities $\frac{\delta\mathcal{E}}{\delta t}, \left[\frac{\delta\mathcal{E}}{\delta x}\right]$ are components of the variational derivative $\frac{\delta\mathcal{E}}{\delta\xi}$ of the energy of the particles in a variation $\delta\xi$. In a continuous process this variational derivative should be null. This can be understood in another way. We have assumed that the state of the particle can be modelled on a section $\psi \in \mathfrak{X}(Q[E \otimes F, \vartheta])$ in order to vary the trajectory. The energy of the particle reads :

$\delta E = C_I \frac{1}{M}\frac{1}{i}\sum_{\beta=0}^3 V^\beta \langle\psi, \nabla_\beta\psi\rangle = C_I \frac{1}{M}\frac{1}{i}\sum_{\beta,j=0}^3 P_j^\beta U^j \langle\psi, \nabla_\beta\psi\rangle$

$C_I \frac{1}{i}\frac{1}{M}\langle\psi, \nabla_\beta\psi\rangle$ is the variational derivative with respect to V^β, the component of V, which can vary with the tetrad, U being fixed, or with U, with a fixed tetrad. In a continuous process the integral curves U of a given section are curves of constant energy : $\langle\psi, \nabla_V\psi\rangle = 0$. So $\frac{\delta\mathcal{E}}{\delta t} = 0, \left[\frac{\delta\mathcal{E}}{\delta x}\right] = 0$ is equivalent to the condition that the particle stays on the same integral curve.

And the equations sum up to :

$$\begin{gathered} Tr\{[Y_{RR}][g_3]\} = Tr\left([Y_{WW}][g_3]^{-1}\right) = 0 \\ [g_3][Y_{RW}] = [Y_{RW}]^t[g_3] \\ [Y_{WW}] = -[g_3][Y_{RR}][g_3]\det[g_3]^{-1} \end{gathered} \tag{7.41}$$

We get 20 equations, from which the metric can be computed directly without differential equations.

For instance with the specification of the metric used before :

$$g_3 = \begin{bmatrix} \lambda_1^2 & a_3\lambda_1\lambda_2 & a_2\lambda_1\lambda_3 \\ a_3\lambda_1\lambda_2 & \lambda_2^2 & a_1\lambda_2\lambda_3 \\ a_2\lambda_1\lambda_3 & a_1\lambda_2\lambda_3 & \lambda_3^2 \end{bmatrix}$$

The equation $[g_3][Y_{RW}] = [Y_{RW}]^t[g_3]$ gives 3 linear equations :
$a_1 = \lambda_3^{-1}\lambda_2^{-1}\left\{\lambda_1^2 A_{11} + \lambda_2^2 A_{12} + \lambda_3^2 A_{13}\right\}$
$a_2 = \lambda_3^{-1}\lambda_1^{-1}\left\{\lambda_1^2 A_{21} + \lambda_2^2 A_{22} + \lambda_3^2 A_{23}\right\}$
$a_3 = \lambda_2^{-1}\lambda_1^{-1}\left\{\lambda_1^2 A_{31} + \lambda_2^2 A_{32} + \lambda_3^2 A_{33}\right\}$
where A_{pq} depends on $[Y_{RW}]$ only. Then the second equation gives λ.

The results have no simple expression, but they are explicit, the computations are straightforward and do not involve differential equations.

Remarks :

i) These equations involve only the fields and the metric : the metric is defined by the value of the fields, even if they are themselves defined through their interactions with particles. The physical universe has a finite number of symmetries and these symmetries hold also for the fields, which could be expected from their propagation curves.

ii) The equations are not independent : we have overall 20 equations for 6 parameters in $[g]$. So the 21 coefficients in the matrices $[Y_{WW}], [Y_{RR}], [Y_{WR}]$ are related. They involve all the fields on the same footing, that is, practically the EM and the gravitational fields. It implies that, even if the fields are assumed to not interact with each others, their value is not independent. This is not totally new : in the Einstein's theory of gravitation one can include the EM field in the energy-momentum tensor, however the mechanism involved is then more classic (the EM field contributes to the general equilibrium). Here we have a more direct mechanism, which opens the way to act on the gravitational field through the EM field, as a recent experiment shows (H.White and alii). Moreover the possibility to act on the metric opens also many possibilities. Of course these results should be checked, but due to their potential importance, this path deserves to be explored further.

7.6 CONCLUSION

The implementation of the Principle of Least Action provides several sets of equations, to which must be added the laws of propagation. They can be used separately, according to the problem, or together. The computation of a general solution is always complicated in GR, but, with the numerous tools which are provided, one can go much further than usual.

The continuous models, in spite of their limitations, comfort the framework which has been introduced here, and gives a profound meaning to the relationship between matter fields and force fields, both from a mathematical and physical point of view.

$$\begin{array}{ccc} \text{Matter Fields} & & \text{Force Fields} \\ \sigma : \text{Sections of } P_G & & \mathbf{G}, \grave{\mathbf{A}} : \text{ equivariant sections of } J^1P_G, J^1P_U \\ \downarrow & & \downarrow \\ \text{Motion : trajectories} & & \mathcal{F} : \text{ Propagation} \\ \downarrow & & \downarrow \\ \text{Spinors, } \psi_0 : \text{ momenta} & & \text{Potential : force, torque} \\ \downarrow & & \downarrow \\ \text{Currents } J & \longleftrightarrow & \text{Currents } \phi \end{array} \tag{7.42}$$

At first everything seems to oppose particles, moving on curves, and force fields, defined everywhere. But, in a system at equilibrium, the trajectories and the motion of particles can be represented through sections of P_G defined everywhere, meanwhile the fields propagate along Killing curves. Sections of P_G give, with a unique constant ψ_0, matter fields and momenta, principal connections are equivariant sections of a the first jet prolongation of principal bundle, their Lie derivative give the strength $\mathcal{F}$, which rules the propagation, and the potentials which define forces and torques. Particles and fields generate currents, which are defined everywhere, and are equal. And the general equilibrium is achieved through the metric, which defines the symmetries of the physical universe.

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. 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 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 on 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. But we will start with collisions, which are the basic discontinuous processes.

8.1 COLLISIONS

By a collision we mean the encounter of two (or more, which should be very unusual) particles which at some time, occupy the same location. It is “elastic” when the kinetic energy is preserved, which has a meaning for deformable solids : no energy is spent in the deformation. We will consider only particles, then an elastic collision means that the particles keep their fundamental state ψ_0 : for elementary particles there is no creation or annihilation, and for other material bodies the inertial spinors S_0 are preserved. In non elastic collisions it is necessary to involve the forces and charges of the particles, directly or through phenomenological laws.

8.1.1 Collisions in Newtonian Mechanics

Solving the problem of collision between particles is commonly said to come from the Principle of Conservation of Momentum, but this is deceptive. The key point is that, in Galilean Geometry, it is possible to define a center of mass G for any system of material points : $\left(\sum_a m_a\right)\overrightarrow{OG} = \sum_a m_a \overrightarrow{OM}_a$. Then the system is equivalent to a particle of mass $\sum_a m_a$ located at G and the sum $\overrightarrow{F}_G = \sum_a \overrightarrow{F}_a$, exercised at G, has a physical meaning. And the Law of Mechanics can be written, by derivation :

$\sum_a \frac{d\overrightarrow{p}_a}{dt} = \frac{d\overrightarrow{p}_G}{dt} = \overrightarrow{F}_G$

In the collision of two particles, the sum of the momenta : $\overrightarrow{p}_1 + \overrightarrow{p}_2$ is conserved only if $\frac{d\overrightarrow{p}_G}{dt} = \overrightarrow{F}_G = 0$. Then with

$m_1\overrightarrow{v}_1 + m_2\overrightarrow{v}_2 = m_1\overrightarrow{v'}_1 + m_2\overrightarrow{v'}_2 = (m_1+m_2)\,\overrightarrow{v}_G$

and the conservation of the kinetic energy, if the collision is elastic :

$m_1\left\|\overrightarrow{v}_1\right\|^2 + m_2\left\|\overrightarrow{v}_2\right\|^2 = m_1\left\|\overrightarrow{v'}_1\right\|^2 + m_2\left\|\overrightarrow{v'}_2\right\|^2$

we have 4 equations, for 6 unknown variables. So this is not enough to solve the problem. We need to account for a rotational momentum. The total torque on the system is : $\sum_{a=1,2}\tau_a(O) = \sum_a \tau_a(G)$. If it is null then the total rotational momentum is conserved :

$\sum_a \Gamma_a(O) = \sum_a \Gamma_a(G) = Ct$

$\overrightarrow{OM_1}\times\overrightarrow{p}_1 + \overrightarrow{OM_2}\times\overrightarrow{p}_2 = Ct$

At the point of collision : $\overrightarrow{OG}\times(m_1\overrightarrow{v}_1 + m_2\overrightarrow{v}_2) = \overrightarrow{OG}\times(m_1\overrightarrow{v}'_1 + m_2\overrightarrow{v}'_2)$, which is equivalent to say that $\overrightarrow{v'}_1, \overrightarrow{v'}_2$ are in the plane defined by $\overrightarrow{v}_1, \overrightarrow{v}_2$: we have only 4 unknown variables, and the problem is solved.

This solution is commonly extended to Special Relativity (the conservation of kinetic energy comes then from the 4th component), but this cannot be done in RG.

Using this method, it is possible to define a “collision operator” to represent elastic collisions between particles (the operator gives $\overrightarrow{v'}_1, \overrightarrow{v'}_2$ from $\overrightarrow{v}_1, \overrightarrow{v}_2$), which is then incorporated in general models based on the Principle of Least Action. The oldest are the kinetic models. They usually derive from a hydrodynamic model (similar to the continuous models) and 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).

8.1.2 Collisions in RG

Particles in RG are represented by maps $\mathbb{R} \to J^1Q[E\otimes F,\vartheta] :: (q(t), \psi(t), \delta\psi(t))$. So the location is part of the definition. The variation of momentum is then $\delta\psi(t) = \vartheta(\upsilon(X_r, X_w)\cdot\sigma, \varkappa)\psi_0$ where $\upsilon(X_r, X_w) \in T_1Spin(3,1)$ and $\upsilon(X_r, X_w) = \frac{d\sigma}{dt}\cdot\sigma^{-1}$ in a continuous motion. $\psi(t), \delta\psi(t) \in E\otimes F$, that is a fixed vector space (this is the advantage of the fiber bundle representation). The link with the physical, located, quantity is done through the gauge of the observer at $q(t)$.

The tetrad attached to the particle is such that : $e_i(t) = \mathbf{Ad}_{\sigma(t)}\varepsilon_0$ where ε_0 is a fixed vector. The relation between $\sigma(t)$ and the velocity V goes through the tetrad P of the observer at $q(t)$

$V = \frac{dq}{dt} = c\varepsilon_0 + \overrightarrow{v} = \sum_{j,\alpha=0}^{3} P_j^\alpha U^j \partial\xi_\alpha$

$U = -\frac{c}{\langle \mathbf{Ad}_\sigma\varepsilon_0, \varepsilon_0\rangle_{Cl}}\mathbf{Ad}_\sigma\varepsilon_0 = \sum_{j=0}^{3} U^j\varepsilon_j$

A spinor can be computed for a deformable solid, defined by a section $\sigma \in \mathfrak{X}(P_G)$, an inertial spinor S_0 and a density μ by :

$S(t) = \gamma C\left(\int_{\omega(t)} \sigma(m)\mu(m)\varpi_3(m)\right)S_0$

This can be extended to a matter field with a fundamental state ψ_0. In these representations the assumptions are that the particles have the same characteristics (charge and kinematic) and their trajectories do not cross.

However the representation holds, at least formally, for a collection of individual particles $p = 1...N$.

To each particle is associated a section of $\mathfrak{X}(Q[E\otimes F,\vartheta])$, and the particle is a map :

$Z_p : [0,T] \to \mathbb{C}^3 :: Z(t)$ such that : $\psi(t) = \psi(Z(t))$

The density is a function $\mu : M \to \mathbb{R}$ but actually we can consider a map valued in the set of scalar measures on M : $\mu(t)\varpi_4(m)$. Then the density is equal to 0 or 1, this is equivalent to a Dirac's function δ_p for each particle.

The aggregation is done by the measure $\sum_p \delta_p \times \varpi_4$.

$\psi(t) = \int_{\Omega(t)}\sum_{p=1}^N \psi_p(Z(t))\delta_p(p(t))\varpi_3 = \int_{\Omega(t)}\sum_{p=1}^N \psi_p(Z(t))\delta_p(\varphi_o(t, x_p(t)))\varpi_3(t)$

$\psi(t) \in E\otimes F$ and we can define :

$\psi(t) = \vartheta(\sigma(t), \varkappa)\psi_0 = \int_{\Omega(t)}\sum_{p=1}^N \vartheta(\sigma_p(t), \varkappa)\psi_{0p}\delta_p(p(t))\varpi_3$

with a fixed gauge in F (which identifies the flavor of particles). It sums up to take :

$\psi(t) = \vartheta(\sigma(t), \varkappa)\psi_0 = \sum_{p=1}^N \vartheta(\sigma_p(t), \varkappa)\psi_{0p}$

In a continuous motion :

$\frac{d}{dt}\psi(t) = \sum_{p=1}^N \vartheta\left(\frac{d}{dt}\sigma_p(t), \varkappa\right)\psi_{0p}$

$= \sum_{p=1}^N \vartheta\left(\frac{d}{dt}\sigma_p(t)\cdot\sigma_p(t)^{-1}, \varkappa\right)\psi_{0p} = \sum_{p=1}^N \vartheta(\upsilon(X_{rp}, X_{wp}), \varkappa)\psi_p = \vartheta(\upsilon(X_r, X_w), \varkappa)\psi(t)$

ψ_0 is fixed along the trajectories with $\upsilon(X_r, X_w) = \sum_{p=1}^N \upsilon(X_{rp}, X_{wp})$.

There is no specific location $q(t)$, equivalent to a center of mass, attached to the collection of particles. But we can define a momentum $\delta\psi = \vartheta(\upsilon(X_r, X_w), \varkappa)\psi(t)$ which is equal to the sum of the momentum of the particles.

The momentum of each particle changes with the actions of the fields :

$\delta\psi_p \to \delta\psi_p + \sum_{\alpha=0}^3 V_p^\alpha\vartheta(\sigma_p, \varkappa)\left([\gamma C(\mathbf{Ad}_{\sigma^{-1}}(G_\alpha(q_p(t))))]\psi_{0p} + [\psi_{0p}]\left[Ad_\varkappa\grave{A}_\alpha(q_p(t))\right]\right)$

so the momenta and the total momentum $\delta\psi$ are not conserved in the presence of fields.

The specificity of a collision is that the particles are at the same location, so the location of the center of mass is defined, moreover in the process of collision the fields are not involved : the action of the fields comes from the motion of the particles, which entails a change in the potentials. In a collision the particles exchange only kinetic momentum. So we can write :

$\sum_{p=1}^N \delta\psi_p = \sum_{p=1}^N \delta\widetilde{\psi}_p$ where $\widetilde{\psi}_p$ is the state of the particle after the collision.

Moreover, in an elastic collision, the fundamental states do not change, and :

$\widetilde{\psi}(t_0) = \vartheta(\widetilde{\sigma}(t_0), \varkappa)\psi_0 = \sum_{p=1}^N \vartheta(\widetilde{\sigma}_p(t_0), \varkappa)\psi_{0p}$

The collision is an isolated point, occurring at $m_0 = \varphi_o(t_0, x_0)$: before and after the collision the particles have a continuous motion :

$v(X_{rp}, X_{wp}) = \frac{d\sigma_p}{dt} \cdot \sigma_p^{-1}$

$v\left(\widetilde{X}_{rp}, \widetilde{X}_{wp}\right) = \frac{d\widetilde{\sigma_p}}{dt} \cdot \widetilde{\sigma_p}^{-1}$

And we have the equations :

$\sum_{p=1}^{N} \vartheta(\sigma_p(t_0), \varkappa)\psi_{0p} = \sum_{p=1}^{N} \vartheta(\widetilde{\sigma_p}(t_0), \varkappa)\psi_{0p}$

$\sum_{p=1}^{N} \vartheta\left(\frac{d}{dt}\sigma_p(t_0) \cdot \sigma_p(t_0)^{-1}, \varkappa\right)\psi_p(t_0) = \sum_{p=1}^{N} \vartheta\left(\frac{d}{dt}\widetilde{\sigma_p}(t_0) \cdot \widetilde{\sigma_p}(t_0)^{-1}, \varkappa\right)\psi_p(t_0)$

$\Leftrightarrow \sum_{p=1}^{N} \vartheta(v(X_{rp}, X_{wp}), \varkappa)\psi_p(t_0) = \sum_{p=1}^{N} \vartheta\left(v\left(\widetilde{X}_{rp}, \widetilde{X}_{wp}\right), \varkappa\right)\psi_p(t_0)$

Moreover the total kinetic energy of the particles is conserved :

$\sum_{p=1}^{N} \frac{1}{i}\frac{1}{M_p}\langle\psi_p, \delta\psi_p\rangle = \sum_{p=1}^{N} \frac{1}{i}\frac{1}{M_p}\left\langle\widetilde{\psi}_p, \delta\widetilde{\psi}_p\right\rangle$

For spinors we have a set of 17 real scalar equations for 12 unknown variables for each particle. As in Newtonian Mechanics we need an additional equation, and it comes from the conservation of the rotational momentum. For spinors :

$\delta S_R = \sum_{\alpha=0}^{3} \gamma C(v(X_r, 0))S$ is the equivalent of a change of rotational momentum or an inertial torque.

$\delta S_T = \sum_{\alpha=0}^{3} \gamma C(v(0, X_w))S$ is the equivalent of a change of translational momentum or a translational inertial force.

In the collision the conservation of the momenta is equivalent to the fact that the forces and torques exercised on the "out" particles are equal to the forces and torques exercised by the "in" particles. So we must replace the equation :

$\sum_{p=1}^{N} \vartheta(v(X_{rp}, X_{wp}), \varkappa)\psi_p(t_0) = \sum_{p=1}^{N} \vartheta\left(v\left(\widetilde{X}_{rp}, \widetilde{X}_{wp}\right), \varkappa\right)\psi_p(t_0)$

by the 2 equations :

$\sum_{p=1}^{N} \vartheta(v(X_{rp}, 0), \varkappa)\psi_p(t_0) = \sum_{p=1}^{N} \vartheta\left(v\left(\widetilde{X}_{rp}, 0\right), \varkappa\right)\psi_p(t_0)$

$\sum_{p=1}^{N} \vartheta(v(0, X_{wp}), \varkappa)\psi_p(t_0) = \sum_{p=1}^{N} \vartheta\left(v\left(0, \widetilde{X}_{wp}\right), \varkappa\right)\psi_p(t_0)$

and we have 24 real scalar equations, which solves the problem for the collision of 2 particles. One can check that then the kinetic energy : $\delta K = -\frac{M_p}{2}k_0^t \operatorname{Re} \mathbf{Ad}_{\sigma^{-1}} v(X_r, X_w)$ is conserved.

It is convenient to introduce auxiliary variables. The tetrad attached to each particle is such that : $e_i(t) = \mathbf{Ad}_{\sigma(t)}\varepsilon_0$. At m_0 :

$e_i(t_0) = \mathbf{Ad}_{\sigma(t_0)}\varepsilon_0$

$\widetilde{e}_i(t_0) = \mathbf{Ad}_{\widetilde{\sigma}(t_0)}\varepsilon_0$

and there is a fixed $s \in Spin(3,1)$ such that $\widetilde{\sigma}(t_0) = s \cdot \sigma(t_0)$

$U(t_0) = -\frac{c}{\langle \mathbf{Ad}_\sigma \varepsilon_0, \varepsilon_0 \rangle_{Cl}} \mathbf{Ad}_{\sigma(t_0)}\varepsilon_0$

$\widetilde{U}(t_0) = -\frac{c}{\langle \mathbf{Ad}_s \mathbf{Ad}_\sigma \varepsilon_0, \varepsilon_0 \rangle_{Cl}} \mathbf{Ad}_s \mathbf{Ad}_{\sigma(t_0)}\varepsilon_0 = \lambda \mathbf{Ad}_s U(t_0) = \lambda[h(s)]U(t_0)$

$[h(s)] \in SO(3,1)$

with the additional variable $\lambda = \frac{\langle \mathbf{Ad}_\sigma \varepsilon_0, \varepsilon_0 \rangle_{Cl}}{\langle \mathbf{Ad}_s \mathbf{Ad}_\sigma \varepsilon_0, \varepsilon_0 \rangle_{Cl}}$

The velocity, in the holonomic basis, is then :

$\widetilde{V}^\alpha = \sum_j P_j^\alpha \widetilde{U}^j = \lambda \sum_j P_j^\alpha [h(s)]_k^j U^k = \lambda \sum_{j,\beta} P_j^\alpha [h(s)]_k^j P_\beta'^k V^\beta$

$= \lambda \sum_{\beta=0}^{3} ([P][h(s)][P'])_\beta^\alpha V^\beta = \lambda \sum_{\beta=0}^{3} [X]_\beta^\alpha V^\beta$

8.1.3 Solution by the gravitational currents

There is another way to proceed. Let us consider 2 particles A, B colliding in a point O.

It has been assumed that the gravitational current ϕ_G is continuous. At any point in the neighborhood of O, it reflects the sum of the gravitational currents J_A, J_B associated to A and B.

Then we have :

before the collision : $\phi_G(O) = J_A + J_B$

after the collision : $\phi_G(O) = J'_A + J'_B$

The currents are defined in $T_1 Spin(3,1) \otimes TM$, they depend on 16 parameters $\upsilon(r,w), V$:

$J_A = \frac{C_I}{16C_G} \epsilon M_A \upsilon(-[A(w_A)][C(r_A)]k_{A0}, [B(w_A)][C(r_A)]k_{A0}) \otimes V_A$

and similarly for B.

Then we have :

$M_A \upsilon(-[A(w_A)][C(r_A)]k_{A0}, [B(w_A)][C(r_A)]k_{A0}) V_A^\alpha$
$+M_B \upsilon(-[A(w_B)][C(r_B)]k_{B0}, [B(w_B)][C(r_B)]k_{B0}) V_B^\alpha$
$= M_A \upsilon(-[A(w'_A)][C(r'_A)]k_{A0}, [B(w'_A)][C(r'_A)]k_{A0}) V_A'^\alpha$
$+M_B \upsilon(-[A(w'_B)][C(r'_B)]k_{B0}, [B(w'_B)][C(r'_B)]k_{B0}) V_B'^\alpha$

that is 24 equations, for the 12 variables $(r'_A, w'_A), (r'_B, w'_B)$ (V depends then on w).

It is obvious that, through the gravitational field, the motion of the particles will adjust before the collision, so that the variables are not independent. However this simple (but coming from a long way...) model shows that, actually, the process of collision is determinist. One could proceed independently to the same computation with the currents related to the other fields, but not for the EM field because $\phi_{EM} = 0$.

We have the estimate :.

$J_G \simeq \frac{C_I}{16C_G} \epsilon M \upsilon \left(-[C(r)]k_0, \left(1 + \frac{1}{2}\frac{\|\overrightarrow{v}\|^2}{c^2}\right) j\left(\frac{\overrightarrow{v}}{c}\right)[C(r)]k_0\right) \otimes V$

and for non rotating particles : $J_G \simeq \frac{C_I}{16C_G} \epsilon M \upsilon(-k_0, 0) \otimes V$

so that :

$M_A \upsilon(-k_{A0}, 0) \otimes V_A + M_B \upsilon(-k_{B0}, 0) \otimes V_B = M_A \upsilon(-k_{A0}, 0) \otimes V'_A + M_B \upsilon(-k_{B0}, 0) \otimes V'_B$
$M_A \upsilon(-k_{A0}, 0) \otimes \delta V_A + M_B \upsilon(-k_{B0}, 0) \otimes \delta V_B = 0$

and we get back the usual equation for the momentum.

Notice that, in both methods, one accounts for the possible rotations of the particles, something which is difficult to achieve even in Classic Mechanics.

8.1.4 Scattering

The typical experiment in Particles Physics is the impact of a beam of particles on a target. Discontinuous processes occur when the incoming particles interact with the atoms of the target. This is represented as a transition between a population of incoming particles in "in" states, and outgoing particles in "out" states, called a **scattering**. It is generally assumed that before and after the interactions the beam is in equilibrium, and the particles follow a continuous motion. So the mechanisms are collisions, but the number of particles is not necessarily fixed : annihilations and creations of particles can occur, so more complex models are necessary.

When the weak or strong interactions are involved, additional rules apply, empirical or 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 guide in predicting the outcome. Moreover the strong interaction and electromagnetic interaction seem to be invariant under the combined CP operation,

8.2 BOSONS

The topic of this section is discontinuity in the force fields. One of the characteristic of force fields is that they are defined everywhere and propagate, so to get an idea about the representation of a discontinuity it is useful to look at how this is done in a continuous medium (a fluid or a deformable solid).

8.2.1 Discontinuity in a continuous medium

Material medium, deformable solids or fluids, can transport a "signal" which comes from a specific motion (spin or vibration) of its molecules. Wave propagations, wave packets, bursts 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 signal is still represented by smooth maps. Discontinuities are different : the maps are no longer smooth. At the macroscopic level we have shock waves. The model of deformable solid (in the GR framework !) is actually well suited for this study.

The key property of fluids, as well as deformable solids, is that they are comprised of material points which can be identified by their location x at $t = 0$ and follow trajectories which do not cross, along a vector field V. A tetrad is attached to each material point, and the system is then represented by a section $S \in \mathfrak{X}(P_G[E, \gamma C])$. Along an integral curve of V :

$\frac{de_i}{dt} = [\upsilon(X_r, X_w), e_i]$

$\frac{dU}{dt} = \frac{U}{c} \langle [\upsilon(X_r, X_w), U], \varepsilon_0 \rangle_{Cl} + [\upsilon(X_r, X_w), U]$

with $\upsilon(X_r, X_w) = \sum_{\alpha=0}^{3} V^\alpha \partial_\alpha \sigma \cdot \sigma^{-1}$

So the material points (say molecules) can have a spinning or vibrating motion, which is continuously differentiable.

A shock wave does not disrupt (usually) the continuity of the medium, but the derivatives $\frac{dV}{dt}, \frac{de_i}{dt}$ are no longer continuous. One characteristic of a shock wave is that it propagates : this is typically the sonic "boom" which occurs when the shock propagates faster than V.

Because the model is based on material points, identified by their location $x \in \Omega_3(0)$ at $t = 0$, there is a function :

$\theta : \Omega_3(0) \to \mathbb{R} :: \theta(x) = \tau$

which tells that the shock occurs to the particle x at the time (for the observer) : $\tau = \theta(x)$.

The shock wave is then located at t on the points $\omega(t) = \{\varphi_o(ct, x) :: t = \theta(x)\}$. And spatially it propagates on $\omega_3(t) = \{t = \theta(x)\}$. The equation $t = \theta(x)$ defines a foliation of $\Omega_3(0)$ in 2 dimensional hypersurfaces, which represent the waves. The spatial speed of propagation is given by the gradiant of the function θ.

The discontinuity can then be represented by some map :

$\delta X : \Omega_3(0) \to T_1 Spin(3,1) :: \delta X(\varphi_o(\theta(x), x))$

and :

$\upsilon(X_r(\varphi_o(t,x)), X_w(\varphi_o(t,x))) = \sum_{\alpha=0}^{3} V^\alpha \partial_\alpha \sigma \cdot \sigma^{-1} + \delta X(\varphi_o(\theta(x), x))$

The discontinuity appears as a map, valued in the same vector space as $\upsilon(X_r, X_w)$, null everywhere but on the waves, and which is added to the continuous derivative. Mathematically this is the usual representation of a discontinuous derivative with distributions (or generalized functions) : the jump in the derivative appears as a Dirac's function.

We could consider to extend this scheme to force fields. But there is a major difference : there is no "material points" in force fields, and they propagates along lines which are not integral curves of a single vector field. Actually there are infinitely many such curves which originate from any given point. So the propagation of discontinuities in a force field occurs along lines, and not 3 dimensional waves. And this is at the root of the bosons.

We have seen in the study of the field equations that, even in a continuous model, there is an issue to find solutions which both meet the conditions in the vacuum (that is, in a real world, almost everywhere) and at the location of particles (at least in models which identify individual particles, in model of the first kind there is no propagation to speak of). The root of the problem is in the concepts of fields and particles : according to the Principle of Causality we should distinguish an incoming field and an outgoing field before and after it encounters a particle. The adjustment implies some discontinuity, because of the identification of the particle with a geometric point. Albeit the concept of field and its propagation implies its continuity, at least at some level. We have solved this issue, from a mathematical point of view, by assuming that the variables $\mathcal{F}$ are distinct, continuous variables (which is the requirement for the propagation). This solution is acceptable when the purpose of the model is to compute the properties of the field over some extended area, that is at a macroscopic scale, in what is already a dreadful endeavour. However we need a more robust solution, which goes beyond the computational necessities. We will proceed by coming back to the definition of $\mathcal{F}$ from the basic variable which is the connection itself, and we will take as example a general connection $\grave{\mathbf{A}}$ associated to the group U (whose precise definition does not matter here).

8.2.2 Mathematical representation

The mathematical representation of discontinuities of force fields

Our purpose is to represent a discontinuity in the derivative of the connection $\grave{\mathbf{A}}$ on P_U, which propagates. For this we start from the definition of $\mathcal{F}$ (see Chapter 5).

A vector $X(p) \in T_pP_U$ reads $X(p) = \sum_{\alpha=0}^{3} X_m^\alpha(p)\, \partial m_\alpha + \zeta(X_U)(p)$ where $\zeta(X_U)(p)$ is a fundamental vector located at $p \in P_U$ and defined by $X_U = \sum_{a=1}^{m} X_U^a(p)\, \overrightarrow{\theta}_a \in T_1U$.

The connection is a tensor acting on vectors of T_pP_U and valued in the vertical bundle :

$$\grave{\mathbf{A}}(p)(X_m + \zeta(X_U)(p)) = \zeta\left(\widehat{\grave{A}}(p)(X)\right)(p)$$

with the connection form $\widehat{\grave{A}} \in \mathbf{\Lambda}_1(TP_U; T_1U)$.

For a principal connection :

$$\widehat{\grave{A}}(\varphi_U(m,g))((X_m + \zeta(X_U))(\varphi_U(m,g))) = X_U + Ad_{g^{-1}}\left(\grave{A}(m)\, X_m\right)$$

where $\grave{A}(m)$ is the potential $\grave{A} \in \Lambda_1(TM; T_1U)$.

The derivative of the connection at a point $p \in P_U$ is defined along a vector field $W \in \mathfrak{X}(TP_U)$ through :

$$\Delta_R(s) = \frac{1}{s}\left(\Phi_W(s,p)^*\, \widehat{\grave{A}}(p) - \widehat{\grave{A}}(p)\right)$$

$$\Delta_L(s) = \frac{1}{s}\left(\widehat{\grave{A}}(p) - \Phi_W(-s,p)^*\, \widehat{\grave{A}}(p)\right)$$

If $\lim_{s\to0} \Delta_R(s) = \lim_{s\to0} \Delta_L(s)$ then the connection is differentiable at p and $\pounds_W \widehat{\grave{A}}(p) = \lim_{s\to0} \Delta_R(s)$.

But the quantities may have limits which are not equal : we have a discrepancy in the derivative, which can be measured by :

$$\Delta_W\left(\grave{\mathbf{A}}(p)\right) = \lim_{s\to0} \frac{1}{s}\left(\Phi_W(s,p)^*\, \widehat{\grave{A}}(p) - \Phi_W(-s,p)^*\, \widehat{\grave{A}}(p)\right)$$

$\mathcal{F}$ is defined as a derivative with respect to a displacement in M. So the derivative is for a section $S \in \mathfrak{X}(P_U)$ and the horizontal lift χ_L of a vector field V on TM.

Let us just take a section $\mathbf{P} \in \mathfrak{X}(P_U) : \mathbf{P}(m) = \varphi_U(m, \gamma(m))$ and a projectable vector field W on $TP_U : \pi'_U(p)\, W(p) = V(\pi_U(p))$ $(\chi_L(p(m))(V(m))$ is projectable). The affine parameter s is the same along the integral curves of V, W :

$$\Phi_W(s, \mathbf{P}(m)) = \Phi_W(s, \varphi_U(m, \gamma(m))) = \varphi_U(\Phi_V(s,m), \gamma(\Phi_V(s,m))) = \mathbf{P}(\Phi_V(s,m))$$

$$\Phi_W(s, \mathbf{P}(m))' = \mathbf{P}'(\Phi_V(s,m))\, \Phi_V(s,m)'$$

$$\Phi_W(s, \mathbf{P}(m))'((X_m + \zeta(X_U))(\mathbf{P}(m))) = \Phi_V(s,m)'\, X_m + \zeta\left(\left(L'_{\gamma^{-1}}\gamma\right) X_U\right)(\mathbf{P}(m))$$

$\widehat{A}(\mathbf{P}(m))((X_m+\zeta(X_U))(\mathbf{P}(m)))=X_U+Ad_{\gamma(m)^{-1}}\grave{A}(m)X_m$
$\Phi_W(s,p)^*\widehat{A}(p)(X(p))=\widehat{A}(\Phi_W(s,p))\Phi_W(s,p)'(X(p))$
$\Phi_W(s,\mathbf{P}(m))^*\widehat{A}(\mathbf{P}(m))(X(\mathbf{P}(m)))=\left(L'_{\gamma^{-1}}\gamma\right)X_U+Ad_{\gamma^{-1}}\left(\grave{A}(\Phi_V(s,m))\Phi_V(s,m)'(X_m)\right)$
So the derivative is computed by :
$\left(\Phi_W(s,p)^*\widehat{A}(p)-\Phi_W(-s,p)^*\widehat{A}(p)\right)(X(\mathbf{P}(m)))$
$=\left(\left(L'_{\gamma^{-1}}\gamma\right)(\Phi_V(s,m))-\left(L'_{\gamma^{-1}}\gamma\right)(\Phi_V(-s,m))\right)X_U$
$+Ad_{\gamma(\Phi_V(s,m))^{-1}}\left\{\left(\grave{A}(\Phi_V(s,m))\Phi_V(s,m)'(X_m)\right)-\left(\grave{A}(\Phi_V(-s,m))\Phi_V(s,m)'(X_m)\right)\right\}$
If there is a discontinuity, let us define :
$\Delta_W\left(\grave{\mathbf{A}}(\mathbf{P}(m))\right)(X(\mathbf{P}(m)))$
$=Ad_{\gamma(m)^{-1}}\left\{\lim_{s\to 0}\frac{1}{s}\left(\grave{A}(\Phi_V(s,m))-\grave{A}(\Phi_V(-s,m))\right)\right\}\Phi_V(0,m)'(X_m)$
thus $\Delta_W\left(\grave{\mathbf{A}}(\mathbf{P}(m))\right)\in\Lambda_1(TM;T_1U)$, in particular with the standard gauge $\mathbf{p}(m)=\varphi_U(m,1)$:
$\Delta_W\left(\grave{\mathbf{A}}(\mathbf{p}(m))\right)=\left\{\lim_{s\to 0}\frac{1}{s}\left(\grave{A}(\Phi_V(s,m))-\grave{A}(\Phi_V(-s,m))\right)\right\}\Phi_V(0,m)'$
so that : $\Delta_W\left(\grave{\mathbf{A}}(\mathbf{P}(m))\right)=Ad_{\gamma(m)^{-1}}\Delta_W\left(\grave{\mathbf{A}}(\mathbf{p}(m))\right)$.Because $\Delta_W\grave{\mathbf{A}}$ is defined by a difference, it transforms by $Ad_{\chi^{-1}}$, as $\mathcal{F}$ (for the same reasons) : it can be seen as a 1 form on TM valued in T_1U. 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), and in QTF bosons are represented like the potential, and the transformations rules are one of the main motivations for the introduction of the Higgs boson. But this restriction applies no longer to $\Delta_W\grave{\mathbf{A}}$.

$\Delta_W\left(\grave{\mathbf{A}}(\mathbf{p}(m))\right)$ can be written :
$\Delta\widehat{A}(\mathbf{p}(m),W)=\sum_{a=1}^m\sum_{\alpha,\beta=0}^3\Delta\grave{A}^a_\beta(m)\left[\Phi_V(0,m)'\right]^\beta_\alpha d\xi^\alpha\otimes\overrightarrow{\theta}_a$

so actually it does not depend on the choice of the projectable vector field W, but of course, to be consistent with $\mathcal{F}$ we can choose the horizontal lift of any vector field V, or of the tangent to a curve defined by V.

$\Delta\widehat{A}(\mathbf{p}(m),\chi_L(V))=\sum_{a=1}^m\sum_{\alpha,\beta=0}^3\Delta\grave{A}^a_\beta(m)\left[\Phi_V(0,m)'\right]^\beta_\alpha d\xi^\alpha\otimes\overrightarrow{\theta}_a$
that we denote : $\Delta\grave{A}(m)=\sum_{a=1}^m\sum_{\beta=0}^3\Delta\grave{A}^a_\beta(m)d\xi^\beta\otimes\overrightarrow{\theta}_a$
If there is no discontinuity then $\Delta_W\left(\grave{\mathbf{A}}(\mathbf{p}(m))\right)=0$.

So far $\Delta\grave{A}(m,V)$ is a covector on T_mM valued in T_1U. We assume that the discontinuity propagates. We can proceed as for a discontinuity in a continuous medium : it is represented as a quantity which is added to the continuous derivative, here $\mathcal{F}$. We need also to define the vectors V. We have seen that fields propagate along Killing curves at a constant speed. Clearly discontinuities emanate from a point (usually the interaction with a particle) and the propagation is along such a curve. So this is very similar to the propagation of a signal.

The propagation of the strength $\mathcal{F}$ follows specific rules :
$[\delta\mathcal{F}^a(\tau)]=\theta(\tau)[K(\tau)]^t[\delta\mathcal{F}^a(O)][K(\tau)]$

The signal $\delta\mathcal{F}^a$ originating at O is transported along the curve, thus its components change through $[K(\tau)]$. This process depends only on the curve, it is the same for all the components a. But there is an additional process, linked to the attenuation of the signal along its propagation, and represented by θ.

We can assume safely that the discontinuity propagates along the same curves as the signals of the same type of fields. There are Killing curves, with a speed w depending on the type of field.

The attenuation process is clearly linked to the fact that the propagation of a signal occurs on hypersurfaces $S_3(O,\tau)$: the energy is spread. We have nothing equivalent here, the discontinuity propagates along a curve. So we can safely assume that there is no attenuation[1]. And actually there is none for the photons (the "red shift" is just a Doppler effect, and the attenuation in Astrophysics is assumed to come from the expansion of the Universe). So there is no θ involved and the motion of bosons is characterized by $\pounds_V \Delta \grave{A} = 0$.

And we state :

Proposition 103 *Discontinuities of fields can be represented as maps $\Delta \grave{A} \in \Lambda_1 (TM; T_1U)$, with support a Killing curve and propagate at the same speed as the type of fields to which they belong, by transport such that $\pounds_V \Delta \grave{A} = 0$.*

We have a picture similar to particles : an object living on a curve, with a constant velocity $\langle V, V \rangle = w^2 - c^2$ 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 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).

Motion

The discontinuity is actually a discontinuity in the derivative of the potential, and not of the potential or the strength. We have noticed that the field can be represented in the jet formalism by : $\left(m, \grave{A}^a_\alpha, \delta_\beta \grave{A}^a_\alpha, a = 1...m, \alpha, \beta = 0...3\right)$ where $\grave{A}^a_\alpha, \delta_\beta \grave{A}^a_\alpha$ are independent variables, and if the potential is continuously differentiable then $\delta_\beta \grave{A}^a_\alpha = \partial_\beta \grave{A}^a_\alpha$. Because we represent the discontinuity as added to a underlying, smooth, field, we keep the definition of the strength as $\mathcal{F}^a_{\alpha\beta} = \partial_\alpha \grave{A}^a_\beta - \partial_\beta \grave{A}^a_\alpha + 2 \left[\grave{A}_\alpha, \grave{A}_\beta\right]^a$, that is a smooth variable.

The motion of the boson is represented as for particles in the jet formalism.

The trajectory is given by the curve $q(t) = \Phi_V(t, O)$ with tangent V, and the motion itself by the map $\Phi'_V(t, O) :: T_O M \to T_m M$

Because $\pounds_V \Delta \grave{A} = 0 : \Delta \grave{A}(\Phi_V(t, O)) = \Phi_V(t, .)_* \Delta \grave{A}(O)$ and with a matrix $[K(t)]$ representing the transport along the curve :

$\Delta \grave{A}_\alpha(t) = \sum_{\beta=0}^3 [K(t)]^\beta_\alpha \Delta \grave{A}_\beta(O)$

that we can represent in matrix form :

$$\left[\Delta \grave{A}(t)\right] = \left[\begin{array}{cc} \left[\Delta \grave{A}_0(t)\right]_{m\times 1} & [\Delta a(t)]_{m\times 3} \end{array}\right]$$

and in the standard chart :

$$[K(t)] = \left[\begin{array}{cc} 1 & 0 \\ 0 & [k]_{3\times 3} \end{array}\right]$$

$$\left[\Delta \grave{A}(t)\right] = \left[\begin{array}{cc} \left[\Delta \grave{A}_0(0)\right] & [\Delta a(0)][k] \end{array}\right]$$

So the *motion* of the boson can be represented by : $\left(q(t), V^\alpha, [K(t)]^\alpha_\beta, \alpha, \beta = 0...3\right) \in J^2 TM$.

To particles one can associate a section of P_G, this is based on the fact that any time like, future oriented, vector can be represented in the tetrad as $U = -\frac{c}{\langle \mathbf{Ad}_\sigma \varepsilon_0, \varepsilon_0 \rangle_{Cl}} \mathbf{Ad}_\sigma \varepsilon_0$. This is no longer possible, at least for photons and gravitons whose trajectories are null curves.

Fundamental state

A boson originates from a point and travels on a precise curve. The vector V is part of the definition of the boson.

Along a propagation curve both $\Delta \grave{A}^a, V$ are transported :

[1] This is the opposite to shock waves in deformable solids, which expand over 2 dimensional areas and dissipate quickly their energy.

$\left[\Delta \grave{A}^a\left(q\left(\tau\right)\right)\right]=\left[\Delta \grave{A}^a\left(q\left(0\right)\right)\right]\left[K\left(q\left(\tau\right)\right)\right]$
$\left[V\left(q\left(\tau\right)\right)\right]=\left[K\left(q\left(\tau\right)\right)\right]^{-1}\left[V\left(q\left(0\right)\right)\right]$
so the quantity :
$\sum_{a=1}^m\sum_{\alpha=0}^3 V^\alpha\Delta\grave{A}_\alpha^a\left(m\right)\overrightarrow{\theta}_a\in T_1U$
is preserved :
$\left[\Delta \grave{A}^a\left(q\left(\tau\right)\right)\right]\left[V\left(q\left(\tau\right)\right)\right]=\left[\Delta \grave{A}^a\left(q\left(0\right)\right)\right]\left[V\left(q\left(0\right)\right)\right]=B_A=Ct$
It does not depend on the chart :
$B_A=c\Delta\grave{A}_0^a\left(q\left(0\right)\right)+\left[\Delta \grave{A}^a\left(q\left(0\right)\right)\right]\left[v\right]$
B_A is similar to the fundamental state of a particle : it is preserved along the propagation.

$$\sum_{\alpha=0}^3 V^\alpha\Delta\grave{A}_\alpha^a=B_A=Ct\in T_1U \tag{8.1}$$

8.2.3 Quantization of Bosons

Types of bosons

For an observer bosons are maps :
$\Delta\grave{A}:\left[0,T\right]\rightarrow TM^*\otimes T_1U$
They belong to a normed vector space F, invariant by a global change of gauge on P_U:
$\mathbf{p}_U\left(m\right)=\varphi_U\left(m,1\right)\rightarrow\widetilde{\mathbf{p}}_U\left(m\right)=\widetilde{\varphi}_U\left(m,1\right)=\mathbf{p}_U\left(m\right)\cdot\chi\left(m\right)^{-1}$
$\Delta\grave{A}\rightarrow Ad_\chi\Delta\grave{A}$
Along the curve : $\sum_{\alpha=0}^3 V^\alpha\Delta\grave{A}_\alpha^a=B_A=Ct\in T_1U$ which depends on the boson.
Let be an observable $\Phi:F\rightarrow F_0$ where F_0 is a finite dimensional vector space of F.
We can implement the Theorem 24 of the Chapter 2.
F is isomorphic to an open of a Hilbert space H : $\Upsilon:F\rightarrow H$ and the action Ad has for image :
$\widehat{Ad}\in\mathcal{L}\left(H;H\right):\widehat{Ad}_\chi=\Upsilon\circ Ad_\chi\circ\Upsilon^{-1}$
$\widehat{Ad}\left(\Upsilon\left(F_0\right)\right)=\widehat{F}_0$
$\left(H,\widehat{Ad}\right)$ is a unitary representation of U, $\left(\widehat{F}_0,\widehat{Ad}\right)$ is a finite dimensional, unitary representation of U.
The vector space F_0 is invariant by Ad, and $\left(F_0,Ad\right)$ is a representation of U.
The relation of equivalence :
$R:\Delta\grave{A}\sim\Delta\grave{A}'\Leftrightarrow\sum_\alpha V^\alpha\Delta\grave{A}_\alpha=\sum_\alpha V^\alpha\Delta\grave{A}'_\alpha$
defines on the sets F,H a partition and for a given value of B_A the corresponding subset $H\left(B_A\right)$ of H is invariant by U. The specification of $\Phi:F\rightarrow F_0$ chosen to measure $\Delta\grave{A}$ corresponds to an irreducible representation.

As a consequence each observable Φ corresponds to a definite value of $B_A\in T_1U$. The corresponding vector space F_0 can be identified by B_A, it is finite dimensional, and we can assume that it can be identified with a vector $\overrightarrow{\theta}_a$ of the basis of T_1U : there are as many kinds of bosons as the dimension of U. Bosons can be labelled by the vectors θ_a and $B^a=b^a\overrightarrow{\theta_a}$ where b^a is constant along the propagation and changes in a change of gauge according to the rules in $P_U\left[T_1U,Ad\right]$.
For $U=U\left(1\right)$, $T_1U\left(1\right)=\mathbb{R}$: photons are characterized by a fixed scalar $\nu\in\mathbb{R}$.
For $Spin\left(3,1\right)$ we should have 6 kinds of graviton, one for each vector $\overrightarrow{\kappa}_a$. But $Spin\left(3,1\right)$ is not compact. Its only unitary, irreducible representations are infinite dimensional, parametrized by $k\in\mathbb{R},z\in\mathbb{Z}$. So we cannot have a unitary, finite dimensional representation $\left(\widehat{F}_0,\widehat{\mathbf{Ad}}\right)$, whatever the Hilbert space H, for all gravitons. However $T_1Spin\left(3,1\right)=L_0\oplus P_0$ which are globally invariant by $Spin\left(3\right)$, the scalar product is definite (positive or negative) and preserved by $\mathbf{Ad}$, so L_0,P_0

are 3 dimensional Hilbert spaces, and for each choice of ε_0, $(L_0, \mathbf{Ad}), (P_0, \mathbf{Ad})$ are 3 dimensional unitary representations of $Spin(3)$. So, for a given observer (who defines ε_0) there could be 3 types of gravitons differentiated by $a = 1, 2, 3$, and another category of gravitons differentiated by $a = 4, 5, 6$. In a change of spatial gauge (that is with the same observer), gravitons stay in one or the other category. In time reversal the categories change into the other. We have a situation similar to the distinction particles / antiparticles and we will call gravitons the bosons for $a = 1, 2, 3$, antigravitons the bosons for $a = 4, 5, 6$.

And we can state :

Proposition 104 *There are :*

3 kinds of gravitons, associated to the vectors $\overrightarrow{\kappa}_a = 1, 2, 3 \in T_1 Spin(3,1)$

3 kinds of antigravitons, associated to the vectors $\overrightarrow{\kappa}_a = 4, 5, 6 \in T_1 Spin(3,1)$

There is only one type of photon, characterized by a scalar $\nu \in \mathbb{R}$.

There is one type of boson associated to each vector $\overrightarrow{\theta}_a \in T_1 U$

A boson is then represented by a map :

Photon :$\Delta\varphi : [0, T] \to TM^* :: \Delta\varphi(t) = \sum_{\alpha,\beta=0}^{3} [K(t)]_{\alpha}^{\beta} \Delta\varphi_\beta d\xi^\alpha (q(t))$

Graviton : $\Delta\Gamma : [0, T] \to TM^* \otimes L_0 :: \Delta\Gamma(t) = \sum_{\alpha,\beta=0}^{3} [K(t)]_{\alpha}^{\beta} \Delta\Gamma_\beta d\xi^\alpha (q(t)) \otimes \overrightarrow{\kappa}_a$ with a equal either $1, 2, 3$

Antograviton : $\Delta\overline{\Gamma} : [0, T] \to TM^* \otimes P_0 :: \Delta\overline{\Gamma}(t) = \sum_{\alpha,\beta=0}^{3} [K(t)]_{\alpha}^{\beta} \Delta\overline{\Gamma}_\beta d\xi^\alpha (q(t)) \otimes \overrightarrow{\kappa}_a$ with a equal either $4, 5, 6$

Other bosons : $\Delta\grave{A} : [0, T] \to TM^* \otimes T_1 U :: \Delta\grave{A}(t) = \sum_{\alpha,\beta=0}^{3} [K(t)]_{\alpha}^{\beta} \Delta\grave{A}_\beta^a d\xi^\alpha (q(t)) \otimes \overrightarrow{\theta}_a$

where $[K(t)]$ is associated to the propagation curve.

The vectors $\Gamma_\beta \overrightarrow{\kappa}_a, ...$ change in a change of gauge as usual. φ_β is constant.

And the fundamental states are the real scalars :

Photon :$\sum_{\beta=0}^{3} V^\beta \Delta\varphi_\beta$

Graviton : $\sum_{\beta=0}^{3} \Delta\Gamma_\beta V^\beta$

Antigraviton : $\sum_{\beta=0}^{3} \Delta\overline{\Gamma}_\beta V^\beta$

Other bosons : $\sum_{\alpha,\beta=0}^{3} \Delta\grave{A}_\beta^a V^\beta$

where V is the tangent to the propagation curve.

Spin

The spin of a particle comes from the representation of its motion, which distinguishes two rotations with respect to the velocity. The motion of bosons with respect to the holonomic basis of a chart is given by the matrix $[K(t)]$. For bosons the equivalent is the rotation of the holonomic basis with $\Phi'_V(t, O)$. The map is assumed to be smooth, so over a connected interval of time the sign of the determinant $\det[K]$ has a fixed value ± 1, and we have two possible spinning motions. *Bosons have a spin 1, with two possible states of the spin.* This can be seen equivalently as two classes of irreducible representations in the quantization of the maps $\Delta\grave{A}(t)$.

And this should hold also for the gravitons. Moreover if gravitons are associated to vectors of $T_1 Spin(3)$ with the group $Spin(3)$ a change of gauge by -1 has the same effect as the inversion of the rotation.

Anti-particles

Anti-particles have been introduced to account for the two possible representations of fermions by spinors. Except for the gravitons, we have nothing equivalent here and the other bosons can be their own anti-particles. But time reversal has a clear meaning for particles which travels at the speed of light : it sums up to inverse the trajectory. Mathematically this is equivalent to take the opposite spin, and photons are their own anti-particles. But the situation is less obvious for the other bosons.

Charge of the bosons

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. Because bosons are represented by a vector of their Lie algebra, they inherit the charges which are imputed to the corresponding particles in the representation (F, ϱ). For each of them there are 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 have no charge.

Observables of bosons

Whatever the primary observable Φ, the measure which is observed is an eigen vector of the operator Φ, with a probability depending on the state of the boson, that is on the map $\Delta\grave{A} : [0,T] \to TM^* \otimes T_1U$ which belongs to a subset of F characterized by the kind of boson.

A discontinuity is certainly a singular phenomenon. A discontinuous process is a transition between states of equilibrium which can be represented as the result of continuous processes. The continuous models are still useful, but their equations must be understood in the meaning of distributions, as giving the result of measures taken in a 4 dimensional area which encompasses the singular point. Any observable of a boson is related to measures which can be similarly performed on force fields, such as energy. The detection and measures of bosons are then the detection and measures of a singular phenomenon, with respect to the continuous fields which exist in the background. We have seen in the 2nd chapter that, to be successful, there is a condition related to the signal to noise ratio. Bosons will not be detected if this ratio is too low. It does not mean that they do not exist, but that any attempt to detect them will fail (or more precisely deemed not "scientifically conclusive").

Build an observable sums up to choose a specification for the variable. A Killing curve is defined by the value of the tangent at a point, and a boson is then defined by a vector $v \in \mathbb{R}^3$ and the 4 components $\Delta\grave{A}^a_\beta$ along a vector of the Lie algebra. And on this point there is an important difference between the photon and the other bosons. There is only one type of boson attached to the EM field, so any measure about it will involve only the components v^α, φ_β. But there are different types of bosons attached to the other force fields (6 for the gravitational fields, 3 for the weak interactions and 8 for the strong interactions). Even if a given boson belongs only to one type, its type must be identified. The only way to identify a boson is to compare its behavior with bosons or fermions interacting with the same force field. The observables are then maps from the whole of $TM^* \otimes T_1U$ to F_0.

Bosonic fields

To a particle one can associate a, non unique, matter field with the fundamental state ψ_0 and a section $\sigma \in \mathfrak{X}(P_G)$ such its trajectory is one of the integral curves of the associated vector field.

A boson is defined by its point of origin O, a curve which is a Killing curve with a future oriented vector of specific Lorentz length, and a tensor $\Delta\grave{A}^a_\alpha$ which is propagated with a definite law, depending only on the curve, and transforms regularly in a change of gauge. So to a boson at a given location, one can associate a bosonic field, which is not unique. And conversely a Killing vector field (with the adequate vector) and a single vector $\Delta\grave{A}^a_\beta(m)$ at a point m defines uniquely a boson.

Bosons in SR Geometry

In SR geometry a spherical chart centered at a point O is : $m = ct\varepsilon_0 + \rho\overrightarrow{u}$ where $\overrightarrow{u}$ is a unitary vector $u \in \mathbb{R}^3$ normal to $S_3(O,\rho)$. The Killing curves are straight lines, and a signal originating at

O is then a wave with wave vector $\overrightarrow{u}$: $[\delta\mathcal{F}(t+\tau, w\tau, u)] = \theta(\tau)[\delta\mathcal{F}(t,0,u)]$ and similarly for the potential : $\left[\delta\grave{A}(t+\tau, w\tau, u)\right] = \left[\delta\grave{A}(t,0,u)\right][\theta(\tau)]$

So that a boson propagates along a straight line with fixed spatial direction $\overrightarrow{u}$ as :
$\left[\Delta\grave{A}(t+\tau, w\tau, u)\right] = \left[\Delta\grave{A}(t,0,u)\right]$

The components $\Delta\grave{A}_\beta^a$ are constant.

The fundamental state is with : $V = c\varepsilon_0 + w\overrightarrow{u}$
$B_A = \sum_{\alpha=0}^3 V^\alpha \Delta\grave{A}_\alpha^a = c\Delta\grave{A}_0^a + w\sum_{\beta=1}^3 u^\beta \Delta\grave{A}_\alpha^a$

QED and QTF

The EM and gravitational fields have an infinite range, meanwhile the range of the weak and strong interactions is very short, and this has a direct consequence on the way the related bosons are considered practically.

Whenever weak or strong interactions are involved, the motion of the boson does not matter. The dominant feature is their interaction with fermions as a localized potential $\Delta\grave{A}$, modelled through a lagrangian. This is the topic of the Quantum Theory of Fields (QTF). Photons are then composite particles, and gravitons are not considered.

The photons, and the gravitons as far as we can assume, propagate without attenuation on great distances, and their propagation curves are part of their definition. When the weak or strong interactions are not involved their dominant feature is as discontinuity of an underlying field and their action is then modelled as collisions, through the Energy-Momentum tensor. This is what is done in Quantum Electro-dynamics (QED).

8.3 BOSONS IN QED

In QED only the photons, which are then elementary particle, are considered, but gravitons should have similar properties.

The potential $\grave{A}$ of the EM field is a one form : $\grave{A} = \sum_\alpha \grave{A}_\alpha d\xi^\alpha \in \mathfrak{X}(TM^*)$:

$\mathcal{F}_{\alpha\beta} = \sum_\alpha \left(\partial_\alpha \grave{A}_\beta - \partial_\beta \grave{A}_\alpha\right) d\xi^\alpha \wedge d\xi^\beta$

In the standard chart $\left[d\grave{A}^r\right]$ is related to the magnetic field, and $\left[d\grave{A}^w\right]$ to the electric field :

$[\mathcal{F}^r_{EM}] = -[B]^t \det Q' = \left[d\grave{A}^r\right]$

$[\mathcal{F}^w_{EM}] = [E]^t [g_3] = \left[d\grave{A}^w\right]$

Using the same decomposition in the tetrad of an observer, a photon has an "electric component" $\Delta_E\varphi$ in the time direction ε_0 and a "magnetic component" $\Delta_B\varphi$ in the 3 dimensional space, such that :

$\Delta\varphi = (\Delta_E\varphi)\,\varepsilon^0 + \sum_{j=1}^3 (\Delta_{Bj}\varphi)\,\varepsilon^j$

$c\Delta_E\varphi + \Delta_B\varphi(v) = Ct$

8.3.1 Interaction of photons with particles

Photons appear or disappear during interactions of the EM field with particles - essentially electrons - and their properties are defined through these discontinuous interactions.

In condensed matter there are electrons which are free, or whose link with atoms is weak. They are the carriers of the usual electric current. Atoms and molecules can also vibrate. In normal circumstances their interaction with the EM field is represented in a continuous process, as in our 2 models. The field modifies the motion of the particle, and the change in the field propagates following a law which smears out the local discrepancy. The effects of this continuous process are :

- a reflection and refraction of the incoming field

- a thermal effect : the field transfers kinetic energy to the particles, increasing the temperature of the solid, and conversely the thermal agitation of the electrons induces a field which is then reemitted as radiance heat

- a pressure on the solid : the change in the motion of the particles is transferred to the solid through the internal links.

At a macroscopic scale these phenomena appear as continuous, and their amplitude depends only on the intensity of the field.

However, the states of a system composed of microsystems interacting are quantized, as seen in the 2nd Chapter. In an atom electrons are arranged in shells, corresponding to specific states of a common matter field, and notably to different level of energy. Free electrons in condensed matter show similar properties, as well as the vibration modes of crystals. When the levels of energy are close (as for electrons on the "exterior" shells) the change appears as continuous, but when the discrepancy is too big, the change involves a discontinuous process.

When the electron loses energy the balance is provided to the field. The point where the interaction occurs is singular, as in any interaction. The local excess of energy is dissipated through a mechanism involving a fixed law (the matrix θ) which has a limited capacity to do it in a given volume. When this capacity is exceeded a discontinuity appears in the field, and the excess of energy is carried away as a photon, propagating on a line.

Conversely an electron can be upgraded to a shell with a greater level of energy, but this requires a supply of energy localized at a point, which is not possible with a continuous field. This supply of energy can only come from a discontinuity, that is from a photon.

These discontinuous processes are common, and indeed there are the dominant phenomena at the atomic level. Their main characteristic is the existence of a threshold : if their effect at the macroscopic scale still depends on the intensity of the field, they happen only if some conditions, linked essentially to the energy, are met. The system which should be considered comprises the electrons, atoms and the fields, so it is complicated, and any conceivable model is based on specifications of each of the components, as we have done for particles with matter fields, and of interacting microsystems. Even if the individual processes are not random, any measure which can be done appears as the result of a process in which a transition happens with some probability.

These mechanisms explain phenomena such as :

- photo-electric emission : if the energy brought by the field is sufficient, the link with the atom is broken and a free electron appears. The conversion of energy between the incoming field and the current collected is not 100 % : above the threshold there is only a probability, increasing with the intensity, that an electron is ejected
- black-body radiation : thermal agitation of electrons causes the emission of a EM field, as radiance heat. Its spectrum in frequencies depends on the balance between the emission / reabsorbtion of the field by the electrons, internal to the material. The existence of a threshold has an impact to the aspect of the spectrum.
- Compton effect : the encounter between a photon and a free charged particle changes both the trajectory of the particle and the characteristic of the field.

8.3.2 Representations of the mechanisms

The tetrad equation

The mechanisms above can be represented as transitions between two states of equilibrium, so the continuous models are useful. In particular the tetrad equation, which gives the balance of energy, understood in the meaning of distributions, still holds. Particles are represented as belonging to a matter field : their trajectories are integral curves of a given vector field. The variation of the energy of a particle is given by the tensor :

$$\sum_{\alpha,\beta=0}^{3} V^{\alpha}\frac{\delta\mathcal{E}}{\delta\xi_{\beta}}\partial\xi_{\alpha}\otimes d\xi^{\beta}=\left(\sum_{\alpha=0}^{3}V^{\alpha}\partial\xi_{\alpha}\right)\otimes\left(\sum_{\beta=0}^{3}\frac{\delta\mathcal{E}}{\delta\xi_{\beta}}d\xi^{\beta}\right)\in TM\otimes TM^{*}$$

with the variational derivative $\frac{\delta\mathcal{E}}{\delta\xi_{\beta}}$. The integral curves associated to a given section are curve of constant energy : $\langle\psi,\nabla_{V}\psi\rangle=0$. A change of the energy of the particle implies that the particle goes from one integral curve to another, and we can no longer assume that $\frac{\delta\mathcal{E}}{\delta x}=0,\frac{\delta\mathcal{E}}{\delta t}=0$, it implies a change in the quantities on the left hand side of the tetrad equations. One can assume that the metric does not change : this is a continuous variable, defined all over M, so the balance is imputed to the fields. Actually it could be either the EM field or the gravitational field, the discrepancy between the intensity of the fields makes that the change, as it can be measured, is imputed to the EM field. However one cannot discount a change in the gravitational field, difficult if not impossible to measure.

The Energy-Momentum tensor

Moreover the change involves other variables than the energy, and the most complete picture is given by the Energy-Momentum tensor, which gives the impact of a change of any variable of the system along a vector field V.

$$T=$$
$$-\frac{1}{2}C_{I}\mu MV\otimes k_{0}^{t}\operatorname{Re}D\left(-Z\right)dZ+4\sum_{\alpha\beta\gamma}\left(4C_{G}\left\langle\mathcal{F}_{G}^{\alpha\gamma},\partial_{\beta}G_{\gamma}\right\rangle_{Cl}+C_{A}\left\langle\mathcal{F}_{A}^{\alpha\gamma},\partial_{\beta}\grave{A}_{\gamma}\right\rangle_{T_{1}U}\right)\partial\xi_{\alpha}\otimes d\xi^{\beta}$$

The balance is such that $T\left(V\right)=0$.For a particle, which belongs to a matter field, the partial derivatives $\partial_{\beta}Z$ are defined for any direction $d\xi^{\beta}$. For the field this is a change $\Delta\grave{A}$ along a definite curve : it is null except along the curve. So that $T_{EM}\left(V\right)=0$ except in the direction of the propagation of the photon. Which is similar to imparting a momentum to the photon in the direction of its propagation

The energy is given by the trace of the tensor :

$Tr(T) = L_{Particles} + 4L_{Fields}$
and at equilibrium :
$\langle \mathcal{F}_{EM}, \mathcal{F}_{EM} \rangle = G_1 \left(\delta \mathcal{F}_{EM}, \grave{A} \right)$
For the underlying field the equation $-\delta \mathcal{F}_{EM} = 2J^*_{EM}$ holds, so that the variation of energy can be written as :

$$\delta E_{Field} = -2\Delta \grave{A} \left(J_{EM} \right) = -\frac{1}{4} C_I \delta K \quad (8.2)$$

where the current is understood as the current induced by the underlying matter field (this is the current related to one of the integral curves of the matter field). This is clearly what happens to an electron in an atom : its trajectories are periodic motions with different periods, a change of kinetic energy implies a change of trajectory, inside the same section of P_G .

The usual representations

In SR geometry fields and photons propagate along straight lines, and a photon is characterized by a scalar and a single spatial vector. A photon carries energy, it looks also like it carries a momentum, so the natural representation is as a particle with momentum travelling along a straight line. Moreover, because the usual paradigm of the EM field is the "plane wave", the picture was easy to hold. However there are two points which need some explanation : the reference to a frequency, as in the celebrated Plank's law, and the null mass of the photon.

The use of a frequency to describe a photon has historical reasons : its existence came in light with the study of the spectrum of the black-body, in which the frequency is a central variable. The distortion of the spectrum, due to the threshold effect, is then explained by the idea that the exchanges between particles and field could happen only in definite "quanta", expressed as multiple of the frequency.

There is another motivation. The existence of a photon is acknowledged through one of the mechanisms above. A photon appears as a discrepancy in the EM field, a continuous field which exists everywhere and to which the discontinuity is added. Whatever the observable, a singularity is acknowledged as such as a deviation from a continuous phenomenon. Any measure of the EM field is done over a definite period of time. For a periodic field, over any interval of time any observable of the field changes with some frequency ν. Practically a sample is taken over some period of time, and the values measured $(z_1, z_2, ...z_N)$ are compared to the values $(x_1, x_2, ...x_N)$ which would be expected according to the underlying field. If there is a discrepancy : $z_p = y_p \neq x_p$ then a discontinuity has been detected, and the experiment can be repeated : over more than one period the probability to detect the anomaly is always the same, and is proportional to the frequency. This is similar to what has been shown in the 2nd Chapter on QM. The threshold for the detection of the photon decreases with the frequency. And actually it is quite impossible to detect a photon with fields of large wave lengths. Because the observable is usually linked to the energy, this is the basic interpretation of the Planck's law : $E = h\nu$.

A mass could be given to the photon, it could be deduced from the energy, as corresponding to a "mass at rest" in the traditional picture, however this not would be compatible with a propagation at the speed of light. The usual theoretical justification for the null mass of the photon is that a non null mass would violate the Maxwell's laws. Actually this reasoning is based on a model with an interacting term where the photon is represented as the potential, with its affine transformation law (see Guidry p.81) but, as we have seen, the right representation of boson avoids this issue. Actually the idea to give a mass to the photon comes from the want to agree with its assumed momentum. But the definition of the mass of a particle is somewhat conventional. For particles it comes from the inertial spinor, and more precisely from the inertial vector. We have nothing equivalent here, and there is no need to invent a mass for the photon, null or otherwise.

8.3.3 Extension to gravitons

Gravitons should have properties similar to the photon, as they propagate also at the speed of light. Because they have never been observed, any idea on the subject is speculative.

There would be 3 kinds of gravitons, and 3 kinds of antigraviton, with Spin 1.

The equations for the energy is :

$\delta E_{Field} = -8C_G \Delta\Gamma\left(J_G\right) = -\frac{1}{4} C_I \delta K$

with the currents J_G associated to the gravitational field.

The quantity $\Delta\Gamma\left(J_G\right) = \sum_{\beta=0}^{3} \left\{ \sum_{a=1}^{3} \left(\Delta\Gamma_{\beta}^{a} J_G^{a,\beta} \right) - \sum_{a=4}^{6} \left(\overline{\Delta\Gamma}_{\beta}^{a} J_G^{a,\beta} \right) \right\}$

so gravitons and antigravitons should have opposite interactions with particles, as it appears also in the Energy-Momentum tensor.

All that has been said about periodic motions represented by matter fields holds also the gravitational field. In a star system the orbits of planets are organized in a configuration which is usually stable. A change of orbit occurring in a short period of time, as in a collision, would result in a significant discontinuity in the gravitational field. Gravitons and anti-gravitons would propagate in opposite directions, so that, at a given point, only one kind would appear.

According to the tetrad equation the impact of a discontinuity in the gravitational field can be imparted to the field itself or to the metric. With a photon the choice seems obvious, however with a graviton it is less so. A discontinuous process is a transition between 2 states of equilibrium, and it is clear that in a phenomenon such as a collision in a star system the metric in the new state would not be the same. But, far away, the metric is defined by local conditions and only the local gravitational field would be affected. Which gives some credence to the idea of "gravitational waves". Notice that there is always an EM field present, and a discontinuity of the gravitational field could also, through the tetrad equation, create a discontinuity in the EM field, that is the emission of photons of strong energy. This could be one explanation for the bursts of X or gamma rays.

8.4 QUANTUM THEORY OF FIELDS

Whenever the weak or strong interactions are considered, due to their very small range, the focus is on the interactions between fermions and bosons, in which the particles usually do not keep their fundamental state. QTF is a Theory of its own, and encompasses many topics, using several tools and models. We will just give an overview of them and see how they can be consistent with our picture.

8.4.1 Micro-Systems interacting

Lagrangian

In the 2nd continuous model, with individual particles, the equations for the particles involve only the value of the potentials at the location of the particle. The location (in M) of the events is not involved : only the value of the maps. So, when the topic is focused on the particles (without consideration of the propagation of the fields in the vacuum) and on the transitions between states, it makes sense to use a model where the fields are themselves represented as bosons.

Bosons interact with particles when they meet : the interaction occurs at a point m_0 common to their respective trajectory. Bosons act on particles through the same mechanism as the fields, and according to the charge of the particles. All known elementary particles, except the neutrinos, have an electric charge, so they interact with photons.

The action of a boson on the momentum of a particle with velocity U is :

$\delta\psi_B = \vartheta(\sigma,\varkappa)[\psi_0]\left[Ad_{\varkappa}\left(\sum_{\alpha=0}^{3} U^{\alpha}\Delta\grave{A}_{\alpha}\right)\right]$ with bosons

If the particle keeps its fundamental state ψ_0 its momentum changes as :

$\vartheta\left(\upsilon\left(\widetilde{X}_r,\widetilde{X}_w\right)\cdot\widetilde{\sigma},\varkappa\right)\widetilde{\psi} = \vartheta(\upsilon(X_r,X_w)\cdot\sigma,\varkappa)\psi + \delta\psi_B$

Then, as in a collision, we can assume that the particle has a continuous motion before and after interacting with the boson. So $\psi,\widetilde{\psi}$ follow the same PDE, the only difference is the initial condition, defined at m_0 for $\widetilde{\psi}$.

The change of kinetic energy is : $\delta K = \frac{1}{M_p}\frac{1}{i}\langle\psi,\delta\psi\rangle$, so it is not necessarily equal to the energy of the boson, and different outcomes are possible. The boson can disappear, or it can survive the encounter, but with a loss of energy. Then it follows one of the propagation curve of the field emanating from the particle.

The variables $\psi,\widehat{\grave{A}},\widehat{G},\Delta\grave{A},\Delta G$, considered as maps on $\mathbb{R}$ valued in the respective vector spaces, can be introduced in a lagrangian, either as "bosonic fields", similar to "ferminonic fields" (similar to the first model), or as isolated bosons on their trajectories (similar to the second model). This is actually the lagrangian of the standard model. However in QTF bosons are introduced with the same format as the potential, and its affine transformation law in a change of gauge. This complication, and the use of the Dirac's operator, require the introduction of the Higgs boson. In our picture one can build a lagrangian including bosons without the need for the Higgs boson.

Micro-Systems interacting

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. We can consider micro systems, comprising one particle, the fields and possibly one boson represented by the variables which enter the lagrangian : $\psi,\mathcal{F}_A,\mathcal{F}_G,\widehat{\grave{A}},\widehat{G},B,\Gamma$. These variables are seen as fermionic, bosonic or force fields but, because their location does not matter, this is the value in the associated vector spaces which is considered. If the conditions of the theorem 29 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.

For a system of N such microsystems, where 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 32 and 34. 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

$\psi, \widehat{\grave{A}}, \widehat{G}, \Delta\grave{A}, \Delta G$) 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.

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 and is represented by a variable associated to a vector of $T_1 Spin(3)$. The action of $s \in Spin(3)$ and $-s \in Spin(3)$ 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 microsystems 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

So far, in all the models the number of interacting particles (fermions or bosons) is constant, and the particles keep their fundamental state. The main topic of QTF is the study of events when these rules do not hold any longer : creation or annihilation of particles. Each of these events occur at a given location and involve discontinuous processes, the fields are represented as bosons, so we have a system consisting of a variable number of micro-systems, represented by vectors in some Hilbert space H, interacting.

Interacting microsystems are represented in the tensorial product of H, but because their number is not fixed, we need to consider the Fock space, defined as $\mathcal{F} = \oplus_{k=0}^{\infty} (\otimes_k H)$ (Maths.12.5.8). 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 quantization). However the creation and annihilation operators have extensions with specific commutation rules :

Canonical commutation rules (CCR) in the Bose space:
$[a_+(u), a_+(v)] = \left[a_+^*(u), a_+^*(v)\right] = 0$
$\left[a_+(u), a_+^*(v)\right] = \langle u, v \rangle 1$

Canonical anticommutation rules (CAR) in the Fermi space :
$\{a_+(u), a_+(v)\} = \left\{a_-^*(u), a_-^*(v)\right\} = 0$
$\left\{a_+(u), a_+^*(v)\right\} = \langle u, v \rangle 1$
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. There is nothing equivalent for the bosons.

8.4.2 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, so the law of evolution is simple when Θ 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 $\mathbb{R}^m$: $\frac{dX}{dt} = \Theta(t) X(t)$:

$$X(t) = \lim_{n\to\infty} \left(\prod_{p=0}^{n} \exp(t_{p+1} - t_p)\Theta(t_p) \right) X(0)$$

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 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 last two cases a trajectory is computed as a path.

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 no 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,...

Chapter 9

CONCLUSION

At the end of this book I hope that the reader has a better understanding of how Theoretical Physics, encompassing the most advanced topics, can be grounded in a deep understanding of the usual concepts and First Principles, with the use of the adequate mathematical tools. Group Representations, Clifford Algebras, Fiber Bundles, Connections, jet prolongations, are a bit abstract, but well suited, and quite efficient to address the issues of modern Physics. The many tools presented (such as the operators j,σ, the complex representation, the decomposition $\mathcal{F}^r, \mathcal{F}^w$, the charts with r, w...) make manageable the problems in RG, without the usual assumptions, made essentially to simplify the computations and not fully physically justified. I hope also to have brought some clarification on Quantum Mechanics, Relativity and gauge theories, as well as on ideas, 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. They 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 an 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 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. We have also shown the necessity to review the concept of motion, incorporating translation and rotations, leading to the general assumption of the existence of a tetrad attached to all material bodies, at any scale, which adds relief to the Geometry of RG. 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. And they are the natural, and necessary, framework to represent the motion of material bodies. They can be used in SR, or even in Galilean Geometry.

3. The enlargement of the concept of motion leads naturally to revisit the concept of momentum.

The framework given in the Chapter 4 is actually the natural prolongation of Classic Mechanics, when the adjustments required by the Geometry of General Relativity are accounted for. They lead to a sound definition of the Spinors, give a clear meaning to the Spin and the introduction of anti-particles. 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 misses 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 results presented in the Chapter 5 are significant, and should be useful for understanding the propagation of the gravitational field. They provide a general specification for the metric which should be precious in all studies involving GR.

5. In the Chapter 6 we have presented the different issues in the implementation of the Principle of Least Action. We have proven that, in the most general lagrangian, 6 variables suffice, but others, and notably the potentials, are excluded. We have given a strong mathematical backing to the functional derivatives calculus, based of an original theory of distributions on vector bundles. This is the key to understand the meaning of the Energy-Momentum tensor. We have shown that the tetrad equation, in the most general context, is equivalent to the conservation of energy, which emphasizes the role of the metric.

6. The two models presented are essentially examples of how the theory of Lagrangians can be used practically. They are the starting point for the concepts of currents. Important theorems have been proven, such that non relativist particles follow geodesics, and we have provided guidelines which can be used to find explicit solutions of the most general problems, in particular for the computation of the fields and of the metric.

7. The idea of bosons as discontinuities in the fields seems more speculative. But it is clear that one cannot reconcile the concepts of localized material bodies and continuous force fields without some discontinuities. The common answer of the two Physics, based on a totally discreet and random vision of the world, on one hand, and a continuous classic and practical Physics on the other hand, lacks both of ambition and imagination. As it has been done on the other topics a deeper understanding of the concepts, and the extension of the well known phenomenon of discontinuities in a continuous medium give a natural solution. The presentation leaves some gaps, which are due to our limited knowledge of the propagation of weak and strong interactions and the non observation of gravitons.

There are some new results in this book : QM, deformable solids in RG, spinors, motion of material bodies, propagation of fields, 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 the concepts of Thermodynamics, 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 most important topic should be 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 the photon, the only well known boson, shows distinct properties than the other bosons. 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 venue than to fight against GR. The gravitational field shows, in all its aspects, two components, and it seems logical that it has a cosmological interpretation : it would be the engine which moves matter on its world line. Both components have different 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 for 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 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* (2016) Amazon E-book

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)

Jan Slovak *Smooth structures on fiber jet spaces* Czechoslovak Mathematical Journal 36 (111) (1986) Praha

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* University of 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)

H.White, P.March, J.Lawrence, J.Vera, A.Sylvester, D.Brady, P.Bailey *Measurement of impulsive thrust from a closed radio-frequency cavity in vacuum at Nasa Johnson Space Center* Journal of propulsion and power (2016)

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). $\varepsilon_5 = \varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$

The numerous formulas involving the operator j are given 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) :

$\upsilon(r,w) \cdot \upsilon(r',w') = \frac{1}{4}\left(w^t w' - r^t r'\right) - \frac{1}{4}\left(w^t r' + r^t w'\right)\varepsilon_5$
$+\frac{1}{4}\left(-r^3 w'^2 + r^2 w'^3 + w^2 r'^3 - w^3 r'^2\right)\varepsilon_0\varepsilon_1 + \frac{1}{4}\left(r^3 w'^1 - w^1 r'^3 + w^3 r'^1 - r^1 w'^3\right)\varepsilon_0\varepsilon_2$
$+\frac{1}{4}\left(r^1 w'^2 - r^2 w'^1 + w^1 r'^2 - w^2 r'^1\right)\varepsilon_0\varepsilon_3$
$+\frac{1}{4}\left(w^2 w'^1 - w^1 w'^2 + r^1 r'^2 - r^2 r'^1\right)\varepsilon_2\varepsilon_1 + \frac{1}{4}\left(-w^3 w'^1 + w^1 w'^3 - r^1 r'^3 + r^3 r'^1\right)\varepsilon_1\varepsilon_3$
$+\frac{1}{4}\left(w^3 w'^2 - w^2 w'^3 + r^2 r'^3 - r^3 r'^2\right)\varepsilon_3\varepsilon_2$
$\upsilon(r,w) \cdot \upsilon(r',w') = \frac{1}{4}\left(w^t w' - r^t r'\right) + \frac{1}{2}\upsilon\left(j(r)r' - j(w)w', j(w)r' + j(r)w'\right) - \frac{1}{4}\left(w^t r' + r^t w'\right)\varepsilon_5$

From there the bracket on the Lie algebra :

$[\upsilon(r,w), \upsilon(r',w')] = \upsilon(r,w) \cdot \upsilon(r',w') - \upsilon(r',w') \cdot \upsilon(r,w)$

$$[\upsilon(r,w), \upsilon(r',w')] = \upsilon\left(j(r)r' - j(w)w', j(w)r' + j(r)w'\right) \tag{A.1}$$

With signature (1,3) :

$\upsilon(r,w) \cdot \upsilon(r',w') = \frac{1}{4}\left(w^t w' - r^t r'\right) - \frac{1}{2}\upsilon\left(-j(r)r' + j(w)w', j(w)r' + j(r)w'\right) - \frac{1}{4}\left(w^t r' + r^t w'\right)\varepsilon_5$

From there the bracket on the Lie algebra :

$$[\upsilon(r,w), \upsilon(r',w')] = -\upsilon\left(j(r)r' - j(w)w', j(w)r' + j(r)w'\right) \tag{A.2}$$

More over, with both signatures : $\upsilon(x,y) \cdot \varepsilon_5 = \varepsilon_5 \cdot \upsilon(x,y) = \upsilon(-y,x)$

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

$$s = a + \upsilon(r,w) + b\varepsilon_5 \tag{A.3}$$

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.

$(a + \upsilon(r,w) + b\varepsilon_5) \cdot (a' + \upsilon(r',w') + b'\varepsilon_5)$

$= aa' - bb' + \upsilon(a'r + ar' - bw' - b'w, a'w + aw' + br' + b'r) + \upsilon(r,w) \cdot \upsilon(r',w') + (a'b + ab')\varepsilon_5$

i) With signature (3,1)

$\upsilon(r,w) \cdot \upsilon(r',w') = \frac{1}{4}(w^t w' - r^t r') + \frac{1}{2}\upsilon(j(r)r' - j(w)w', j(w)r' + j(r)w') - \frac{1}{4}(w^t r' + r^t w')\varepsilon_5$

$(a + \upsilon(r,w) + b\varepsilon_5) \cdot (a' + \upsilon(r',w') + b'\varepsilon_5) = a" + \upsilon(r", w") + b"\varepsilon_5$

$a" = aa' - bb' + \frac{1}{4}(w^t w' - r^t r')$

$b" = (a'b + ab') - \frac{1}{4}(w^t r' + r^t w')$

$r" = a'r + ar' - bw' - b'w + \frac{1}{2}(j(r)r' - j(w)w')$

$w" = a'w + aw' + br' + b'r + \frac{1}{2}(j(w)r' + j(r)w')$

ii) With signature (1,3)

$\upsilon(r,w) \cdot \upsilon(r',w') = \frac{1}{4}(w^t w' - r^t r') - \frac{1}{2}\upsilon(-j(r)r' + j(w)w', j(w)r' + j(r)w') - \frac{1}{4}(w^t r' + r^t w')\varepsilon_5$

$(a + \upsilon(r,w) + b\varepsilon_5) \cdot (a' + \upsilon(r',w') + b'\varepsilon_5) = a" + \upsilon(r", w") + b"\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$

$a" = aa' - bb' + \frac{1}{4}(w^t w' - r^t r')$

$b" = (a'b + ab') - \frac{1}{4}(w^t r' + r^t w')$

$r" = a'r + ar' - bw' - b'w + a'r + ar' - bw' - b'w + \frac{1}{2}(j(r)r' - j(w)w')$

$w" = a'w + aw' + br' + b'r - \frac{1}{2}(j(w)r' + j(r)w')$

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. So 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 $Cl(1,3)$

$$(a + \upsilon(r,w) + b\varepsilon_5)^{-1} = a - \upsilon(r,w) + b\varepsilon_5 \tag{A.4}$$

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 :
$a"=a^2-b^2+\frac{1}{4}(-w^tw+r^tr)=1$
$b"=ab+ba-\frac{1}{4}(-w^tr-r^tw)=0$
$r"=\frac{1}{2}(-j(r)r+j(w)w)+ar-ar-bw+bw=0$
$w"=\frac{1}{2}(-j(w)r-j(r)w)+aw-aw+br-br=0$
$a^2-b^2=1+\frac{1}{4}(w^tw-r^tr)$
So, for any element : $a+\upsilon(r,w)+b\varepsilon_5$ we have :

$$a^2-b^2=1+\frac{1}{4}(w^tw-r^tr) \tag{A.5}$$

$$ab=-\frac{1}{4}r^tw \tag{A.6}$$

and if we keep only 6 free parameters, a,b are defined from r,w, up to sign, with the conditions:
i) $r^tw\neq 0: b=-\frac{1}{4a}r^tw$
$a^2=\frac{1}{2}\left(\left(1+\frac{1}{4}(w^tw-r^tr)\right)+\sqrt{\left(1+\frac{1}{4}(w^tw-r^tr)\right)^2+\frac{1}{4}(r^tw)^2}\right)$
ii) $r^tw=0$:
$(w^tw-r^tr)\geq -4: a=\epsilon\sqrt{1+\frac{1}{4}(w^tw-r^tr)}; b=0$
$(w^tw-r^tr)\leq -4: b=\epsilon\sqrt{-\left(1+\frac{1}{4}(w^tw-r^tr)\right)}; a=0$
So :
if $r=0$ then : $s=\epsilon\sqrt{1+\frac{1}{4}w^tw}+\upsilon(0,w)$
if $w=0$ then
$r^tr\leq 4: s=\epsilon\sqrt{1-\frac{1}{4}r^tr}+\upsilon(r,0)$
$r^tr\geq 4: s=\upsilon(r,0)+\epsilon\sqrt{\frac{1}{4}r^tr-1}\varepsilon_5$

2. With signature (1,3)

We get the same relations.

A.1.3 Adjoint map

The adjoint map : $\mathbf{Ad}: Spin(3,1)\times Cl(3,1)\to Cl(3,1) :: \mathbf{Ad}_s X=s\cdot X\cdot s^{-1}$ is expressed differently when it acts on vectors or elements of the Lie algebra $T_1 Spin(3,1)$.

Action on vectors of F

A straightforward computation gives the following results :
$\forall X\in F, s\in Spin(3,1): \mathbf{Ad}_s X=s\cdot X\cdot s^{-1}$
$X=X_0\varepsilon_0+X_1\varepsilon_1+X_2\varepsilon_2+X_3\varepsilon_3$
$s=a+\upsilon(r,w)+b\varepsilon_5$
$\mathbf{Ad}_s X=(a+\upsilon(r,w)+b\varepsilon_5)\cdot X\cdot(a-\upsilon(r,w)+b\varepsilon_5)$
$=a^2X+ab(X\cdot\varepsilon_5+\varepsilon_5\cdot X)+b^2\varepsilon_5\cdot X\cdot\varepsilon_5+a(\upsilon(r,w)\cdot X-X\cdot\upsilon(r,w))$
$+b(\upsilon(r,w)\cdot X\cdot\varepsilon_5-\varepsilon_5\cdot X\cdot\upsilon(r,w))-\upsilon(r,w)\cdot X\cdot\upsilon(r,w)$
$X\cdot\varepsilon_5=-X_0\varepsilon_1\varepsilon_2\varepsilon_3-X_1\varepsilon_0\varepsilon_2\varepsilon_3+X_2\varepsilon_0\varepsilon_1\varepsilon_3-X_3\varepsilon_0\varepsilon_1\varepsilon_2$
$\varepsilon_5\cdot X=X_0\varepsilon_1\varepsilon_2\varepsilon_3+X_1\varepsilon_0\varepsilon_2\varepsilon_3-X_2\varepsilon_0\varepsilon_1\varepsilon_3+X_3\varepsilon_0\varepsilon_1\varepsilon_2$
$X\cdot\varepsilon_5+\varepsilon_5\cdot X=0$
$\varepsilon_5\cdot X\cdot\varepsilon_5=-X\varepsilon_5\varepsilon_5=X$

$\upsilon(r,w)\cdot X\cdot\varepsilon_5-\varepsilon_5\cdot X\cdot\upsilon(r,w)=-\upsilon(-w,r)\cdot X+X\cdot\upsilon(-w,r)$
$\mathbf{Ad}_sX=\left(a^2+b^2\right)X+a\left(\upsilon(r,w)\cdot X-X\cdot\upsilon(r,w)\right)$
$-b\left(\upsilon(-w,r)\cdot X-X\cdot\upsilon(-w,r)\right)-\upsilon(r,w)\cdot X\cdot\upsilon(r,w)$
$2\upsilon(r,w)\cdot X$
$=X_0\left(y_1\varepsilon_1+y_2\varepsilon_2+y_3\varepsilon_3-x_3\varepsilon_0\varepsilon_1\varepsilon_2+x_2\varepsilon_0\varepsilon_1\varepsilon_3-x_1\varepsilon_0\varepsilon_2\varepsilon_3\right)$
$+X_1\left(y_1\varepsilon_0-y_2\varepsilon_0\varepsilon_1\varepsilon_2-y_3\varepsilon_0\varepsilon_1\varepsilon_3+x_3\varepsilon_2-x_2\varepsilon_3-x_1\varepsilon_1\varepsilon_2\varepsilon_3\right)$
$+X_2\left(y_1\varepsilon_0\varepsilon_1\varepsilon_2+y_2\varepsilon_0-y_3\varepsilon_0\varepsilon_2\varepsilon_3-x_3\varepsilon_1-x_2\varepsilon_1\varepsilon_2\varepsilon_3+x_1\varepsilon_3\right)$
$+X_3\left(y_1\varepsilon_0\varepsilon_1\varepsilon_3+y_2\varepsilon_0\varepsilon_2\varepsilon_3+y_3\varepsilon_0-x_3\varepsilon_1\varepsilon_2\varepsilon_3+x_2\varepsilon_1-x_1\varepsilon_2\right)$
$2X\cdot\upsilon(r,w)$
$=X_0\left(-y_1\varepsilon_1-y_2\varepsilon_2-y_3\varepsilon_3-x_3\varepsilon_0\varepsilon_1\varepsilon_2+x_2\varepsilon_0\varepsilon_1\varepsilon_3-x_1\varepsilon_0\varepsilon_2\varepsilon_3\right)$
$+X_1\left(-y_1\varepsilon_0-y_2\varepsilon_0\varepsilon_1\varepsilon_2-y_3\varepsilon_0\varepsilon_1\varepsilon_3-x_3\varepsilon_2+x_2\varepsilon_3-x_1\varepsilon_1\varepsilon_2\varepsilon_3\right)$
$+X_2\left(y_1\varepsilon_0\varepsilon_1\varepsilon_2-y_2\varepsilon_0-y_3\varepsilon_0\varepsilon_2\varepsilon_3+x_3\varepsilon_1-x_2\varepsilon_1\varepsilon_2\varepsilon_3-x_1\varepsilon_3\right)$
$+X_3\left(y_1\varepsilon_0\varepsilon_1\varepsilon_3+y_2\varepsilon_0\varepsilon_2\varepsilon_3-y_3\varepsilon_0-x_3\varepsilon_1\varepsilon_2\varepsilon_3-x_2\varepsilon_1+x_1\varepsilon_2\right)$

$$\left(\upsilon(r,w)\cdot X-X\cdot\upsilon(r,w)\right)=X_0w+\left(w^tx\right)\varepsilon_0+j(r)x \tag{A.7}$$

$[h(s)]=$
$$\begin{bmatrix} a^2+b^2+\frac{1}{4}\left(r^tr+w^tw\right) & aw^t-br^t+\frac{1}{2}w^tj(r) \\ aw-br+\frac{1}{2}j(r)w & a^2+b^2+\frac{1}{4}\left(r^tr+w^tw\right)+aj(r)+bj(w)+\frac{1}{2}\left(j(r)j(r)+j(w)j(w)\right) \end{bmatrix}$$

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 :
$\mathbf{Ad}_g\upsilon(x,y)=\left(a+\upsilon(r,w)+b\varepsilon_5\right)\cdot\upsilon(x,y)\cdot\left(a-\upsilon(r,w)+b\varepsilon_5\right)$
$=\upsilon\{\left[a^2-b^2+aj(r)-bj(w)\right]x-\left[2ab+aj(w)+bj(r)\right]y,$
$\left[2ab+aj(w)+bj(r)\right]x+\left[a^2-b^2+aj(r)-bj(w)\right]y\}-\upsilon(r,w)\upsilon(x,y)\upsilon(r,w)$
with
$\upsilon(x,y)\varepsilon_5=\varepsilon_5\upsilon(x,y)=\upsilon(-y,x)$
$\varepsilon_5\upsilon(x,y)\varepsilon_5=-\upsilon(x,y)$
$\upsilon(r,w)\cdot\upsilon(x,y)\cdot\upsilon(r,w)$
$=\frac{1}{2}\upsilon\{\left(j(w)j(w)-j(r)j(r)+2\left(a^2-b^2-1\right)\right)x+\left(j(r)j(w)+j(w)j(r)-4ab\right)y,$
$-\left(j(r)j(w)+j(w)j(r)-4ab\right)x+\left(j(w)j(w)-j(r)j(r)+2\left(a^2-b^2-1\right)\right)y\}$
$\mathbf{Ad}_g\upsilon(x,y)$
$=\upsilon\{\left[1+aj(r)-bj(w)+\frac{1}{2}\left(j(r)j(r)-j(w)j(w)\right)\right]x-\left[aj(w)+bj(r)+\frac{1}{2}\left(j(r)j(w)+j(w)j(r)\right)\right]y,$
$\left[aj(w)+bj(r)+\frac{1}{2}\left(j(r)j(w)+j(w)j(r)\right)\right]x+\left[1+aj(r)-bj(w)-\frac{1}{2}\left(j(w)j(w)-j(r)j(r)\right)\right]y\}$
$[\mathbf{Ad}_g]=$
$$\begin{bmatrix} 1+aj(r)-bj(w)+\frac{1}{2}\left(j(r)j(r)-j(w)j(w)\right) & -\left(aj(w)+bj(r)+\frac{1}{2}\left(j(r)j(w)+j(w)j(r)\right)\right) \\ aj(w)+bj(r)+\frac{1}{2}\left(j(r)j(w)+j(w)j(r)\right) & 1+aj(r)-bj(w)+\frac{1}{2}\left(j(r)j(r)-j(w)j(w)\right) \end{bmatrix}$$
With $s_w=a_w+\upsilon(0,w)$
$$[\mathbf{Ad}_s]=\begin{bmatrix} \left[1-\frac{1}{2}j(w)j(w)\right] & -\left[a_wj(w)\right] \\ \left[a_wj(w)\right] & \left[1-\frac{1}{2}j(w)j(w)\right] \end{bmatrix}$$
With $s_r=a_r+\upsilon(r,0)$
$$[\mathbf{Ad}_s]=\begin{bmatrix} \left[1+a_rj(r)+\frac{1}{2}j(r)j(r)\right] & 0 \\ 0 & \left[1+a_rj(r)+\frac{1}{2}j(r)j(r)\right] \end{bmatrix}$$

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 with 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 105 *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 :

$s_r = a + \upsilon(r, w) + b\varepsilon_5$

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 $E = \{s \in Spin(3,1) : s \cdot \varepsilon_0 = \varepsilon_0 \cdot s\} = Spin(3)$

$s_r = a + \upsilon(r, w) + b\varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3$

If $s_r \cdot \varepsilon_0 = \varepsilon_0 \cdot s_r$

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 106 *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 $[s] = \left\{s' = s \cdot \left(\epsilon\sqrt{1 - \frac{1}{4}\rho^t\rho} + \upsilon(\rho,0)\right), \rho^t\rho \le 4\right\}$

i) In $Spin(3,1)$:

$s = a + \upsilon(r,w) + b\varepsilon_5$

$s' = a' + \upsilon(r',w') + b'\varepsilon_5$

$a' = a\epsilon\sqrt{1 - \frac{1}{4}\rho^t\rho} - \frac{1}{4}r^t\rho$

$b' = b\epsilon\sqrt{1 - \frac{1}{4}\rho^t\rho} - \frac{1}{4}w^t\rho$

$r' = \frac{1}{2}j(r)\rho + r\epsilon\sqrt{1 - \frac{1}{4}\rho^t\rho} + a\rho$

$w' = \frac{1}{2}j(w)\rho + w\epsilon\sqrt{1 - \frac{1}{4}\rho^t\rho} + b\rho$

$a^2 - b^2 = 1 + \frac{1}{4}\left(w^t w - r^t r\right)$

$ab = -\frac{1}{4}r^t w$

ii) We can always choose in the class an element s' such that : $r' = 0$. It requires : $\left(\frac{1}{2}j(r) + aI\right)\rho = -r\epsilon\sqrt{1 - \frac{1}{4}\rho^t\rho}$

$x = \left(\frac{1}{a} - \frac{b}{a^2 + b^2 r^t r}j(r) - \frac{b^2}{a(a^2 + b^2 r^t r)}j(r)j(r)\right)y$

This linear equation in ρ has always a unique solution :

$\rho = -\epsilon\sqrt{1 - \frac{1}{4}\rho^t\rho}\frac{1}{a}r$

$\rho^t\rho = \left(1 - \frac{1}{4}\rho^t\rho\right)\frac{1}{a^2}\left(r^t r\right) \Rightarrow$

$\left(a^2 + \frac{1}{4}\left(r^t r\right)\right)\rho^t\rho = \left(r^t r\right)$

$\rho^t\rho = \frac{4(r^t r)}{4a^2+(r^t r)} \leq 4$

$\sqrt{1-\frac{1}{4}\rho^t\rho} = \sqrt{\frac{4a^2}{4a^2+r^t r}} = \frac{2a}{\sqrt{4a^2+r^t r}}$

$\rho = -\epsilon\frac{2}{\sqrt{4a^2+r^t r}}r$

$\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho} + \upsilon(\rho,0) = \epsilon\frac{2a}{\sqrt{4a^2+r^t r}} - \upsilon\left(\epsilon\frac{2}{\sqrt{4a^2+r^t r}}r, 0\right) = \epsilon\left(\frac{2a}{\sqrt{4a^2+r^t r}} - \upsilon\left(\frac{2}{\sqrt{4a^2+r^t r}}r, 0\right)\right)$

$a' = a\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho} - \frac{1}{4}r^t\rho = \frac{1}{2}\frac{\epsilon}{\sqrt{4a^2+r^t r}}\left(4a^2+r^t r\right) = \frac{1}{2}\epsilon\sqrt{4a^2+r^t r}$

$w' = \frac{1}{2}j(w)\rho + w\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho} + b\rho = \epsilon\frac{2}{\sqrt{4a^2+r^t r}}\left(\frac{1}{2}j(r)w + aw - br\right)$

$b' = b\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho} - \frac{1}{4}w^t\rho = \epsilon\frac{2}{\sqrt{4a^2+r^t r}}\left(ab + \frac{1}{4}w^t r\right) = 0$

$s' = s_w = \frac{1}{2}\epsilon\sqrt{4a^2+r^t r} + \upsilon\left(0, \epsilon\frac{2}{\sqrt{4a^2+r^t r}}\left(\frac{1}{2}j(r)w + aw - br\right)\right)$

$= \epsilon\left(\frac{1}{2}\sqrt{4a^2+r^t r} + \upsilon\left(0, \frac{2}{\sqrt{4a^2+r^t r}}\left(\frac{1}{2}j(r)w + aw - br\right)\right)\right)$

$s' = s\cdot\left(\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho} + \upsilon(\rho,0)\right)$

$s = s'\cdot\left(\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho} + \upsilon(\rho,0)\right)^{-1} = s_w.\left(\epsilon\sqrt{1-\frac{1}{4}\rho^t\rho} - \upsilon(\rho,0)\right)$

$= \epsilon\left(\frac{1}{2}\sqrt{4a^2+r^t r} + \upsilon\left(0, \frac{2}{\sqrt{4a^2+(r^t r)}}\left(\frac{1}{2}j(r)w + aw - br\right)\right)\right)\cdot\epsilon\left(\frac{2a}{\sqrt{4a^2+r^t r}} + \upsilon\left(\frac{2}{\sqrt{4a^2+(r^t r)}}r, 0\right)\right)$

$s = a + \upsilon(r,w) + b\varepsilon_5 = s_w\cdot s_r$

iii) In $Cl(1,3)$ we have the same decomposition with the same components.

$s = a + \upsilon(r,w) + b\varepsilon_5 = s_w\cdot s_r$

$r" = \frac{1}{2}\epsilon\sqrt{4a^2+r^t r}\epsilon\frac{2}{\sqrt{4a^2+(r^t r)}}r = r$

$w" = \frac{1}{2}j\left(\left(\epsilon\sqrt{4a^2+r^t r}\right)\epsilon\frac{2}{4a^2+(r^t r)}\left(\frac{1}{2}j(r)w + aw - br\right)\right)\left(\epsilon\frac{2}{\sqrt{4a^2+(r^t r)}}\right)r$

$+\left(\epsilon\frac{2}{\sqrt{4a^2+(r^t r)}}\right)a\left(\epsilon\sqrt{4a^2+r^t r}\right)\epsilon\frac{2}{4a^2+(r^t r)}\left(\frac{1}{2}j(r)w + aw - br\right)$

$= 2j\left(\epsilon\frac{1}{4a^2+(r^t r)}\left(\frac{1}{2}j(r)w + aw - br\right)\right)r + a\epsilon\frac{4}{4a^2+(r^t r)}\left(\frac{1}{2}j(r)w + aw - br\right)$

$= \left(\epsilon\frac{2}{4a^2+(r^t r)}\right)\left(\frac{1}{2}j(j(r)w)r - aj(w)r + aj(r)w + 2a^2 w - 2abr\right)$

$= \left(\epsilon\frac{2}{4a^2+(r^t r)}\right)\left(\frac{1}{2}(wr^t - rw^t)r + 2a^2 w + \frac{1}{2}(r^t w)r\right)$

$= \left(\epsilon\frac{2}{4a^2+(r^t r)}\right)\left(\frac{1}{2}w(r^t r) - \frac{1}{2}r(w^t r) + 2a^2 w + \frac{1}{2}(r^t w)r\right)$

$= \left(\epsilon\frac{1}{4a^2+(r^t r)}\right)\left(\left(4a^2 + (r^t r)\right)w\right) = w$ ■

$a^2 - b^2 = 1 + \frac{1}{4}(w^t w - r^t r) \Rightarrow 4a^2 + r^t r = 4 + w^t w + 4b^2$

So any element of $Spin(3,1)$ can be written uniquely (up to sign) :

$s = a + \upsilon(r,w) + b\varepsilon_5 = \epsilon s_w\cdot\epsilon s_r = \epsilon(a_w + \upsilon(0,w_w))\cdot\epsilon(a_r + \upsilon(0,r_r))$

$s_w = a_w + \upsilon(0,w_w) = \frac{1}{2}\sqrt{4a^2+r^t r} + \upsilon\left(0, \frac{2}{\sqrt{4a^2+r^t r}}\left(\frac{1}{2}j(r)w + aw - br\right)\right)$

$s_r = (a_r + \upsilon(r_r,0)) = \frac{2a}{\sqrt{4a^2+r^t r}} + \upsilon\left(\frac{2}{\sqrt{4a^2+(r^t r)}}r, 0\right)$

$\epsilon a_r a_w a > 0$

Remark : the elements $\pm s_w$ are equivalent :

$(a_w + \upsilon(0,w_w)) \sim -(a_w + \upsilon(0,w_w))$

Take $s_r = -1 \in Spin(3) : -s_w = s_w\cdot s_r$

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.

Decomposition of the Lie algebra

To each Clifford bundle $Cl(3)$ is associated a unique Lie algebra $T_1 Spin(3)$ which is a subset of $Cl(3)$ and thus of $Cl(3,1)$. In any orthonormal basis an element of $T_1 Spin(3,1)$ reads :

$X = \upsilon(r,0) + \upsilon(0,w)$ and $\upsilon(r,0) \in T_1 Spin(3), \upsilon(0,w) \in T_1 SW$

The vectors r, w depends on the basis (they are components), however the elements $\upsilon(r,0), \upsilon(0,w) \in T_1 Spin(3,1)$ 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 :

$\theta(\varepsilon_0) : T_1 Spin(3,1) \to T_1 Spin(3,1) : \theta(\varepsilon_0)(X) = \varepsilon_0 \cdot X \cdot \varepsilon_0$

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 :

$L_0 = \{X \in T_1 Spin(3,1) : \theta(\varepsilon_0)(X) = -X\} = \{\upsilon(r,0), r \in \mathbb{R}^3\}$

$P_0 = \{X \in T_1 Spin(3,1) : \theta(\varepsilon_0)(X) = X\} = \{\upsilon(0,w), w \in \mathbb{R}^3\}$

$T_1 Spin(3,1) = L_0 \oplus P_0$

Thus L_0, P_0 and the decomposition depend only on the choice of ε_0 and $L_0 = T_1 Spin(3), P_0 \simeq T_1 SW$.

$\theta(\varepsilon_0)$ commutes with the action of the elements of $Spin(3)$:

$\forall s_r \in Spin(3), X \in T_1 Spin(3,1)$:

$\mathbf{Ad}_{s_r}\theta(\varepsilon_0)(X) = s_r \cdot \varepsilon_0 \cdot X \cdot \varepsilon_0 \cdot s_r^{-1} = \varepsilon_0 \cdot s_r \cdot X \cdot s_r^{-1} \cdot \varepsilon_0 = \theta(\varepsilon_0)(\mathbf{Ad}_{s_r}(X))$

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 $(\varepsilon_i)_{i=1}^3$.

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.

L_0, P_0 are globally invariant by $Spin(3)$, the scalar product is definite (positive or negative) and preserved by $\mathbf{Ad}$, so L_0, P_0 are 3 dimensional Hilbert spaces, and for each choice of ε_0 $(L_0, \mathbf{Ad}), (P_0, \mathbf{Ad})$ are 3 dimensional unitary representations of $Spin(3)$.

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 $Spin(3)$:

$\forall s_r \in Spin(3), X \in T_1 Spin(3,1)$:

$\pi_L(\varepsilon_0)(\mathbf{Ad}_{s_r}(X)) = \mathbf{Ad}_{s_r}(\pi_L(\varepsilon_0)(X))$

$\pi_P(\varepsilon_0)(\mathbf{Ad}_{s_r}(X)) = \mathbf{Ad}_{s_r}(\pi_P(\varepsilon_0)(X))$

$\theta(\varepsilon_0)$ 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 :

$$\|\upsilon(r,w)\| = \frac{1}{2}\sqrt{r^t r + w^t w} = \frac{1}{2}\sqrt{\langle \pi_L(X), \pi_L(X) \rangle_{Cl} - \langle \pi_P(X), \pi_P(X) \rangle_{Cl}} \tag{A.8}$$

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.

In any Lie algebra there is a bilinear symmetric form B called the Killing form,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 $so(3,1)$:

$B(\upsilon(r,w), \upsilon(r',w')) = 4(r^t r' - w^t w') = 16 \langle \upsilon(r,w), \upsilon(r',w') \rangle_{Cl}$

A.2 LIE DERIVATIVE

For the definition and properties of Lie derivatives see Maths.16.2.

Vector

A vector field $U \in \mathfrak{X}(TM)$ is Lie transported along the vector field $V \in \mathfrak{X}(TM)$ if the commutator (Maths.1437) :

$\pounds_V U = [V,U] = 0 = \sum_{\alpha,\beta=0}^{3} \left(V^\beta \partial_\beta U^\alpha - U^\beta \partial_\beta V^\alpha \right) \partial\xi_\alpha = -\pounds_U V$

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 :

$\pounds_V \lambda(m) = \sum_{a=1}^{m} \sum_{\{\alpha\beta\}} \pounds_V (\lambda_\alpha^a(m) d\xi^\alpha) \otimes \overrightarrow{\theta}_a$
Using the properties of the Lie derivative :
$\pounds_V (\lambda_\alpha^a(m) d\xi^\alpha)$
$= \sum_{\alpha=0}^{3} (\pounds_V \lambda_\alpha^a(m)) d\xi^\alpha + \lambda_\alpha^a(m) \pounds_V (d\xi^\alpha)$
$\pounds_V \lambda_\alpha^a(m) = \sum_{\gamma=0}^{3} (V^\gamma \partial_\gamma \lambda_\alpha^a)$
$\pounds_V (d\xi^\alpha) = i_V d(d\xi^\alpha) + d(i_V d\xi^\alpha) = dV^\alpha = \sum_\gamma \partial_\gamma V^\alpha d\xi^\gamma$
$\pounds_V (\lambda_\alpha^a(m) d\xi^\alpha) = \sum_{\alpha=0}^{3} \left(\sum_{\gamma=0}^{3} (V^\gamma \partial_\gamma \lambda_\alpha^a) \right) d\xi^\alpha + \lambda_\alpha^a \sum_{\gamma=0}^{3} \partial_\gamma V^\alpha d\xi^\gamma$
$= \sum_{\alpha=0}^{3} \left(\sum_{\gamma=0}^{3} (V^\gamma \partial_\gamma \lambda_\alpha^a) \right) d\xi^\alpha + \sum_{\alpha=0}^{3} \lambda_\gamma^a \sum_{\gamma=0}^{3} \partial_\alpha V^\gamma d\xi^\alpha$
$= \sum_{\alpha,\gamma=0}^{3} \left(V^\gamma \partial_\gamma \lambda_\alpha^a + \lambda_\gamma^a \partial_\alpha V^\gamma \right) d\xi^\alpha$

$$\pounds_V \lambda(m) = \sum_{a=1}^{m} \sum_{\alpha,\gamma=0}^{3} \left(V^\gamma \partial_\gamma \lambda_\alpha^a + \lambda_\gamma^a \partial_\alpha V^\gamma \right) d\xi^\alpha \otimes \overrightarrow{\theta}_a \tag{A.9}$$

2 form

The Lie derivative of a 2 form on M :$\mathcal{F}(m)=\sum_{a=1}^{m}\sum_{\{\alpha\beta\}}\mathcal{F}_{\alpha\beta}^{a}(m)\,d\xi^{\alpha}\wedge d\xi^{\beta}\otimes\overrightarrow{\theta}_{a}\in\Lambda_{2}(M;T_{1}U)$ valued in a 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 :

$$\pounds_{V}\mathcal{F}(m)=\sum_{a=1}^{m}\left(\sum_{\{\alpha\beta\}}\sum_{\gamma=0}^{3}V^{\gamma}\partial_{\gamma}\mathcal{F}_{\alpha\beta}^{a}d\xi^{\alpha}\wedge d\xi^{\beta}+\mathcal{F}_{\alpha\beta}^{a}\left(\partial_{\gamma}V^{\alpha}d\xi^{\gamma}\wedge d\xi^{\beta}+\partial_{\gamma}V^{\beta}d\xi^{\alpha}\wedge d\xi^{\gamma}\right)\right)\otimes\overrightarrow{\theta}_{a}\tag{A.10}$$

A straightforward computation gives :

$$\left[(\pounds_{V}\mathcal{F})^{r}\right]=\sum_{\gamma=0}^{3}V^{\gamma}\left[\partial_{\gamma}\mathcal{F}^{r}\right]+\left[\mathcal{F}^{aw}\right]j\left(\partial V^{0}\right)+\left[\mathcal{F}^{ar}\right]\left(-\left[\partial v\right]^{t}+(div(v))\,I_{3}\right)\tag{A.11}$$

$$\left[(\pounds_{V}\mathcal{F})^{w}\right]=\sum_{\gamma=0}^{3}V^{\gamma}\left[\partial_{\gamma}\mathcal{F}^{w}\right]+\left[\mathcal{F}^{aw}\right]\left(\partial_{0}V^{0}+\left[\partial v\right]\right)-\left[\mathcal{F}^{ar}\right]j\left(\partial_{0}V\right)\tag{A.12}$$

with :

$[\partial v]=\begin{bmatrix}\partial_{1}V^{1}&\partial_{2}V^{1}&\partial_{3}V^{1}\\\partial_{1}V^{2}&\partial_{2}V^{2}&\partial_{3}V^{2}\\\partial_{1}V^{3}&\partial_{2}V^{3}&\partial_{3}V^{3}\end{bmatrix}$

$[\partial_{0}V]=\begin{bmatrix}\partial_{0}V^{1}&\partial_{0}V^{2}&\partial_{0}V^{3}\end{bmatrix}$

$\left[\partial V^{0}\right]=\begin{bmatrix}\partial_{1}V^{0}\\\partial_{2}V^{0}\\\partial_{3}V^{0}\end{bmatrix}$

$\left[(\pounds_{V}\mathcal{F})^{r}\right]=\sum_{\gamma=0}^{3}V^{\gamma}\left[\partial_{\gamma}\mathcal{F}^{r}\right]+\left[\mathcal{F}^{aw}\right]j\left(\partial V^{0}\right)+\left[\mathcal{F}^{ar}\right]\left(-\left[\partial v\right]^{t}+(div(v))\,I_{3}\right)$

$\left[(\pounds_{V}\mathcal{F})^{w}\right]=\sum_{\gamma=0}^{3}V^{\gamma}\left[\partial_{\gamma}\mathcal{F}^{w}\right]+\left[\mathcal{F}^{aw}\right]\left(\partial_{0}V^{0}+\left[\partial v\right]\right)-\left[\mathcal{F}^{ar}\right]j\left(\partial_{0}V\right)$

Lie derivative of the metric

We have similarly :

$\pounds_{V}\left(\sum_{\alpha\beta}g_{\alpha\beta}(m)\,d\xi^{\alpha}\otimes d\xi^{\beta}\right)$

$=\sum_{\alpha\beta}\left(\pounds_{V}g_{\alpha\beta}(m)\right)d\xi^{\alpha}\otimes d\xi^{\beta}+g_{\alpha\beta}(m)\,\pounds_{V}\left(d\xi^{\alpha}\otimes d\xi^{\beta}\right)$

$=\sum_{\alpha\beta}\left(\sum_{\gamma=0}^{3}V^{\gamma}\partial_{\gamma}g_{\alpha\beta}d\xi^{\alpha}\otimes d\xi^{\beta}\right)+g_{\alpha\beta}\left((\pounds_{V}d\xi^{\alpha})\otimes d\xi^{\beta}+d\xi^{\alpha}\otimes\pounds_{V}d\xi^{\beta}\right)$

$\pounds_{V}(d\xi^{\alpha})=i_{V}d(d\xi^{\alpha})+d(i_{V}d\xi^{\alpha})=dV^{\alpha}=\sum_{\gamma}\partial_{\gamma}V^{\alpha}d\xi^{\gamma}$

$\pounds_{V}\left(d\xi^{\alpha}\otimes d\xi^{\beta}\right)=$

$\left(\sum_\gamma \partial_\gamma V^\alpha d\xi^\gamma\right) \otimes d\xi^\beta + d\xi^\alpha \otimes \sum_\gamma \partial_\gamma V^\beta d\xi^\gamma = \sum_\gamma \left(\partial_\gamma V^\alpha d\xi^\gamma \otimes d\xi^\beta + \partial_\gamma V^\beta d\xi^\alpha \otimes d\xi^\gamma\right)$
$\sum_{\alpha\beta} g_{\alpha\beta} \pounds_V \left(d\xi^\alpha \otimes d\xi^\beta\right) = \sum_{\alpha\beta} g_{\alpha\beta} \sum_\gamma \left(\partial_\gamma V^\alpha d\xi^\gamma \otimes d\xi^\beta + \partial_\gamma V^\beta d\xi^\alpha \otimes d\xi^\gamma\right)$
The Lie derivative is a symmetric tensor of 2nd order :
$\pounds_V \left(\sum_{\alpha\beta} g_{\alpha\beta}(m)\, d\xi^\alpha \otimes d\xi^\beta\right) = \sum_{\alpha\beta} \widetilde{g}_{\alpha\beta}(m)\, d\xi^\alpha \otimes d\xi^\beta$
$\widetilde{g}_{\alpha\beta}(m) = \sum_{\gamma=0}^3 V^\gamma \partial_\gamma g_{\alpha\beta} + g_{\gamma\beta}\partial_\alpha V^\gamma + g_{\alpha\gamma}\partial_\beta V^\gamma$

A.3 HODGE DUAL

(Maths.4.6.)

On M the Hodge dual of a scalar :
0 form (a function) : $*f = f\varpi_4$
1 form : $*\lambda_1 = *\left(\sum_{\alpha=0}^3 \lambda_\alpha d\xi^\alpha\right) = \sum_{\alpha=0}^3 (-1)^\alpha g^{\alpha\beta}\lambda_\beta d\xi^0 \wedge ...\widehat{d\xi^\alpha}... \wedge d\xi^3$
2 form :
$*\mathcal{F}^r = -\left(\mathcal{F}^{01} d\xi^3 \wedge d\xi^2 + \mathcal{F}^{02} d\xi^1 \wedge d\xi^3 + \mathcal{F}^{03} d\xi^2 \wedge d\xi^1\right) \det P'$
$*\mathcal{F}^w = -\left(\mathcal{F}^{32} d\xi^0 \wedge d\xi^1 + \mathcal{F}^{13} d\xi^0 \wedge d\xi^2 + \mathcal{F}^{21} d\xi^0 \wedge d\xi^3\right) \det P'$
$\mathcal{F}^{\alpha\beta} = \sum_{\lambda\mu=0}^3 g^{\alpha\lambda} g^{\beta\mu} \mathcal{F}_{\lambda\mu}$
3 form : $F = F_0 d\xi^1 \wedge d\xi^2 \wedge d\xi^3 + F_1 d\xi^0 \wedge d\xi^2 \wedge d\xi^3 + F_2 d\xi^0 \wedge d\xi^1 \wedge d\xi^3 + F_3 d\xi^0 \wedge d\xi^1 \wedge d\xi^2$
$*F = \sum_{\alpha=0}^3 (-1)^\alpha F^\alpha (\det P')\, d\xi^\alpha$
$F^0 = F^{123} = \sum_{\alpha\beta\gamma} g^{1\alpha} g^{2\beta} g^{3\gamma} F_{\alpha\beta\gamma} = \sum_{\alpha\beta\gamma} g^{1\alpha} g^{2\beta} g^{3\gamma} \epsilon(\alpha,\beta,\gamma) F_{123} = Min_{00} [g]^{-1} F_1$
$F^1 = F^{023} = \sum_{\alpha\beta\gamma} g^{0\alpha} g^{2\beta} g^{3\gamma} \epsilon(\alpha,\beta,\gamma) F_{123} = Min_{11} [g]^{-1} F_1$
$F^2 = F^{013} = \sum_{\alpha\beta\gamma} g^{0\alpha} g^{1\beta} g^{3\gamma} \epsilon(\alpha,\beta,\gamma) F_{123} = Min_{22} [g]^{-1} F_2$
$F^3 = F^{012} = \sum_{\alpha\beta\gamma} g^{0\alpha} g^{1\beta} g^{2\gamma} \epsilon(\alpha,\beta,\gamma) F_{123} = Min_{33} [g]^{-1} F_3$
$F^\alpha = Min_{\alpha\alpha} [g]^{-1} F_\alpha$ where $Min_{\alpha\alpha} [g]^{-1}$ is the determinant of the matrix $[g]^{-1}$ with the removal of the line and column α.
$g^{\alpha\gamma} = \frac{1}{\det[g]} (-1)^{\alpha+\gamma} (Min_{\alpha,\gamma} [g])$ (Maths.472)
$g_{\alpha\gamma} = \det[g] (-1)^{\alpha+\gamma} \left(Min_{\alpha,\gamma} [g]^{-1}\right)$
$g_{\alpha\alpha} = \det[g] (-1)^{\alpha+\alpha} \left(Min_{\alpha,\alpha} [g]^{-1}\right)$
$\left(Min_{p,p} [g]^{-1}\right) = \frac{1}{\det g} g_{\alpha\alpha} = (\det P)^2 g_{\alpha\alpha}$
$*F = \sum_{\alpha=0}^3 (-1)^\alpha g_{\alpha\alpha} F_\alpha (\det P)\, d\xi^\alpha$
4 form : $*\lambda_4 = *\left(\lambda d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3\right) = -(\det P)\lambda$

A.4 CODIFFERENTIAL

$\delta\lambda_r = (-1)^{r(4-r)-r} * d * \lambda_r = *d*\lambda_r$
For any form : $\delta^2\lambda = 0$
0 form : $\delta f = 0$
1 form : $\delta\left(\sum_{\alpha=0}^3 \lambda_\alpha d\xi^\alpha\right) = (\det P) \sum_{\alpha,\beta=0}^3 \partial_\alpha \left(g^{\alpha\beta}\lambda_\beta (\det P')\right)$
2 form : $\delta\mathcal{F} = *d*\mathcal{F}$
$d*\mathcal{F} = \left(\sum_{\alpha=0}^3 (-1)^\alpha \left(\sum_{\beta=0}^3 \partial_\beta \left(\mathcal{F}^{\alpha\beta} \det P'\right)\right) d\xi^0 \wedge ...\widehat{d\xi^\alpha}... \wedge d\xi^3\right)$
$\delta\mathcal{F} = \sum_{\alpha=0}^3 (-1)^\alpha g_{\alpha\alpha} \left(\sum_{\beta=0}^3 \partial_\beta \left(\mathcal{F}^{\alpha\beta} \det P'\right)\right) (\det P)\, d\xi^\alpha$
3 form : $\delta\left(\sum_{\alpha=0}^3 F_\alpha d\xi^0 \wedge ...\widehat{d\xi^\alpha}... \wedge d\xi^3\right) = *d\left(\sum_{\alpha=0}^3 (-1)^\alpha g_{\alpha\alpha} F_\alpha (\det P)\, d\xi^\alpha\right)$
$= *\left\{\sum_{\beta=0}^3 \partial_\alpha \left(\sum_{\beta=0}^3 (-1)^\beta g_{\beta\beta} F_\beta (\det P)\right) d\xi^\alpha \wedge d\xi^\beta\right\}$

$\delta F^r = -\left(F^{01} d\xi^3 \wedge d\xi^2 + F^{02} d\xi^1 \wedge d\xi^3 + F^{03} d\xi^2 \wedge d\xi^1\right) \det P'$
$\delta F^w = -\left(F^{32} d\xi^0 \wedge d\xi^1 + F^{13} d\xi^0 \wedge d\xi^2 + F^{21} d\xi^0 \wedge d\xi^3\right) \det P'$
$F^{\alpha\beta} = \sum_{\lambda\mu=0}^{3} g^{\alpha\lambda} g^{\beta\mu} \partial_\lambda \left(\sum_{\mu=0}^{3} (-1)^\mu g_{\mu\mu} F_\mu (\det P)\right)$
4 form : $\delta\left(\lambda d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3\right) = *d\left(-(\det P)\lambda\right) = -*\sum_{\beta=0}^{3} \partial_\beta (\lambda \det P)\, d\xi^\beta$
$= -\sum_{\alpha=0}^{3} (-1)^\alpha g^{\alpha\beta} \left(\sum_{\beta=0}^{3} \partial_\beta (\lambda \det P)\right) d\xi^0 \wedge ...\widehat{d\xi^\alpha} ... \wedge d\xi^3$

A.5 FORMULAS

A.5.1 ALGEBRA

Operator j

$r \in \mathbb{C}^3, w \in \mathbb{C}^3$:

$$[j(r)][w] = \begin{bmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} r_2 w_3 - r_3 w_2 \\ -r_1 w_3 + r_3 w_1 \\ r_1 w_2 - r_2 w_1 \end{bmatrix}$$

$$[w]^t [j(r)] = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} -r_2 w_3 + r_3 w_2 & r_1 w_3 - r_3 w_1 & -r_1 w_2 + r_2 w_1 \end{bmatrix}$$

$$[j(r)]^\alpha_\beta = -\epsilon(\alpha, \beta, \gamma) r_\gamma$$

$$[j(r) w]^a = \sum_{b,c=1}^3 \epsilon(a,b,c) r_b w_c$$

$$[wj(r)]_a = -\sum_{b,c=1}^3 \epsilon(a,b,c) r_b w_c$$

$$[j(r)]^t = -[j(r)] = [j(-r)]$$

$$[j(x)][y] = -[j(y)][x]$$

$$[y]^t (j(x)) = -[x]^t [j(y)]$$

$$[j(x)][y] = 0 \Leftrightarrow \exists k \in \mathbb{R} : y = kx$$

$$[x]^t [j(y)][z] = -\det \begin{bmatrix} x_1 & y_2 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix}$$

$$\begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} - \begin{bmatrix} M_{11} & M_{21} & M_{31} \\ M_{12} & M_{22} & M_{32} \\ M_{13} & M_{23} & M_{33} \end{bmatrix} = j \begin{pmatrix} -M_{23} + M_{32} \\ M_{13} - M_{31} \\ -M_{12} + M_{21} \end{pmatrix}$$

Eigenvectors of j(r)

$$0 : \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}$$

$$ir : \begin{bmatrix} -(-r_1 r_2 + i r_3 r) \\ -\left(r_1^2 + r_3^2\right) \\ r_2 r_3 + i r_1 r \end{bmatrix}$$

$$-ir : \begin{bmatrix} -(r_1 r_2 + i r_3 r) \\ \left(r_1^2 + r_3^2\right) \\ -r_2 r_3 + r_1 i r \end{bmatrix}$$

Eigen vectors of j(r)j(r)

$0 : [r]$

$$-\left(r_1^2 + r_2^2 + r_3^2\right) : \begin{bmatrix} -r_2 \\ r_1 \\ 0 \end{bmatrix}, \begin{bmatrix} -r_3 \\ 0 \\ r_1 \end{bmatrix}$$

Identities

With 2 operators :

$$[j(x)][j(y)] = [y][x]^t - \left([y]^t [x]\right) I$$

$$Tr([j(x)][j(y)]) = -2 [x]^t [y]$$

$[x]^t [j(r)][j(s)][y] = \left([x]^t [s]\right)\left([r]^t [y]\right) - \left([x]^t [y]\right)\left([r]^t [s]\right)$

With 3 operators :

$[j(x)][j(y)][j(x)] = -\left([y]^t [x]\right)[j(x)]$

$[j(y)][j(x)][j(x)] + [j(x)][j(x)][j(y)] = -\left([y]^t [x]\right)[j(x)] - \left([x]^t [x]\right)[j(y)]$

Powers :

$k > 0 : [j(r)]^{2k} = \left(-[r]^t [r]\right)^{k-1} [j(r)][j(r)]$

$k \geq 0 : [j(r)]^{2k+1} = \left(-[r]^t [r]\right)^{k} [j(r)]$

$\exp[j(r)] = I_3 + \frac{\sin\sqrt{r^t r}}{\sqrt{r^t r}}[j(r)] + \frac{1-\cos\sqrt{r^t r}}{r^t r}[j(r)][j(r)]$

Iteration :

$[j(j(x)y)] = [y][x]^t - [x][y]^t = [j(x)][j(y)] - [j(y)][j(x)]$

$\left[j([w]^t j(r))\right] = -\left[j([j(r)w]^t)\right] = -[w][r]^t + [r][w]^t = -[j(r)][j(w)] + [j(w)][j(r)]$

$[j(j(x)j(x)y)] = \left([y]^t [x]\right)[j(x)] - \left([x]^t [x]\right)[j(y)]$

With Matrices :

$[M],[X] \in L(3)$:

$[M]^t [j(r)][M] = j\left([M]^{-1}[r]\right)\det M = j\left([r]^t \left([M]^t\right)^{-1}\right)\det M$

$[j([M][r])] = \left([M]^{-1}\right)^t [j(r)][M]^{-1}\det M$

$j\left([r]^t [M]\right) = \left([M]^{-1}\right)[j(r)]\left([M]^{-1}\right)^t \det M$

$([M]_1)^t [j([M]_2)][M]_3 = \det M$

$\left[\ j([M]_2)[M]_3 \quad j([M]_3)[M]_1 \quad j([M]_1)[M]_2\ \right] = (\det M)\left[M^{-1}\right]^t$

$([X]_1)^t [M]^t j([M][X]_2)[M][X]_3 = (\det[M])(\det[X])$

$Tr\left([M]^t [j(x)][M]\right) = 0$

$M \in O(3) : j([M][x])[M][y] = [M][j(x)][y] \Leftrightarrow Mx \times My = M(x \times y)$

Miscellaneous :

$$\begin{bmatrix} 0 & j(z) & -j(y) \\ -j(z) & 0 & j(x) \\ j(y) & -j(x) & 0 \end{bmatrix}^{-1} = \frac{1}{2x^t j(y) z}\begin{bmatrix} xx^t & 2yx^t - xy^t & 2zx^t - xz^t \\ 2xy^t - yx^t & yy^t & 2zy^t - yz^t \\ 2xz^t - zx^t & 2yz^t - zy^t & zz^t \end{bmatrix}$$

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.

$(a + bj(z) + cj(z)j(z))(a' + b'j(z) + c'j(z)j(z))$

$= aa' + (ab' + a'b - (z^t z)(c'b + b'c))j(z) + (ac' + a'c + b'b - (z^t z)c'c)j(z)j(z)$

$\det(a + bj(z) + cj(z)j(z)) = a\left(a^2 + \left(b^2 + c^2(z^t z) - 2ac\right)(z^t z)\right)$

$[a + bj(z) + cj(z)j(z)]^{-1} = \left[\frac{1}{a}I - \frac{ab}{\det P}j(z) - \frac{\left(ac - b^2 - c^2\left(z^t z\right)\right)}{\det P}j(z)j(z)\right]$

eigenvectors of $P(z)$: the only real eigen value is a with eigen vector z

Matrices on SO(3,1)

signature $(3,1) : \langle u, v\rangle = u^1v^1 + u^2v^2 + u^3v^3 - u^0v^0$

signature $(1,3) : \langle u, v\rangle = -u^1v^1 - u^2v^2 - u^3v^3 + u^0v^0$

$[\kappa]^t [\eta][\kappa] = [\eta] \Leftrightarrow [\kappa] \in SO(3,1) \equiv SO(1,3)$

$[\chi] = \exp[K(w)]\exp[J(r)]$

$\exp\left[K\left(w\right)\right]=I_4+\frac{\sinh\sqrt{w^tw}}{\sqrt{w^tw}}K(w)+\frac{\cosh\sqrt{w^tw}-1}{w^tw}K(w)K(w)$

$[\eta]=\begin{bmatrix}-1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$

Basis of $so\left(3,1\right)\equiv so\left(1,3\right)$

$[\kappa_1]=\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&-1\\0&0&1&0\end{bmatrix};[\kappa_2]=\begin{bmatrix}0&0&0&0\\0&0&0&1\\0&0&0&0\\0&-1&0&0\end{bmatrix};[\kappa_3]=\begin{bmatrix}0&0&0&0\\0&0&-1&0\\0&1&0&0\\0&0&0&0\end{bmatrix}$

$[\kappa_4]=\begin{bmatrix}0&1&0&0\\1&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix};[\kappa_5]=\begin{bmatrix}0&0&1&0\\0&0&0&0\\1&0&0&0\\0&0&0&0\end{bmatrix};[\kappa_6]=\begin{bmatrix}0&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&0\end{bmatrix}$

$[\kappa]=\left[J\left(r\right)\right]+\left[K\left(w\right)\right]\in so\left(3,1\right)$ with

$\left[J\left(r\right)\right]=\begin{bmatrix}0&0&0&0\\0&0&-r_3&r_2\\0&r_3&0&-r_1\\0&-r_2&r_1&0\end{bmatrix};\left[K\left(w\right)\right]=\begin{bmatrix}0&w_1&w_2&w_3\\w_1&0&0&0\\w_2&0&0&0\\w_3&0&0&0\end{bmatrix}$

Dirac's matrices

$\sigma_1=\begin{bmatrix}0&1\\1&0\end{bmatrix};\sigma_2=\begin{bmatrix}0&-i\\i&0\end{bmatrix};\sigma_3=\begin{bmatrix}1&0\\0&-1\end{bmatrix};\sigma_0=\begin{bmatrix}1&0\\0&1\end{bmatrix}$

$\begin{bmatrix}1\backslash 2&\sigma_1&\sigma_2&\sigma_3\\\sigma_1&\sigma_0&i\sigma_3&-i\sigma_2\\\sigma_2&-i\sigma_3&\sigma_0&i\sigma_1\\\sigma_3&i\sigma_2&-i\sigma_1&\sigma_0\end{bmatrix}$

$\sigma_j^*=\sigma=\sigma_j^{-1}$

$\sigma_j\sigma_k+\sigma_k\sigma_j=2\delta_{jk}\sigma_0$

$j\neq k,l=1,2,3:\sigma_j\sigma_k=\epsilon\left(j,k,l\right)i\sigma_l$

$\sigma_1\sigma_2\sigma_3=i\sigma_0$

Matrices $\sigma\left(z\right)$

$\sum_{a=1}^3 z_a\sigma_a=\sigma\left(z\right)$ with $z\in\mathbb{C}^3$

$\sigma\left(z\right)=\begin{bmatrix}z_3&z_1-iz_2\\z_1+iz_2&-z_3\end{bmatrix}\in sl\left(\mathbb{C},2\right)$

$\left(\sigma\left(z\right)\right)^*=\sigma\left(\overline{z}\right)$

$\sigma\left(z\right)\sigma\left(z'\right)=\sigma\left(j\left(z\right)z'\right)+z^tz'\sigma_0$

$\sigma\left(z\right)^{-1}=\frac{1}{z^tz}\sigma\left(z\right)$

$\sigma\left(z\right)\sigma\left(z'\right)-\sigma\left(z'\right)\sigma\left(z\right)=2\sigma\left(j\left(z\right)z'\right)$

$\sigma\left(z'\right)\sigma\left(z\right)\sigma\left(z'\right)=\left(\left(z'\right)^tz'\right)\sigma\left(z\right)$

$\sigma\left(z\right)=k\sigma_0,k\in\mathbb{C}\Rightarrow z=0$

eigenvectors of $\sigma\left(z\right):\epsilon=\pm1:\epsilon\sqrt{z^tz}:\begin{bmatrix}z_1-iz_2\\\epsilon\sqrt{z^tz}-z_3\end{bmatrix}$

γ matrices

$\gamma_0=\begin{bmatrix}0&-i\sigma_0\\i\sigma_0&0\end{bmatrix};\gamma_1=\begin{bmatrix}0&\sigma_1\\\sigma_1&0\end{bmatrix};\gamma_2=\begin{bmatrix}0&\sigma_2\\\sigma_2&0\end{bmatrix};\gamma_3=\begin{bmatrix}0&\sigma_3\\\sigma_3&0\end{bmatrix};$

$$\gamma_i\gamma_j + \gamma_j\gamma_i = 2\delta_{ij}I_4$$
$$\gamma_i = \gamma_i^* = \gamma_i^{-1}$$
$$j = 1,2,3 : \widetilde{\gamma}_j = \begin{bmatrix} \sigma_j & 0 \\ 0 & \sigma_j \end{bmatrix}$$
$$\begin{bmatrix} 1\backslash 2 & \gamma_0 & \gamma_1 & \gamma_2 & \gamma_3 \\ \gamma_0 & \gamma_0 & -i\gamma_5\widetilde{\gamma}_1 & -i\gamma_5\widetilde{\gamma}_2 & -i\gamma_5\widetilde{\gamma}_3 \\ \gamma_1 & i\gamma_5\widetilde{\gamma}_1 & \gamma_0 & i\widetilde{\gamma}_3 & -i\widetilde{\gamma}_2 \\ \gamma_2 & i\gamma_5\widetilde{\gamma}_2 & -i\widetilde{\gamma}_3 & \gamma_0 & i\widetilde{\gamma}_1 \\ \gamma_3 & i\gamma_5\widetilde{\gamma}_3 & i\widetilde{\gamma}_2 & -i\widetilde{\gamma}_1 & \gamma_0 \end{bmatrix}$$
$$\gamma_1\gamma_2\gamma_3 = i\begin{bmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{bmatrix}$$
$$\gamma_5 = \gamma_0\gamma_1\gamma_2\gamma_3 = \begin{bmatrix} \sigma_0 & 0 \\ 0 & -\sigma_0 \end{bmatrix}$$
$$\gamma_5\gamma_5 = I$$
$$\gamma_5\gamma_a = -\gamma_a\gamma_5$$

γC matrices

$$Cl(3,1) : \gamma C(\varepsilon_j) = \gamma_j, j = 1,2,3; \gamma C(\varepsilon_0) = i\gamma_0; \gamma C(\varepsilon_5) = i\gamma_5$$
$$Cl(1,3) : \gamma C'(\varepsilon_j) = i\gamma_j, j = 1,2,3; \gamma C'(\varepsilon_0) = \gamma_0; \gamma C'(\varepsilon_5) = \gamma_5$$
$$a = 1,2,3 : \gamma C(\overrightarrow{\kappa}_a) = -\frac{1}{2}i\begin{bmatrix} \sigma_a & 0 \\ 0 & \sigma_a \end{bmatrix} = -\frac{1}{2}i\widetilde{\gamma}_a$$
$$a = 4,5,6 : \gamma C(\overrightarrow{\kappa}_a) = \frac{1}{2}\begin{bmatrix} \sigma_a & 0 \\ 0 & -\sigma_a \end{bmatrix} = -\frac{1}{2i}\gamma_0\gamma_j$$
$$\gamma C(\upsilon(r,w)) = -\frac{1}{2}i\begin{bmatrix} \sigma(r+iw) & 0 \\ 0 & \sigma(r-iw) \end{bmatrix} = -\frac{1}{2}i\begin{bmatrix} \sigma(Z) & 0 \\ 0 & \sigma(\overline{Z}) \end{bmatrix}$$
$$\gamma C(a + \upsilon(r,w) + b\varepsilon_5)$$
$$= \begin{bmatrix} a + ib - \frac{1}{2}i\sigma(r+iw) & 0 \\ 0 & a - ib - \frac{1}{2}i\sigma(r-iw) \end{bmatrix}$$
$$= \begin{bmatrix} A - \frac{1}{2}i\sigma(Z) & 0 \\ 0 & \overline{A} - \frac{1}{2}i\sigma(\overline{Z}) \end{bmatrix}$$

A.5.2 CLIFFORD ALGEBRA

$$\varepsilon_i \cdot \varepsilon_j + \varepsilon_j \cdot \varepsilon_i = 2\eta_{ij}$$
$$\varepsilon_5 \cdot \varepsilon_5 = -1$$
$$X \cdot \varepsilon_5 + \varepsilon_5 \cdot X = 0$$

Adjoint map

$$\forall X \in Cl(3,1), s \in Spin(3,1) : \mathbf{Ad}_s X = s \cdot X \cdot s^{-1}$$
$$\langle \mathbf{Ad}_s X, \mathbf{Ad}_s Y \rangle = \langle X, Y \rangle$$
$$\mathbf{Ad}_s \circ \mathbf{Ad}_{s'} = \mathbf{Ad}_{s \cdot s'}$$
$$\forall s \in Spin(3,1) : \mathbf{Ad}_s \varepsilon_5 = \varepsilon_5$$

Action of the Adjoint map on vectors :

$$V = \sum_{i=0}^{3} V^i \varepsilon_i \to \widetilde{V} = \mathbf{Ad}_s V = \sum_{i=0}^{3} \widetilde{V}^i \varepsilon_i$$
$$\widetilde{V}^i = \sum_{j=0}^{3} [h(s)]^i_j v^j$$
$$[h(s)] =$$
$$\begin{bmatrix} a^2 + b^2 + \frac{1}{4}(r^t r + w^t w) & a w^t - b r^t + \frac{1}{2} w^t j(r) \\ a w - b r + \frac{1}{2} j(r) w & a^2 + b^2 + \frac{1}{4}(r^t r + w^t w) + a j(r) + b j(w) + \frac{1}{2}(j(r) j(r) + j(w) j(w)) \end{bmatrix}$$

$[h(s)]^t [\eta] [h(s)] = [\eta]$

If $s = a_w + \upsilon(0,w)$

$[h(s)] = \begin{bmatrix} 2a_w^2 - 1 & a_w w^t \\ a_w w & 2a_w^2 - 1 + \frac{1}{2} j(w) j(w) \end{bmatrix}$

If $s = a_r + \upsilon(r,0)$

$[h(s)] = \begin{bmatrix} 1 & 0 \\ 0 & 1 + a_r j(r) + \frac{1}{2} j(r) j(r) \end{bmatrix}$

$[C(r)] = 1 + a_r j(r) + \frac{1}{2} j(r) j(r) \in SO(3)$

Action of the adjoint map on the Lie algebra:

$Z = \sum_{a=1}^{6} Z_a \overrightarrow{\kappa}_a \to \widetilde{Z} = \sum_{a=1}^{6} \widetilde{Z}_a \overrightarrow{\kappa}_a$

With :

$Z = \upsilon(X,Y) \to \widetilde{Z} = \upsilon\left(\widetilde{X}, \widetilde{Y}\right)$

$\begin{bmatrix} \widetilde{X} \\ \widetilde{Y} \end{bmatrix} = [\mathbf{Ad}_s] \begin{bmatrix} X \\ Y \end{bmatrix}$

$[\mathbf{Ad}_s] =$

$$\begin{bmatrix} 1 + aj(r) - bj(w) + \frac{1}{2}(j(r)j(r) - j(w)j(w)) & -\left(aj(w) + bj(r) + \frac{1}{2}(j(r)j(w) + j(w)j(r))\right) \\ aj(w) + bj(r) + \frac{1}{2}(j(r)j(w) + j(w)j(r)) & 1 + aj(r) - bj(w) + \frac{1}{2}(j(r)j(r) - j(w)j(w)) \end{bmatrix}$$

With $s_w = a_w + \upsilon(0,w)$

$[\mathbf{Ad}_s] = \begin{bmatrix} \left[1 - \frac{1}{2} j(w) j(w)\right] & -[a_w j(w)] \\ [a_w j(w)] & \left[1 - \frac{1}{2} j(w) j(w)\right] \end{bmatrix} = \begin{bmatrix} A & -B \\ B & A \end{bmatrix}$

$A = A^t, B^t = -B, AB = BA$

$A^2 + B^2 = I$

$[\mathbf{Ad}_{s_w}]^{-1} = \left[\mathbf{Ad}_{s_w^{-1}}\right] = \begin{bmatrix} A & B \\ -B & A \end{bmatrix}$

With $s_r = a_r + \upsilon(r,0)$

$[\mathbf{Ad}_s] = \begin{bmatrix} \left[1 + a_r j(r) + \frac{1}{2} j(r) j(r)\right] & 0 \\ 0 & \left[1 + a_r j(r) + \frac{1}{2} j(r) j(r)\right] \end{bmatrix} = \begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix}$

$CC^t = C^t C = I_3$

$[\mathbf{Ad}_{s_r}]^{-1} = \left[\mathbf{Ad}_{s_r^{-1}}\right] = \begin{bmatrix} C^t & 0 \\ 0 & C^t \end{bmatrix}$

Lie Algebra

$\upsilon(r,w) = \frac{1}{2}\left(w^1 \varepsilon_0 \cdot \varepsilon_1 + w^2 \varepsilon_0 \cdot \varepsilon_2 + w^3 \varepsilon_0 \cdot \varepsilon_3 + r^3 \varepsilon_2 \cdot \varepsilon_1 + r^2 \varepsilon_1 \cdot \varepsilon_3 + r^1 \varepsilon_3 \cdot \varepsilon_2\right)$

$\overrightarrow{\kappa}_1 = \upsilon((1,0,0),(0,0,0)) = \frac{1}{2}\varepsilon_3 \cdot \varepsilon_2,$

$\overrightarrow{\kappa}_2 = \upsilon((0,1,0),(0,0,0)) = \frac{1}{2}\varepsilon_1 \cdot \varepsilon_3,$

$\overrightarrow{\kappa}_3 = \upsilon((0,0,1),(0,0,0)) = \frac{1}{2}\varepsilon_2 \cdot \varepsilon_1,$

$\overrightarrow{\kappa}_4 = \upsilon((0,0,0),(1,0,0)) = \frac{1}{2}\varepsilon_0 \cdot \varepsilon_1,$

$\overrightarrow{\kappa}_5 = \upsilon((0,0,0),(0,1,0)) = \frac{1}{2}\varepsilon_0 \cdot \varepsilon_2,$

$\overrightarrow{\kappa}_6 = \upsilon((0,0,0),(0,0,1)) = \frac{1}{2}\varepsilon_0 \cdot \varepsilon_3$

Multiplication table :

$$\begin{bmatrix} & \kappa_1 & \kappa_2 & \kappa_3 \\ \kappa_1 & -\frac{1}{4} & \frac{1}{2}\kappa_3 & -\frac{1}{2}\kappa_2 \\ \kappa_2 & -\frac{1}{2}\kappa_3 & -\frac{1}{4} & \frac{1}{2}\kappa_1 \\ \kappa_3 & \frac{1}{2}\kappa_2 & -\frac{1}{2}\kappa_1 & -\frac{1}{4} \end{bmatrix}$$

$\upsilon(r,w) \cdot \varepsilon_5 = \varepsilon_5 \cdot \upsilon(r,w) = \upsilon(-w,r)$

$V = V^0 \varepsilon_0 + v :$

$[\upsilon(r,w), V] = \upsilon(r,w) \cdot V - V \cdot \upsilon(r,w) = \frac{1}{2}\left(V^0 r + (r^t v)\varepsilon_0 - j(w) v\right)$

In $Cl(3,1)$:

$\upsilon(r',w')\cdot\upsilon(r,w)$
$=\frac{1}{4}(w^tw'-r^tr')+\frac{1}{2}\upsilon(-j(r)r'+j(w)w',-j(w)r'-j(r)w')-\frac{1}{4}(w^tr'+r^tw')\varepsilon_5$
$[\upsilon(r,w),\upsilon(r',w')]=\upsilon(j(r)r'-j(w)w',j(w)r'+j(r)w')$
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 $Cl(1,3)$:

$r"=\frac{1}{2}(j(r)r'-j(w)w')+a'r+ar'+b'w+bw'$
$w"=-\frac{1}{2}(j(w)r'+j(r)w')+a'w+aw'+b'r+br'$

Complex structure

$Cl(3,1)$:

$$\begin{bmatrix} \text{real} & \text{imaginary} \\ E_j & E_j \cdot \varepsilon_5 = iE_j \\ 1 & \varepsilon_5 \\ \varepsilon_1 & \varepsilon_0 \cdot \varepsilon_3 \cdot \varepsilon_2 \\ \varepsilon_2 & \varepsilon_0 \cdot \varepsilon_1 \cdot \varepsilon_3 \\ \varepsilon_3 & \varepsilon_0 \cdot \varepsilon_2 \cdot \varepsilon_1 \\ \varepsilon_3 \cdot \varepsilon_2 & \varepsilon_0 \cdot \varepsilon_1 \\ \varepsilon_1 \cdot \varepsilon_3 & \varepsilon_0 \cdot \varepsilon_2 \\ \varepsilon_2 \cdot \varepsilon_1 & \varepsilon_0 \cdot \varepsilon_3 \\ \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3 & \varepsilon_0 \end{bmatrix}$$

Lie Algebra

$\upsilon(r,w) = \sum_{a=1}^{3} (r_a + iw_a) \overrightarrow{\kappa}_a = \sum_{a=1}^{3} Z^a \overrightarrow{\kappa}_a = (r + iw)$

$Z' \cdot Z = -\frac{1}{4} Z^t Z' + \frac{1}{2} j(Z') Z$

$[\upsilon(r,w), \upsilon(r',w')] = j(Z) Z'$

Spin group

$g = a + \upsilon(r,w) + b\varepsilon_5 = A + Z$

$A^2 = 1 - \frac{1}{4} Z^t Z$

Derivative :

$\frac{\partial \sigma}{\partial x} \cdot \sigma^{-1} = D(Z) \frac{\partial Z}{\partial x}$

$\sigma^{-1} \cdot \frac{\partial \sigma}{\partial x} = D(-Z) \frac{\partial Z}{\partial x}$

$D(Z) = \frac{1}{A} + \frac{1}{2} j(Z) + \frac{1}{4A} j(Z) j(Z)$

$[D(Z)]^{-1} = A - \frac{1}{2} j(Z)$

Adjoint map :

$\mathbf{Ad}_s iX = i\mathbf{Ad}_s X$

$\mathbf{Ad}_s \upsilon(r,w) = Ad(Z)[X] = \left(1 + Aj(Z) + \frac{1}{2} j(Z) j(Z)\right)[X]$

$[Ad(Z)][D(Z)] = D(-Z)$

A.5.3 GEOMETRY

Pull back, push forward

$f^* = (f_*)^{-1}$

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 : $f_* V = Tf(V)$

$f_* : V \in TM \to f_* V \in TN :: f_* V(f(m)) = f'(m) V(m)$

pull back of a 1 form : $f^* \lambda = \lambda(Tf)$

$f^* :: \lambda \in TN^* \to f^* \lambda \in TM^* :: f^* \lambda(m)(u_m) = \lambda(f(m)) f'(m) u_m$

If f is a diffeomorphism :

pull back of a vector field : $f^* V = (Tf)^{-1}(V)$

$f^* :: V \in TN \to f^* V \in TM :: f^* V(m) = (f')^{-1}(n) V(n)$

push forward of a 1 form : $f_* \lambda = \lambda\left((Tf)^{-1}\right)$

$f_* : \lambda \in TM^* \to f_* \lambda \in TN^* :: f_* \lambda(f(m))\left(u_{f(m)}\right) = \lambda\left(f^{-1}(n)\right)(f'(n))^{-1}(u_n)$

Flow of a vector field

$\Phi_V(\tau, a) : \mathbb{R} \times M \to M$

$\frac{\partial}{\partial \tau} \Phi_V(\tau, a)|_{\tau=\theta} = V(\Phi_V(\theta, a))$

$\Phi_V(0, a) = a$

$\Phi_V(\tau+\tau',a)=\Phi_V(\tau,\Phi_V(\tau',a))$
$\frac{\partial}{\partial m}\Phi_V(\tau,m):T_mM\to T_{\Phi_V(\tau,m)}M$
$\Phi'_{Vm}(\tau,m)(\partial\xi_\alpha(q(0)))=\sum_{\lambda=0}^3[J(\tau)]_\alpha^\lambda\partial\xi_\lambda(q(\tau))$
$\Phi'_{Vm}(\tau,m)(d\xi^\alpha(q(0)))=\sum_{\lambda=0}^3[K(\tau)]_\alpha^\lambda d\xi^\alpha(q(\tau))$ with $[K(\tau)]=[J(\tau)]^{-1}$
$\pounds_VS=0\Leftrightarrow S(\Phi_V(\tau,m))=\Phi'_{Vm}(\tau,m)_*S(m)\Leftrightarrow S(\Phi_V(0,m))=\Phi'_{Vm}(\tau,m)^*S(m)$
Integral curve : $q(\tau)=\Phi_V(\tau,a)$

Divergence of a vector field V

$\pounds_V\varpi_4=div(V)\varpi_4$
$divV=\frac{1}{\sqrt{-\det g}}\sum_{\alpha=0}^3\partial_\alpha\left(V^\alpha\sqrt{-\det g}\right)=\sum_{\alpha=0}^3\frac{\partial V^\alpha}{\partial\xi^\alpha}+\frac{1}{2}V^\alpha\sum_{\beta\gamma=0}^3g^{\beta\gamma}\partial_\alpha g_{\beta\gamma}$

Hodge dual of r-forms

Metric :
$G_r(\lambda,\mu)=\sum_{\{\alpha_1...\alpha_r\}\{\beta_1...\beta_r\}}\lambda_{\alpha_1...\alpha_r}\mu_{\beta_1...\beta_r}\det\left[g^{-1}\right]_{\{\beta_1...\beta_r\}}^{\{\alpha_1...\alpha_r\}}$
Hodge dual :
$\forall\lambda,\mu\in\Lambda_r(M;\mathbb{R}):*\lambda\wedge\mu=G_r(\lambda,\mu)\varpi_4=G_r(\mu,\lambda)\varpi_4$
0 form (a function) : $*f=f\varpi_4$
1 form : $*\lambda_1=*\left(\sum_{\alpha=0}^3\lambda_\alpha d\xi^\alpha\right)=\sum_{\alpha=0}^3(-1)^\alpha g^{\alpha\beta}\lambda_\beta d\xi^0\wedge...\widehat{d\xi^\alpha}...\wedge d\xi^3$
2 form :
$*\mathcal{F}^r=-\left(\mathcal{F}^{01}d\xi^3\wedge d\xi^2+\mathcal{F}^{02}d\xi^1\wedge d\xi^3+\mathcal{F}^{03}d\xi^2\wedge d\xi^1\right)\det P'$
$*\mathcal{F}^w=-\left(\mathcal{F}^{32}d\xi^0\wedge d\xi^1+\mathcal{F}^{13}d\xi^0\wedge d\xi^2+\mathcal{F}^{21}d\xi^0\wedge d\xi^3\right)\det P'$
$\mathcal{F}^{\alpha\beta}=\sum_{\lambda\mu=0}^3g^{\alpha\lambda}g^{\beta\mu}\mathcal{F}_{\lambda\mu}$
3 form : $*F=\sum_{\alpha=0}^3(-1)^\alpha g_{\alpha\alpha}F_\alpha(\det P)d\xi^\alpha$
4 form : $*\lambda_4=*\left(\lambda d\xi^0\wedge d\xi^1\wedge d\xi^2\wedge d\xi^3\right)=-(\det P)\lambda$

Codifferential

$\delta\lambda_r=(-1)^{r(4-r)-r}*d*\lambda_r=*d*\lambda_r$
For any form : $\delta^2\lambda=0$
0 form : $\delta f=0$
1 form : $\delta\left(\sum_{\alpha=0}^3\lambda_\alpha d\xi^\alpha\right)=(\det P)\sum_{\alpha,\beta=0}^3\partial_\alpha\left(g^{\alpha\beta}\lambda_\beta(\det P')\right)$
2 form :
$\delta\mathcal{F}=\sum_{\alpha=0}^3(-1)^\alpha g_{\alpha\alpha}\left(\sum_{\beta=0}^3\partial_\beta\left(\mathcal{F}^{\alpha\beta}\det P'\right)\right)(\det P)d\xi^\alpha$

Lie derivative

Vector field :
$\pounds_VU=[V,U]=0=\sum_{\alpha,\beta=0}^3\left(V^\beta\partial_\beta U^\alpha-U^\beta\partial_\beta V^\alpha\right)\partial\xi_\alpha=-\pounds_UV$
One form :
$\pounds_V\lambda(m)=\sum_{a=1}^m\sum_{\alpha,\gamma=0}^3\left(V^\gamma\partial_\gamma\lambda_\alpha^a+\lambda_\gamma^a\partial_\alpha V^\gamma\right)d\xi^\alpha\otimes\overrightarrow{\theta}_a$
2 form :
$[(\pounds_V\mathcal{F})^r]=\sum_{\gamma=0}^3V^\gamma[\partial_\gamma\mathcal{F}^r]+[\mathcal{F}^{aw}]j\left(\partial V^0\right)+[\mathcal{F}^{ar}]\left(-[\partial v]^t+(div(v))I_3\right)$
$[(\pounds_V\mathcal{F})^w]=\sum_{\gamma=0}^3V^\gamma[\partial_\gamma\mathcal{F}^w]+[\mathcal{F}^{aw}]\left(\partial_0V^0+[\partial v]\right)-[\mathcal{F}^{ar}]j(\partial_0V)$
Metric :
$\pounds_V\left(\sum_{\alpha\beta}g_{\alpha\beta}(m)d\xi^\alpha\otimes d\xi^\beta\right)=\sum_{\alpha\beta}\widetilde{g}_{\alpha\beta}(m)d\xi^\alpha\otimes d\xi^\beta$
$\widetilde{g}_{\alpha\beta}(m)=\sum_{\gamma=0}^3V^\gamma\partial_\gamma g_{\alpha\beta}+g_{\gamma\beta}\partial_\alpha V^\gamma+g_{\alpha\gamma}\partial_\beta V^\gamma$

Algebra of two forms

Matricial representations :

$\mathcal{F}=\frac{1}{2}\sum_{\alpha,\beta=0}^{3}\mathcal{F}_{\alpha\beta}d\xi^{\alpha}\wedge d\xi^{\beta}=\sum_{\{\alpha,\beta\}=0}^{3}\mathcal{F}_{\alpha\beta}d\xi^{\alpha}\wedge d\xi^{\beta}$

$\mathcal{F}=\mathcal{F}^{r}+\mathcal{F}^{w}$

$\mathcal{F}^{r}=\mathcal{F}_{32}d\xi^{3}\wedge d\xi^{2}+\mathcal{F}_{13}d\xi^{1}\wedge d\xi^{3}+\mathcal{F}_{21}d\xi^{2}\wedge d\xi^{1}$

$\mathcal{F}^{w}=\mathcal{F}_{01}d\xi^{0}\wedge d\xi^{1}+\mathcal{F}_{02}d\xi^{0}\wedge d\xi^{2}+\mathcal{F}_{03}d\xi^{0}\wedge d\xi^{3}$

$[\mathcal{F}^{r}]=\left[\begin{array}{ccc}\mathcal{F}_{32} & \mathcal{F}_{13} & \mathcal{F}_{21}\end{array}\right];[\mathcal{F}^{w}]=\left[\begin{array}{ccc}\mathcal{F}_{01} & \mathcal{F}_{02} & \mathcal{F}_{03}\end{array}\right]$

$$[\mathcal{F}_{\alpha\beta}]_{\beta=0\ldots3}^{\alpha=0\ldots3}=\begin{bmatrix}0 & \mathcal{F}_{01} & \mathcal{F}_{02} & \mathcal{F}_{03}\\ \mathcal{F}_{10} & 0 & \mathcal{F}_{12} & \mathcal{F}_{13}\\ \mathcal{F}_{20} & \mathcal{F}_{21} & 0 & \mathcal{F}_{23}\\ \mathcal{F}_{30} & \mathcal{F}_{31} & \mathcal{F}_{32} & 0\end{bmatrix}=\left[\begin{array}{cc}0 & [\mathcal{F}^{w}]_{1\times3}\\ -\left([\mathcal{F}^{w}]\right)_{3\times1}^{t} & j\left([\mathcal{F}^{r}]\right)_{3\times3}\end{array}\right]$$

Change of chart :

$\left[\widetilde{\mathcal{F}}\right]=[K]^{t}[\mathcal{F}][K]$

$\left[\begin{array}{cc}\left[\widetilde{\mathcal{F}}^{r}\right] & \left[\widetilde{\mathcal{F}}^{w}\right]\end{array}\right]=\left[\begin{array}{cc}[\mathcal{F}^{r}] & [\mathcal{F}^{w}]\end{array}\right][L_{K}]$

$$[L_{K}]=\left[\begin{array}{cc}\left([k]^{-1}\right)^{t}\det k & -j\left([K_{0}]\right)[k]\\ [k]\,j\left(\left[K^{0}\right]\right) & K_{0}^{0}[k]-[K_{0}]\left[K^{0}\right]\end{array}\right]$$

$[L_{K_{1}K_{2}}]=[L_{K_{1}}][L_{K_{2}}]$

$[\mathcal{F}^{*}]=\left[\mathcal{F}^{\alpha\beta}\right]_{\beta=0..3}^{\alpha=0..3}=[g]^{-1}[\mathcal{F}][g]^{-1}$

$\left[\begin{array}{cc}\mathcal{F}^{*r} & \mathcal{F}^{*w}\end{array}\right]=\left[\begin{array}{cc}\mathcal{F}^{r} & \mathcal{F}^{w}\end{array}\right]\left[L_{g^{-1}}\right]$

Scalar product :

$G_{2}(\mathcal{F},K)=-\frac{1}{2}Tr\left([\mathcal{F}][g]^{-1}[K][g]^{-1}\right)=-\frac{1}{2}Tr\left([\mathcal{F}][K^{*}]\right)$

$=-\frac{1}{\det P'}\left([*\mathcal{F}^{w}][K^{r}]^{t}+[*\mathcal{F}^{r}][K^{w}]^{t}\right)=\sum_{\{\alpha\beta\}}\mathcal{F}^{\alpha\beta}K_{\alpha\beta}=\frac{1}{2}\sum_{\alpha\beta=0}^{3}\mathcal{F}^{\alpha\beta}K_{\alpha\beta}$

Standard chart :

$G_{2}(\mathcal{F},K)=[\mathcal{F}^{w}][g_{3}]^{-1}[K^{w}]^{t}+[\mathcal{F}^{r}][g_{3}][K^{r}]\det[g_{3}]^{-1}$

Hodge duality :

$*\mathcal{F}^{r}=-\left(\mathcal{F}^{01}d\xi^{3}\wedge d\xi^{2}+\mathcal{F}^{02}d\xi^{1}\wedge d\xi^{3}+\mathcal{F}^{03}d\xi^{2}\wedge d\xi^{1}\right)\det P'$

$*\mathcal{F}^{w}=-\left(\mathcal{F}^{32}d\xi^{0}\wedge d\xi^{1}+\mathcal{F}^{13}d\xi^{0}\wedge d\xi^{2}+\mathcal{F}^{21}d\xi^{0}\wedge d\xi^{3}\right)\det P'$

$*\mathcal{F}^{\alpha\beta}=-\mathcal{F}^{\alpha\beta}\det P'=-\left(\sum_{\lambda\mu=0}^{3}g^{\alpha\lambda}g^{\beta\mu}\mathcal{F}_{\lambda\mu}\right)\det P'$

$*\mathcal{F}=*\mathcal{F}^{r}+*\mathcal{F}^{w}$

$[*\mathcal{F}^{r}]=\left[\begin{array}{ccc}*\mathcal{F}^{01} & *\mathcal{F}^{02} & *\mathcal{F}^{03}\end{array}\right]=-\left[\begin{array}{ccc}\mathcal{F}^{01} & \mathcal{F}^{02} & \mathcal{F}^{03}\end{array}\right](\det P')$

$[*\mathcal{F}^{w}]=\left[\begin{array}{ccc}*\mathcal{F}^{32} & *\mathcal{F}^{13} & *\mathcal{F}^{21}\end{array}\right]=-\left[\begin{array}{ccc}\mathcal{F}^{32} & \mathcal{F}^{13} & \mathcal{F}^{21}\end{array}\right](\det P')$

$$[*\mathcal{F}]=\left[\begin{array}{cc}0 & [*\mathcal{F}^{r}]\\ -[*\mathcal{F}^{r}]^{t} & j\left([*\mathcal{F}^{w}]\right)\end{array}\right]=-[\mathcal{F}^{*}]\det P'=-[g]^{-1}[\mathcal{F}][g]^{-1}\det P'$$

$\left[\begin{array}{cc}[*\mathcal{F}^{w}] & [*\mathcal{F}^{r}]\end{array}\right]=-\left[\begin{array}{cc}[\mathcal{F}^{r}] & [\mathcal{F}^{w}]\end{array}\right][L_{H}]\det P'$

$$[L_{H}]=\left[\begin{array}{cc}[h]^{-1}\det h & -j(H_{0})[h]\\ [h]\,j(H_{0}) & H_{0}^{0}[h]-[H_{0}][H_{0}]^{t}\end{array}\right]$$

$[H]_{0}^{0}=[H_{0}]^{t}[h]^{-1}[H_{0}]+(\det g)\det[h]^{-1}$

Standard chart :

$[*\mathcal{F}^{r}]=[\mathcal{F}^{w}][g_{3}]^{-1}\det Q'$

$[*\mathcal{F}^{w}]=-[\mathcal{F}^{r}][g_{3}]\det Q$

Exterior differential of a 2 form

$$d\{\mathcal{F}_{32}d\xi^3\wedge d\xi^2+\mathcal{F}_{13}d\xi^1\wedge d\xi^3+\mathcal{F}_{21}d\xi^2\wedge d\xi^1+\mathcal{F}_{01}d\xi^0\wedge d\xi^1+\mathcal{F}_{02}d\xi^0\wedge d\xi^2+\mathcal{F}_{03}d\xi^0\wedge d\xi^3\}$$
$$=(-\partial_0\mathcal{F}_{21}+\partial_2\mathcal{F}_{01}-\partial_1\mathcal{F}_{02})\,d\xi^0\wedge d\xi^1\wedge d\xi^2+(\partial_0\mathcal{F}_{13}+\partial_3\mathcal{F}_{01}-\partial_1\mathcal{F}_{03})\,d\xi^0\wedge d\xi^1\wedge d\xi^3$$
$$+(-\partial_0\mathcal{F}_{32}+\partial_3\mathcal{F}_{02}-\partial_2\mathcal{F}_{03})\,d\xi^0\wedge d\xi^2\wedge d\xi^3-(\partial_1\mathcal{F}_{32}+\partial_2\mathcal{F}_{13}+\partial_3\mathcal{F}_{21})\,d\xi^1\wedge d\xi^2\wedge d\xi^3$$
$$d\left(*\mathcal{F}\right)=\sum_{\alpha=0}^{3}\left(-1\right)^{\alpha}\left(\sum_{\beta=0}^{3}\partial_\beta\left(\mathcal{F}^{\alpha\beta}\det P'\right)\right)d\xi^0\wedge...\widehat{d\xi^\alpha}...\wedge d\xi^3$$
$$=\left(\sum_{\beta=1}^{3}\partial_\beta\left(-\left[*\mathcal{F}^r\right]_\beta\right)\right)d\xi^1\wedge d\xi^2\wedge d\xi^3$$
$$-\left(\partial_0\left[*\mathcal{F}^r\right]_1+\partial_2\left[*\mathcal{F}^w\right]_3-\partial_3\left[*\mathcal{F}^w\right]_2\right)d\xi^0\wedge d\xi^2\wedge d\xi^3$$
$$+\left(\partial_0\left[*\mathcal{F}^r\right]_2-\partial_1\left[*\mathcal{F}^w\right]_3+\partial_3\left[*\mathcal{F}^w\right]_1\right)d\xi^0\wedge d\xi^2\wedge d\xi^3$$
$$-\left(\partial_0\left[*\mathcal{F}^r\right]_3+\partial_1\left[*\mathcal{F}^w\right]_2-\partial_2\left[*\mathcal{F}^w\right]_1\right)d\xi^0\wedge d\xi^1\wedge d\xi^2$$

Transport of a 2 form by the flow of a vector field

$$\left[\widetilde{\delta\mathcal{F}}^a\left(\Phi_V\left(\tau,a\right)\right)\right]=\left[K\left(\tau\right)\right]^t\left[\delta\mathcal{F}^a\left(a\right)\right]\left[K\left(\tau\right)\right]$$

Killing vector fields and isometries

$\pounds_V g=0$

$$\left[g\left(q\left(\tau\right)\right)\right]=\left[K\left(\tau\right)\right]^t\left[g\left(q\left(0\right)\right)\right]\left[K\left(\tau\right)\right]$$
$$\alpha,\beta=0...3:\sum_{\gamma=0}^{3}V^\gamma\left[\partial_\gamma g\right]^\alpha_\beta+\left[g\right]^\beta_\gamma\left[\partial_\alpha V\right]^\gamma+\left[g\right]^\alpha_\gamma\left[\partial_\beta V\right]^\gamma=0$$

Isometries :

$F:M\to M::g\left(F\left(m\right)\right)=F\left(m\right)$

$\forall u,v\in T_mM:\langle F'\left(m\right)u,F'\left(m\right)v\rangle=\langle u,v\rangle$

$\forall\lambda\in\Lambda_r\left(TM;\mathbb{R}\right):*\left(F^*\lambda\right)=F^*\left(*\lambda\right)$

A.5.4 RELATIVIST GEOMETRY

Standard chart

$\varphi_o:\mathbb{R}^4\to\Omega::\varphi_o\left(ct,\xi^1,\xi^2,\xi^3\right)=\Phi_{c\varepsilon_0}\left(t,x\right)$

$\varphi_\Omega:\mathbb{R}^3\to\Omega_3\left(0\right)::x=\varphi_\Omega\left(\xi^1,\xi^2,\xi^3\right)$

$\xi^0=ct$

$\partial\xi_0\left(m\right)=\varepsilon_0\left(m\right)$

World line of the observer : $p_o\left(t\right)=\varphi_o\left(ct,\xi^1,\xi^2,\xi^3\right)=\Phi_{c\varepsilon_0}\left(t,x\right)$

Tetrad

$\varepsilon_i\left(m\right)=\sum_{\alpha=0}^{3}P_i^\alpha\left(m\right)\partial\xi_\alpha\Leftrightarrow\partial\xi_\alpha=\sum_{i=0}^{3}P'^i_\alpha\left(m\right)\varepsilon_i\left(m\right)$

$\varepsilon^i\left(m\right)=\sum_{i=0}^{3}P'^i_\alpha\left(m\right)d\xi^\alpha\Leftrightarrow d\xi^\alpha=\sum_{i=0}^{3}P_i^\alpha\left(m\right)\varepsilon^i\left(m\right)$

$\left[\widetilde{P}\left(m\right)\right]=\left[P\left(m\right)\right]\left[\chi\left(m\right)\right]^{-1}$

$$\left[P'\right]=\begin{bmatrix}P'_{00}&P'_{10}&P'_{20}&P'_{30}\\P'_{10}&P'_{11}&P'_{12}&P'_{13}\\P'_{20}&P'_{21}&P'_{22}&P'_{23}\\P'_{30}&P'_{31}&P'_{32}&P'_{33}\end{bmatrix};\left[P\right]=\begin{bmatrix}P_{00}&P_{01}&P_{02}&P_{03}\\P_{10}&P_{11}&P_{12}&P_{13}\\P_{20}&P_{21}&P_{22}&P_{23}\\P_{30}&P_{31}&P_{32}&P_{33}\end{bmatrix}$$

Standard chart :

$$\left[P\right]=\begin{bmatrix}1&0\\0&Q\end{bmatrix};\left[Q\right]=\begin{bmatrix}P_{11}&P_{12}&P_{13}\\P_{21}&P_{22}&P_{23}\\P_{31}&P_{32}&P_{33}\end{bmatrix}$$
$$\left[P'\right]=\begin{bmatrix}1&0\\0&Q'\end{bmatrix};\left[Q'\right]=\begin{bmatrix}P'_{11}&P'_{12}&P'_{13}\\P'_{21}&P'_{22}&P'_{23}\\P'_{31}&P'_{32}&P'_{33}\end{bmatrix}$$

$\left[Q\right]\left[Q'\right]=I_3$

Metric

$$[g] = \begin{bmatrix} g_{00} & g_{01} & g_{02} & g_{03} \\ g_{01} & g_{11} & g_{21} & g_{13} \\ g_{02} & g_{21} & g_{22} & g_{32} \\ g_{03} & g_{13} & g_{32} & g_{33} \end{bmatrix} = [g]^t$$

$$[g] = [P']^t [\eta] [P'] \Leftrightarrow [g]^{-1} = [P] [\eta] [P]^t$$

$$\sqrt{-\det [g]} = \det P'$$

$$\partial_\alpha \det P' = (\det P') Tr ([\partial_\alpha P'] [P])$$

Standard chart :

$$[g] = \begin{bmatrix} -1 & 0 \\ 0 & [g]_3 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & [Q']^t [Q'] \end{bmatrix}$$

$$[g]^{-1} = \begin{bmatrix} -1 & 0 \\ 0 & [g]_3^{-1} \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & [Q] [Q]^t \end{bmatrix}$$

$$\det [g_3] = -\det g = (\det Q')^2$$

$$\det [g_3]^{-1} = (\det Q)^2$$

$$\varpi_4 = \det [P'] d\xi^0 \wedge d\xi^1 \wedge d\xi^2 \wedge d\xi^3$$

$$\varpi_3 = \det [Q'] d\xi^1 \wedge d\xi^2 \wedge d\xi^3$$

Fiber bundles

i) $P_G (M, Spin_0 (3,1), \pi_G)$:

$$\mathbf{p} (m) = \varphi_G (m, 1) \to \widetilde{\mathbf{p}} (m) = \mathbf{p} (m) \cdot \chi (m)^{-1} :$$

$$\sigma (m) = \mathbf{p} (m) \cdot \sigma (m) = \widetilde{\mathbf{p}} (m) \cdot \widetilde{\sigma} (m) = \widetilde{\mathbf{p}} (m) \cdot (\chi (m) .\sigma (m))$$

$$\sigma (m) \to \widetilde{\sigma} (m) = \chi (m) \cdot \sigma (m)$$

ii) $Cl (TM)$:

$$(\mathbf{p} (m), X) \sim \left(\widetilde{\mathbf{p}} (m), \mathbf{Ad}_{\chi(m)} X\right)$$

ii) $P_G \left[\mathbb{R}^4, \mathbf{Ad}\right]$:

$$\varepsilon_i (m) = (\mathbf{p} (m), \varepsilon_i) \to \widetilde{\varepsilon}_i (m) = \mathbf{Ad}_{\chi(m)^{-1}} \varepsilon_i (m) = \sum_{j=0}^3 \left[h \left(\chi (m)^{-1}\right)\right]_i^j \varepsilon_j (m)$$

$$v_m = \sum_{i=1}^n v^i \varepsilon_i (m) = \sum_{i=1}^n \widetilde{v}^i \widetilde{\varepsilon}_i (m) \Rightarrow \widetilde{v}^i = \sum_j \left[\mathbf{Ad}_{\chi(m)}\right]_j^i v^j$$

$$\left[\widetilde{P' (m)}\right] = [h (\chi (m))] [P' (m)]$$

$$\left[\widetilde{P} (m)\right] = [P (m)] [\chi (m)]^{-1}$$

Spatial coordinates defined by Killing vector fields

$$g_{32} = a_1 (t, \xi_1, \xi_2, \xi_3) \sqrt{g_{22} (\xi_2) g_{33} (\xi_3)}$$

$$g_{13} = a_2 (t, \xi_1, \xi_2, \xi_3) \sqrt{g_{11} (\xi_1) g_{33} (\xi_3)}$$

$$g_{21} = a_3 (t, \xi_1, \xi_2, \xi_3) \sqrt{g_{22} (\xi_2) g_{11} (\xi_1)}$$

$$\partial_\gamma a_p = -(g_{\gamma\gamma})^{1/2} \partial_0 a_p, \gamma, p = 1, 2, 3$$

$$[g_3] = \begin{bmatrix} \lambda_1^2 & a_3\lambda_1\lambda_2 & a_2\lambda_1\lambda_3 \\ a_3\lambda_1\lambda_2 & \lambda_2^2 & a_1\lambda_2\lambda_3 \\ a_2\lambda_1\lambda_3 & a_1\lambda_2\lambda_3 & \lambda_3^2 \end{bmatrix}$$

$$\det [g_3] = \lambda_3^2\lambda_2^2\lambda_1^2 \left(2a_1a_2a_3 - a_1^2 - a_2^2 - a_3^2 + 1\right)$$

$$[g_3]^{-1} = \frac{1}{a_1^2+a_2^2+a_3^2-1-2a_1a_2a_3} \begin{bmatrix} \frac{a_1^2-1}{\lambda_1^2} & \frac{a_3-a_1a_2}{\lambda_2\lambda_1} & \frac{a_2-a_1a_3}{\lambda_3\lambda_1} \\ \frac{a_3-a_1a_2}{\lambda_2\lambda_1} & \frac{a_2^2-1}{\lambda_2^2} & \frac{a_1-a_2a_3}{\lambda_3\lambda_2} \\ \frac{a_2-a_1a_3}{\lambda_3\lambda_1} & \frac{a_1-a_2a_3}{\lambda_3\lambda_2} & \frac{a_3^2-1}{\lambda_3^2} \end{bmatrix}$$

A.5.5 PARTICLES

Trajectory of a particle

Velocity u of the particle, measured in its proper time :

$u=\frac{dp}{d\tau}$

$\langle u,u\rangle=-c^2$

Trajectory of a particle, in the standard chart of an observer :

$q(t)=\varphi_o\left(\xi^0(t),\xi^1(t),\xi^2(t),\xi^3(t)\right)$

$\xi^0(t)=ct$

Velocity V of a particle as measured by an observer :

$V(t)=\frac{dq}{dt}=c\varepsilon_0(q(t))+\overrightarrow{v}$

$\overrightarrow{v}=\sum_{\alpha=1}^3\frac{d\xi^\alpha}{dt}\partial\xi_\alpha$

Between the proper time τ of a particle and the time t of an observer :

$u=\frac{dq}{d\tau}=\frac{1}{\sqrt{1-\frac{\|\overrightarrow{v}\|^2}{c^2}}}\left(\overrightarrow{v}+c\varepsilon_0(m)\right)$

$\frac{d\tau}{dt}=\sqrt{1-\frac{\|\overrightarrow{v}\|^2}{c^2}}=\frac{1}{c}\sqrt{-\langle V,V\rangle}$

Motion

$e_i=\mathbf{Ad}_\sigma\varepsilon_i$

$U=c\left(\varepsilon_0+\overrightarrow{u}\right)=-\frac{c}{\langle\mathbf{Ad}_\sigma\varepsilon_0,\varepsilon_0\rangle_{Cl}}\mathbf{Ad}_\sigma\varepsilon_0$

$V=\frac{dq}{dt}=c\varepsilon_0+\overrightarrow{v}=\sum_{\alpha,j}[P]_j^\alpha[U]^j\partial\xi_\alpha$

$\langle\overrightarrow{u},\overrightarrow{u}\rangle=1-\left(\frac{1}{\langle\mathbf{Ad}_\sigma\varepsilon_0,\varepsilon_0\rangle_{Cl}}\right)^2$

$\frac{d\sigma}{dt}\cdot\sigma^{-1}=\upsilon(X_r,X_w)\in T_1Spin(3,1)$

$\forall i=0..3:\frac{de_i}{dt}=[\upsilon(X_r,X_w),e_i]$

$\frac{dU}{dt}=\frac{U}{c}\langle[\upsilon(X_r,X_w),U],\varepsilon_0\rangle_{Cl}+[\upsilon(X_r,X_w),U]$

$u^t\frac{du}{dt}=(1-(u^tu))(u^tX_w)$

$\frac{dV^\alpha}{dt}=\sum_{i=0}^3[P']_i^\alpha\frac{dU^i}{dt}$

chart : $\sigma=\sigma_w\cdot\sigma_r=\epsilon(a_w+\upsilon(0,w))\cdot\epsilon(a_r+\upsilon(r,0))$

$U=c\left(\varepsilon_0+\sum_{a=1}^3\frac{a_w}{2a_w^2-1}w_a\varepsilon_a\right)$

$X_r=-\frac{1}{2}j(w)\frac{dw}{dt}+\left[1-\frac{1}{2}j(w)j(w)\right]\left(\frac{1}{a_r}+\frac{1}{2}j(r)+\frac{1}{4a_r}j(r)j(r)\right)\frac{dr}{dt}$

$X_w=\frac{1}{a_w}\left(1-\frac{1}{4}j(w)j(w)\right)\frac{dw}{dt}+[a_wj(w)]\left(\frac{1}{a_r}+\frac{1}{2}j(r)+\frac{1}{4a_r}j(r)j(r)\right)\frac{dr}{dt}$

$\frac{dU}{dt}=cX_w+\left(j(X_r)-(X_w^tv)\frac{1}{c}\right)v$

$a_w\simeq\epsilon\left(1+\frac{1}{8}\frac{\|\overrightarrow{v}\|^2}{c^2}\right)$

$w\simeq\epsilon\left(1+\frac{3}{8}\frac{\|\overrightarrow{v}\|^2}{c^2}\right)\frac{\overrightarrow{v}}{c}$

$V\simeq c\left(\varepsilon_0+\epsilon\left(1-\frac{3}{8}\frac{\|\overrightarrow{v}\|^2}{c^2}\right)\overrightarrow{w}\right)$

complex chart : $\sigma=A+Z$

$U=c\left(\varepsilon_0-\frac{1}{A\overline{A}+\frac{1}{4}Z^t\overline{Z}}\operatorname{Im}\left\{\left(A+\frac{1}{4}j(Z)\right)\overline{Z}\right\}\right)$

$\frac{d\sigma}{dt}\cdot\sigma^{-1}=D(Z)\frac{dZ}{dt}=Y_r+iY_w$

$\frac{dZ}{dt}=\left(A-\frac{1}{2}j(Z)\right)(Y_r+iY_w)$

$\frac{dU}{dt}=cY_w+\left(j(Y_r)-\left([Y_w]^t[v]\right)\frac{1}{c}\right)$

$A \simeq \epsilon a_r \left(1 + \frac{1}{8}\frac{\|\overrightarrow{v}\|^2}{c^2}\right) - i\frac{1}{4}\epsilon r^t \frac{\overrightarrow{v}}{c}$
$Z \simeq \epsilon \left(1 + \frac{1}{8}\frac{\|\overrightarrow{v}\|^2}{c^2}\right) r + i\left(a_r - \frac{1}{2}j(r)\right)\epsilon\frac{\overrightarrow{v}}{c}$

Deformable solid

$\sigma \in \mathfrak{X}(P_G) \to J^1\sigma = \left(m, \sigma(m), \partial_\alpha\sigma \cdot \sigma^{-1}, \alpha = 0...3\right) \in J^1Cl(TM)$
$V = \frac{dq}{dt} = c\varepsilon_0 + \overrightarrow{v} = -\frac{c}{\langle \mathbf{Ad}_\sigma\varepsilon_0, \varepsilon_0\rangle_{Cl}}\mathbf{Ad}_\sigma\varepsilon_0$
$\forall i, \alpha = 0..3:$
$\partial_\alpha e_i = [\upsilon(X_{r\alpha}, X_{w\alpha}), e_i]$
$\partial_\alpha V = \frac{V}{c}\langle[\upsilon(X_{r\alpha}, X_{w\alpha}), V], \varepsilon_0\rangle_{Cl} + [\upsilon(X_{r\alpha}, X_{w\alpha}), V]$
$\alpha > 0 : \partial_\beta V^\alpha = \sum_{j=0}^3 \left(P_j^\alpha - \frac{1}{c}P_j^0 V^\alpha\right)\left[\partial_\beta\sigma \cdot \sigma^{-1}, U\right]^j$
Rigid solid :
$\sigma(\Phi_V(t,x)) = s(t) \cdot \sigma(\Phi_V(0,x))$ with $s(t) \in Spin(3,1)$

Spinor

Spinor bundle $P_G[E, \gamma C]$
$\mathbf{p}(m) = \varphi_G(m,1) \to \widetilde{\mathbf{p}}(m) = \mathbf{p}(m) \cdot \chi(m)^{-1}$:
$\mathbf{e}_i(m) = (\mathbf{p}(m), e_i) \to \widetilde{\mathbf{e}}_i(m) = \gamma C\left(\chi(m)^{-1}\right)\mathbf{e}_i(m)$
$(\mathbf{p}(m), S) \sim (\widetilde{\mathbf{p}}(m), \gamma C(\chi(m))S)$
Scalar product :
$\langle S, S'\rangle_E = [S]^*[\gamma_0][S'] = i\left([S_L]^*[S'_R] - [S_R]^*[S'_L]\right)$
$\langle S, S\rangle_E = -2\operatorname{Im}\left([S_L]^*[S_R]\right)$
Single particle :
$j^1S : \mathbb{R} \to J^1P_G[E, \gamma C] :: j^1S(t) = (q(t), S(t), \delta S(t))$
$S : \mathbb{R} \to J^1P_G[E, \gamma C] :: \left(q(t), S(t), \frac{dS}{dt}(t)\right)$
$\mathcal{M}(q(t), \sigma(t), \upsilon(X_r, X_w)) = (q, S = \gamma C(\sigma)S_0, \delta S = \gamma C(\upsilon(X_r, X_w))S)$
$E_\epsilon = \left\{\begin{bmatrix} S_R \\ S_L \end{bmatrix} \in E : S_L = \epsilon i S_R\right\}$
Mass and kinetic energy :
$M_p = \sqrt{|\langle S_0, S_0\rangle|} = \sqrt{2\left|\operatorname{Im}\left([S_L]^*[S_R]\right)\right|} = \sqrt{2[S_R]^*[S_R]}$
$S_R = \frac{M_p}{\sqrt{2}}\begin{bmatrix} e^{i\alpha_1}\cos\alpha_0 \\ e^{i\alpha_2}\sin\alpha_0 \end{bmatrix}$
$\delta K = \frac{1}{M_p}\frac{1}{i}\langle S, \delta S\rangle = \frac{1}{M_p}\frac{1}{i}\langle S_0, \gamma C(\mathbf{Ad}_{\sigma^{-1}}\upsilon(X_r, X_w))S_0\rangle$
$\frac{dK}{dt} = \frac{1}{M_p}\frac{1}{i}\left\langle S_0, \gamma C\left(\sigma^{-1} \cdot \frac{d\sigma}{dt}\right)S_0\right\rangle$
Inertial vector :
$a = 1,2,3 : k^a = S_L^*\sigma_a S_R = \frac{1}{2}i\langle S_0, (\gamma_0\gamma_a - \widetilde{\gamma}_a)S_0\rangle_E$
$\forall Z \in T_1Spin(3,1) : \langle S_0, \gamma C(Z)S_0\rangle = i\operatorname{Im}k^tZ$
$\delta K = \frac{1}{M_p}\frac{1}{i}\langle S_0, \gamma C(\mathbf{Ad}_{\sigma^{-1}}\upsilon(X_r, X_w))S_0\rangle = \frac{1}{M_p}\operatorname{Im}k^t\mathbf{Ad}_{\sigma^{-1}}\upsilon(X_r, X_w)$
$S_0 \in E_0 : k = -i\epsilon\frac{M_p^2}{2}k_0$
$k_0 = \begin{bmatrix} (\sin 2\alpha_0)\cos(\alpha_1 - \alpha_2) \\ -(\sin 2\alpha_0)\sin(\alpha_1 - \alpha_2) \\ \cos 2\alpha_0 \end{bmatrix}; k_0^tk_0 = 1$
$\delta K = -\epsilon\frac{M_p}{2}k_0^t\operatorname{Re}\mathbf{Ad}_{\sigma^{-1}}\upsilon(X_r, X_w)$
Continuity equation :
$\frac{d\mu}{dt} + \mu divV = 0$

$divV = \sum_{\alpha,j=1}^{3} Q_j^\alpha \left\{ \left[\partial_\alpha \sigma \cdot \sigma^{-1}, U\right]^j + \frac{1}{2} \frac{1}{\det g} U^j \partial_\alpha \left(\det g\right) \right\}$

Charged Particles

Fiber bundles :

i) P_U

$\mathbf{p}_U(m) = \varphi_U(m,1) \rightarrow \widetilde{\mathbf{p}}_U(m) = \widetilde{\varphi}_U(m,1) = \mathbf{p}_U(m) \cdot \chi(m)^{-1}$

$\varkappa(m) = \varphi_U(m, \varkappa(m)) = \widetilde{\varphi}_U(m, \chi(m) \cdot \varkappa(m))$

ii) $P_U[F, \varrho]$

$\mathbf{f}_j(m) = (\mathbf{p}(m), f_j) \rightarrow \widetilde{\mathbf{f}}_j(m) = \varrho\left(\chi(m)^{-1}\right)(\mathbf{f}_j(m))$

$\phi(m) \rightarrow \widetilde{\phi}(m) = \varrho(\chi(m))\phi(m)$

iii) $Q(M, Spin(3,1) \times U, \pi_U)$

$(\varphi_Q(m,(1,1)), \psi) \sim \left(\varphi_Q\left(m, \left(s^{-1}, g^{-1}\right)\right), \vartheta(s,g)\psi\right)$

$(\mathbf{e}_i(m) \otimes \mathbf{f}_j(m))_{i=0..3}^{j=1...n} = (\varphi_Q(m,(1,1)), e_i \otimes f_j)$

$\mathbf{q}(m) = \varphi_Q(m,(1,1)) \rightarrow \widetilde{\mathbf{q}}(m) = \widetilde{\varphi}_Q(m,(1,1)) = \mathbf{q}(m) \cdot \chi(m)^{-1}$

$(\sigma(m), \varkappa(m)) = \varphi_Q(m,(\sigma,\varkappa)) = \widetilde{\varphi}_Q(m,(\widetilde{\sigma},\widetilde{\varkappa})) : (\widetilde{\sigma},\widetilde{\varkappa}) = \chi(m) \cdot (\sigma,\varkappa)$

iv) State of a particle : $Q[E \otimes F, \vartheta]$

$\mathbf{e}_i(m) \otimes \mathbf{f}_j(m) = (\mathbf{p}(m), e_i \otimes f_j) \rightarrow \widetilde{\mathbf{e}}_i(m) \otimes \widetilde{\mathbf{f}}_j(m) = \vartheta\left(\chi(m)^{-1}\right)(\mathbf{e}_i(m) \otimes \mathbf{f}_j(m))$

$\psi(m) = \sum_{i=1}^{4} \sum_{j=1}^{n} [\gamma C(\sigma(m))]_k^i [\varrho(\varkappa(m))]_l^j \psi_0^{kl}(m) \mathbf{e}_i(m) \otimes \mathbf{f}_j(m)$

$[\psi]_{4\times n} = [\gamma C(\sigma)]_{4\times 4} [\psi]_{4\times n} [\rho(\varkappa)]_{n\times n}$

$[\psi(m)] \rightarrow \left[\widetilde{\psi}(m)\right] = \vartheta(\chi(m))[\psi(m)] = [\gamma C(s)][\psi][\varrho(g)]$

$[\psi_0] = \begin{bmatrix} \psi_R \\ \epsilon i \psi_R \end{bmatrix}$

$\langle \psi, \psi' \rangle = Tr\left([\psi]^* [\gamma_0] [\psi']\right)$

Mass : $M_p = \sqrt{\epsilon \langle \psi_0, \psi_0 \rangle} = \sqrt{\epsilon 2 Tr\left(\psi_R^* \psi_R\right)}$

Momentum : $\mathcal{M} = (m, \psi = \vartheta(\sigma,\varkappa)\psi_0, \delta\psi = \vartheta(\upsilon(X_r, X_w) \cdot \sigma, \varkappa)\psi_0) \in J^1 Q[E \otimes F, \vartheta]$

Inertial vector :

$k^a = Tr[\psi_L^*] \sigma_a [\psi_R] = \frac{1}{2} \langle \psi_0, (\gamma_0 \gamma_a - i\widetilde{\gamma}_a)\psi_0 \rangle_E$

$k = -i\epsilon \frac{M_p^2}{2} k_0$

$\langle \psi_0, \vartheta(Z,1)\psi_0 \rangle = i \operatorname{Im} k^t Z = -i\epsilon \frac{M_p^2}{2} k_0^t \operatorname{Re} Z$

Kinetic Energy :

$\delta K = \frac{1}{M_p} \frac{1}{i} \langle \psi, \delta\psi \rangle = \frac{1}{M_p} \frac{1}{i} \langle \psi_0, \vartheta(\mathbf{Ad}_{\sigma^{-1}} \upsilon(X_r, X_w), 1)\psi_0 \rangle$

$\delta K = -\frac{1}{2} \epsilon M_p k_0^t \operatorname{Re}\left(\mathbf{Ad}_{\sigma^{-1}} \upsilon(X_r, X_w)\right)$

Connections

Potential :

$G \in \Lambda_1(M; T_1 Spin(3,1)) : TM \rightarrow T_1 Spin(3,1) ::$

$G(m) = \sum_{a=1}^{6} \sum_{\alpha=0}^{3} G_\alpha^a(m) \overrightarrow{\kappa}_a \otimes d\xi^\alpha = \sum_{\alpha=0}^{3} \upsilon(G_{r\alpha}(m), G_{w\alpha}(m)) d\xi^\alpha$

$\mathbf{p}(m) \rightarrow \widetilde{\mathbf{p}}(m) = \mathbf{p}(m) \cdot \chi(m)^{-1} : G(m) \rightarrow \widetilde{G}(m) = \mathbf{Ad}_\chi\left(G(m) - L'_{\chi^{-1}}(\chi)\chi'(m)\right)$

$[\Gamma_{M\alpha}] = \sum_{a=1}^{6} G_\alpha^a [\kappa_a] = \begin{bmatrix} 0 & G_{w\alpha}^1 & G_{w\alpha}^2 & G_{w\alpha}^3 \\ G_{w\alpha}^1 & 0 & -G_{r\alpha}^3 & G_{r\alpha}^2 \\ G_{w\alpha}^2 & G_{r\alpha}^3 & 0 & -G_\alpha^1 \\ G_{w\alpha}^3 & -G_{r\alpha}^2 & G_{r\alpha}^1 & 0 \end{bmatrix}$

$\grave{A} \in \Lambda_1(M; T_1 U) : TM \rightarrow T_1 U :: \grave{A}(m) = \sum_{\alpha=0}^{3} \sum_{a=1}^{m} \grave{A}_\alpha^a(m) \overrightarrow{\theta}_a \otimes d\xi^\alpha$

$\mathbf{p}_U(m) \rightarrow \widetilde{\mathbf{p}}_U(m) = \widetilde{\varphi}_U(m,1) = \mathbf{p}_U(m) \cdot \chi(m)^{-1}$

$\grave{A}(m) \to \widetilde{\grave{A}}(m) = Ad_\chi\left(\grave{A}(m) - L'_{\chi^{-1}}(\chi)\chi'(m)\right)$
Covariant derivative :
$\nabla^G : \mathfrak{X}(P_G) \to \Lambda_1(M;T_1Spin) :: \nabla^G\sigma = \sum_{\alpha=0}^3 \mathbf{Ad}_{\sigma^{-1}}\left(\partial_\alpha\sigma\cdot\sigma^{-1} + G_\alpha\right)d\xi^\alpha$
$\nabla^G : J^1Cl(TM) \to J^1Cl(TM) : \nabla^G_\alpha j^1\sigma = \sum_{\alpha=0}^3 \mathbf{Ad}_{\sigma^{-1}}(\upsilon(X_{r\alpha}, X_{w\alpha}) + G_\alpha)d\xi^\alpha$
$\nabla^G\sigma \to \widetilde{\nabla^G\sigma} = \nabla^G\sigma$
$\nabla^S : \mathfrak{X}(P_G[E,\gamma C]) \to *_1(M;\mathfrak{X}(P_G[E,\gamma C]))$
$\nabla^S S = \sum_{\alpha=0}^3(\partial_\alpha S + \gamma C(G_\alpha)S)d\xi^\alpha = \sum_{\alpha=0}^3(\partial_\alpha S + \gamma C(\upsilon(G_{r\alpha}, G_{w\alpha}))S)d\xi^\alpha$
$\nabla^S S \to \widetilde{\nabla^S S} = \gamma C(\chi)\nabla^S S$
$\nabla^M V = \sum_{\alpha i=0}^3\left(\partial_\alpha V^i + \sum_{j=0}^3[\Gamma_{M\alpha}(m)]^i_j V^j\right)\varepsilon_i(m)\otimes d\xi^\alpha$
Total connection of a matter field :
$[\nabla_\alpha\psi] = \vartheta(\sigma,\varkappa)\left(\left[\gamma C\left(\mathbf{Ad}_{\sigma^{-1}}\left(\partial_\alpha\sigma\cdot\sigma^{-1} + G_\alpha\right)\right)\right][\psi_0] + [\psi_0]\left[Ad_\varkappa\left(\grave{A}_\alpha\right)\right]\right)$
$[\nabla_\alpha\psi] = \vartheta(\sigma,\varkappa)\left(\left[\gamma C\left(\nabla^G_\alpha\sigma\right)\right][\psi_0] + [\psi_0]\left[Ad_\varkappa\left(\grave{A}_\alpha\right)\right]\right)$
$\left[\nabla_\alpha j^1\psi\right] = \vartheta(\sigma,\varkappa)\left(\left[\gamma C\left(\mathbf{Ad}_{\sigma^{-1}}(\upsilon(X_{r\alpha}, X_{w\alpha}) + G_\alpha)\right)\right][\psi_0] + [\psi_0]\left[Ad_\varkappa\left(\grave{A}_\alpha\right)\right]\right)$
$[\nabla_V\psi] = \sum_{j=0}^3[U]^j\sum_{\alpha=0}^3[P]^\alpha_j[\nabla_\alpha\psi]$
Single particle :
$[\nabla_V\psi] = \vartheta(\sigma,\varkappa)\left(\left[\gamma C\left(\mathbf{Ad}_{\sigma^{-1}}\left(\upsilon(X_r, X_w) + \widehat{G}\right)\right)\right][\psi_0] + [\psi_0]\left[Ad_\varkappa\widehat{\grave{A}}(t)\right]\right)$
$\widehat{G}(t) = \upsilon\left(\widehat{G}_r(t), \widehat{G}_w(t)\right) = \sum_{j=0}^3[U(t)]^j\sum_{\alpha=0i}^3[P(q(t))]^\alpha_j\,\upsilon(G_{r\alpha}(q(t)), G_{w\alpha}(q(t)))$
$\widehat{\grave{A}}(t) = \sum_{j=0}^3[U(t)]^j\sum_{\alpha=0}^3[P(q(t))]^\alpha_j\,\grave{A}_\alpha(q(t))$
Energy :
$\delta E = \frac{1}{M_p}\frac{1}{i}\langle\psi, \nabla_V\psi\rangle$
$= \frac{1}{M_p}\frac{1}{i}\left\langle\psi_0, \left[\gamma C\left(\nabla^G_V\sigma\right)\right][\psi_0]\right\rangle + \frac{1}{M_p}\frac{1}{i}\left\langle\psi_0, [\psi_0]\left[Ad_\varkappa\left(\widehat{\grave{A}}\right)\right]\right\rangle$
$= -\frac{1}{2}\epsilon M_p\left\{k_0^t\,\mathrm{Re}\,\mathbf{Ad}_{\sigma^{-1}}\left(\upsilon(X_r, X_w) + \widehat{G}\right) + k_c^t\left(Ad_\varkappa\widehat{\grave{A}}\right)\right\}$
$a = 1...m : k_c^a = -2\epsilon\frac{1}{M_p^2}\frac{1}{i}\langle\psi_0, [\psi_0][\theta_a]\rangle$
$\delta K = \frac{1}{M_p}\frac{1}{i}\langle\psi, \delta\psi\rangle = \frac{1}{M_p}\frac{1}{i}\langle\psi_0, \vartheta(\mathbf{Ad}_{\sigma^{-1}}\upsilon(X_r, X_w), 1)\psi_0\rangle$
$\delta K = -\frac{1}{2}\epsilon M_p k_0^t\,\mathrm{Re}(\mathbf{Ad}_{\sigma^{-1}}\upsilon(X_r, X_w))$
EM field : $k_c = -2q$

A.5.6 FORCE FIELDS

Gravitational Field

Strength of the field

$\mathcal{F}_G = \sum_{a=1}^6\left(dG^a + \sum_{\alpha\beta=0}^3[G_\alpha, G_\beta]^a d\xi^\alpha\wedge d\xi^\beta\right)\otimes\overrightarrow{\kappa}_a$
$= \sum_{a=1}^6\sum_{\{\alpha,\beta\}=0}^3\left(\partial_\alpha G^a_\beta - \partial_\beta G^a_\alpha + 2[G_\alpha, G_\beta]^a\right)d\xi^\alpha\wedge d\xi^\beta\otimes\overrightarrow{\kappa}_a$
$= \sum_{a=1}^6\sum_{\alpha,\beta=0}^3\mathcal{F}^a_{G\alpha\beta}d\xi^\alpha\wedge d\xi^\beta\otimes\overrightarrow{\kappa}_a$
$= \sum_{\{\alpha,\beta\}=0}^3\upsilon(\mathcal{F}_{r\alpha\beta}, \mathcal{F}_{w\alpha\beta})d\xi^\alpha\wedge d\xi^\beta$
$[\mathcal{F}_{G\alpha\beta}] = \sum_{a=1}^6\mathcal{F}^a_{G\alpha\beta}[\kappa_a] = [K(\mathcal{F}_{w\alpha\beta})] + [J(\mathcal{F}_{r\alpha\beta})]$
$[\mathcal{F}_{\alpha\beta}] = [\partial_\alpha\Gamma_{M\beta}] - [\partial_\beta\Gamma_{M\alpha}] + [\Gamma_{M\alpha}][\Gamma_{M\beta}] - [\Gamma_{M\beta}][\Gamma_{M\alpha}]$
with the signature (3,1) :
$\mathcal{F}_{r\alpha\beta} = \upsilon(\partial_\alpha G_{r\beta} - \partial_\beta G_{r\alpha} + 2(j(G_{r\alpha})G_{r\beta} - j(G_{w\alpha})G_{w\beta}), 0)$
$\mathcal{F}_{w\alpha\beta} = \upsilon(0, \partial_\alpha G_{w\beta} - \partial_\beta G_{w\alpha} + 2(j(G_{w\alpha})G_{r\beta} + j(G_{r\alpha})G_{w\beta}))$
$\mathrm{Re}\,\mathcal{F}_{\alpha\beta} = \mathrm{Re}(\partial_\alpha G_\beta - \partial_\beta G_\alpha + 2j(G_\alpha)G_\beta)$
$\mathrm{Im}\,\mathcal{F}_{\alpha\beta} = \mathrm{Im}(\partial_\alpha G_\beta - \partial_\beta G_\alpha + 2j(G_\alpha)G_\beta)$

$\mathcal{F}_{\alpha\beta} = \partial_\alpha G_\beta - \partial_\beta G_\alpha + 2j(G_\alpha) G_\beta$
Change of gauge : $\mathbf{p}_G(m) = \varphi_G(m,1) \rightarrow \widetilde{\mathbf{p}}_G(m) = \mathbf{p}_G(m) \cdot s(m)^{-1}$:
$\mathcal{F}_{G\alpha\beta} \rightarrow \widetilde{\mathcal{F}}_{G\alpha\beta}(m) = \mathbf{Ad}_{s(m)} \mathcal{F}_{G\alpha\beta}$
Matrix representation :
$[\mathcal{F}]_{6\times 6} = \begin{bmatrix} \mathcal{F}_r^r & \mathcal{F}_r^w \\ \mathcal{F}_w^r & \mathcal{F}_w^w \end{bmatrix} = \left[\mathcal{F}^a_{G\alpha\beta}\right]$
$[\mathcal{F}_r^r]_{3\times 3} = \begin{bmatrix} \mathcal{F}^1_{G32} & \mathcal{F}^1_{G13} & \mathcal{F}^1_{G21} \\ \mathcal{F}^2_{G32} & \mathcal{F}^2_{G13} & \mathcal{F}^2_{G21} \\ \mathcal{F}^3_{G32} & \mathcal{F}^3_{G13} & \mathcal{F}^3_{G21} \end{bmatrix}$
$[\mathcal{F}_r^w]_{3\times 3} = \begin{bmatrix} \mathcal{F}^1_{G01} & \mathcal{F}^1_{G02} & \mathcal{F}^1_{G03} \\ \mathcal{F}^2_{G01} & \mathcal{F}^2_{G02} & \mathcal{F}^2_{G02} \\ \mathcal{F}^3_{G01} & \mathcal{F}^3_{G02} & \mathcal{F}^3_{G03} \end{bmatrix}$
$[\mathcal{F}_w^r]_{3\times 3} = \begin{bmatrix} \mathcal{F}^4_{G32} & \mathcal{F}^4_{G13} & \mathcal{F}^4_{G21} \\ \mathcal{F}^5_{G32} & \mathcal{F}^5_{G13} & \mathcal{F}^5_{G21} \\ \mathcal{F}^6_{G32} & \mathcal{F}^6_{G13} & \mathcal{F}^6_{G21} \end{bmatrix}$
$[\mathcal{F}_w^w]_{3\times 3} = \begin{bmatrix} \mathcal{F}^4_{G01} & \mathcal{F}^4_{G02} & \mathcal{F}^4_{G03} \\ \mathcal{F}^5_{G01} & \mathcal{F}^5_{G02} & \mathcal{F}^5_{G02} \\ \mathcal{F}^6_{G01} & \mathcal{F}^6_{G02} & \mathcal{F}^6_{G03} \end{bmatrix}$
Complex formalism :
$[\mathcal{F}_G^r] = [\mathcal{F}_w^r] + i[\mathcal{F}_w^r]$
$[\mathcal{F}_G^w] = [\mathcal{F}_w^w] + i[\mathcal{F}_w^w]$
$[G_0]_{3\times 1} = \begin{bmatrix} G_0^1 + iG_0^4 \\ G_0^2 + iG_0^5 \\ G_0^3 + iG_0^6 \end{bmatrix}$
$[G]_{3\times 3} = \begin{bmatrix} G_1^1 + iG_1^4 & G_2^1 + iG_2^4 & G_3^1 + iG_3^4 \\ G_1^2 + iG_1^5 & G_2^2 + iG_2^5 & G_3^2 + iG_3^5 \\ G_1^3 + iG_1^6 & G_2^3 + iG_2^6 & G_3^3 + iG_3^6 \end{bmatrix}$
$[\mathcal{F}_G] = \begin{bmatrix} \mathcal{F}_G^r & \mathcal{F}_G^w \end{bmatrix}_{3\times 6} = \begin{bmatrix} [dG^r] - 2(\det[G])[G]^{-1} & [dG^w] + 2[j(G_0)][G] \end{bmatrix}$
Hodge dual : $\begin{bmatrix} [*\mathcal{F}_G^w] & [*\mathcal{F}_G^r] \end{bmatrix} = -\begin{bmatrix} [\mathcal{F}_G^r] & [\mathcal{F}_G^w] \end{bmatrix} [L_H] \det P'$
Riemann tensor :
$\left[\widehat{R}_{\alpha\beta}\right] = [R_{\alpha\beta}] = [P][\mathcal{F}_{G\alpha\beta}][P'] \Leftrightarrow [\mathcal{F}_{G\alpha\beta}] = [P'][R_{\alpha\beta}][P]$
Ricci tensor : $Ric = \sum_{\alpha\beta=0}^3 \sum_{a=1}^6 ([\mathcal{F}^a][P][\kappa_a][P'])^\alpha_\beta \, d\xi^\alpha \otimes d\xi^\beta$
Scalar curvature :
$\mathbf{R} = \sum_{a=1}^6 Tr\left([P]^t [\mathcal{F}^a][P][\kappa_a][\eta]\right)$
$\mathbf{R} = Tr\left\{-2[\mathcal{F}_r^r][Q']^t \det[Q] - 2[\mathcal{F}_r^w][Q] j\left(\left[P^0\right]\right) + [\mathcal{F}_w^r] j([P_0])[Q] - [\mathcal{F}_w^w]\left(P_0^0[Q] - [P_0]\left[P^0\right]\right)\right\}$
Scalar product :
$\langle \mathcal{F}, K\rangle_G = \frac{1}{4}\left(G_2(\mathcal{F}_r, K_r) - G_2(\mathcal{F}_w, K_w)\right) = \frac{1}{4}\sum_{\{\alpha\beta\}} \mathcal{F}_r^{\alpha\beta} K_{r\alpha\beta} - \mathcal{F}_w^{\alpha\beta} K_{w\alpha\beta}$
$= \frac{1}{4}\frac{1}{\det P'}\left([*\mathcal{F}_w^w][K_w^r]^t + [*\mathcal{F}_w^r][K_w^w]^t - \left([*\mathcal{F}_r^w][K_r^r]^t + [*\mathcal{F}^r][K^w]^t\right)\right)$
$= -\frac{1}{8} Tr\left([\mathcal{F}_r][g]^{-1}[K_r][g]^{-1} - [\mathcal{F}_w][g]^{-1}[K_w][g]^{-1}\right)$
$= -\frac{1}{4}\frac{1}{\det P'} \operatorname{Re}\left([*\mathcal{F}^w][K^r]^t + [*\mathcal{F}^r][K^w]^t\right) = -\frac{1}{8}\operatorname{Re} Tr\left([\mathcal{F}][g]^{-1}[K][g]^{-1}\right)$
Standard chart :
$\langle \mathcal{F}, K\rangle_G = \frac{1}{4}\sum_{a=1}^3 \operatorname{Re} Tr\left([\mathcal{F}^w][g_3]^{-1}[K^w]^t + [\mathcal{F}^r][g_3][K^r]^t \det[g_3]^{-1}\right)$
$\langle \mathcal{F}, K\rangle_G \varpi_4 = \frac{1}{4}\sum_{a=1}^3 *\mathcal{F}_r^a \wedge K_r^a - *\mathcal{F}_w^a \wedge K_w^a = \frac{1}{4}\sum_{a=1}^3 *K_r^a \wedge \mathcal{F}_r^a - *K_w^a \wedge \mathcal{F}_w^a$
$\langle X, [Y,Z]\rangle_G = \sum_{\{\alpha\beta\}} \left\langle X^{\alpha\beta}, [Y_\alpha, Z_\beta]\right\rangle_{Cl} = \sum_{\{\alpha\beta\}} \left\langle \left[X^{\alpha\beta}, Y_\alpha\right], Z_\beta\right\rangle_{Cl}$
Chern-Weil theorem :
$Tr\left([\mathcal{F}_r^r]^t[\mathcal{F}_r^w] - [\mathcal{F}_w^r]^t[\mathcal{F}_w^w]\right) = 0 \Leftrightarrow Tr \operatorname{Re}[\mathcal{F}_G^r]^t[\mathcal{F}_G^w] = 0$

Electromagnetic field

$\mathcal{F}_{\alpha\beta} = \sum_{\alpha} \left(\partial_{\alpha} \grave{A}_{\beta} - \partial_{\beta} \grave{A}_{\alpha} \right) d\xi^{\alpha} \wedge d\xi^{\beta}$

$\overrightarrow{E} = \sum_{i=1}^{3} E^i \varepsilon_i = \sum_{i=1}^{3} \sum_{\beta=0}^{3} E^i P_i^{\beta} \partial\xi_{\beta}$

$\overrightarrow{B} = \sum_{i=1}^{3} B^i \varepsilon_i = \sum_{i=1}^{3} \sum_{\beta=0}^{3} B^i P_i^{\beta} \partial\xi_{\beta}$

$[\mathcal{F}_{EM}^r] = -[B]^t [Q']^t \det Q + [E]^t [Q] j\left(\left[P^0\right]\right)$

$[\mathcal{F}_{EM}^w] = [B]^t j([P_0])[Q] + [E]^t \left(P_0^0 [Q] - [P_0]\left[P^0\right]\right)$

$\left[d\grave{A}^r\right] = [\mathcal{F}_{EM}^r] = -[B]^t [Q']^t \det Q + [E]^t [Q] j\left(\left[P^0\right]\right)$

$\left[d\grave{A}^w\right] = [\mathcal{F}_{EM}^w] = [B]^t j([P_0])[Q] + [E]^t \left(P_0^0 [Q] - [P_0]\left[P^0\right]\right)$

Scalar product : $\langle \mathcal{F}_{EM}, \mathcal{F}_{EM} \rangle = [E]^t [g_3][E] + [B][g_3][B]^t$

Other Fields

Strength of the field

$\mathcal{F}_A = \sum_{a=1}^{m} \sum_{\{\alpha,\beta\}} \left(\mathcal{F}_{A\alpha\beta}^a d\xi^{\alpha} \wedge d\xi^{\beta} \right) \otimes \overrightarrow{\theta}_a \in \Lambda_2(M; T_1U)$

$\mathcal{F}_A = \sum_{a=1}^{m} \left(d\left(\sum_{\alpha=0}^{3} \grave{A}_{\alpha}^a d\xi^{\alpha} \right) + \sum_{\alpha\beta} \left[\grave{A}_{\alpha}, \grave{A}_{\beta} \right] d\xi^{\alpha} \wedge d\xi^{\beta} \right) \otimes \overrightarrow{\theta}_a$

$\mathcal{F}_{A\alpha\beta}^a = \partial_{\alpha} \grave{A}_{\beta}^a - \partial_{\beta} \grave{A}_{\alpha}^a + 2\left[\grave{A}_{\alpha}, \grave{A}_{\beta} \right]^a$

Change of gauge : $\mathbf{p}_U(m) = \varphi_{P_U}(m,1) \to \widetilde{\mathbf{p}}_U(m) = \mathbf{p}_U(m) \cdot \varkappa(m)^{-1}$:

$\mathcal{F}_{A\alpha\beta} \to \widetilde{\mathcal{F}}_{A\alpha\beta}(m) = Ad_{\varkappa(m)} \mathcal{F}_{A\alpha\beta}$

Hodge dual : $\begin{bmatrix} [*\mathcal{F}_A^w] & [*\mathcal{F}_A^r] \end{bmatrix} = -\begin{bmatrix} [\mathcal{F}_A^r] & [\mathcal{F}_A^w] \end{bmatrix} [L_H] \det P'$

Scalar product :

$\langle \mathcal{F}, K \rangle_A = \sum_{a=1}^{m} G_2(\mathcal{F}^a, K^a) = \sum_{a=1}^{m} \sum_{\{\alpha\beta\}} \mathcal{F}^{a\alpha\beta} K_{\alpha\beta}^a$

$= -\frac{1}{\det P'} \sum_{a=1}^{m} \left([*\mathcal{F}^{aw}][K^{ar}]^t + [*\mathcal{F}^{ar}][K^{aw}]^t \right)$

$= -\frac{1}{\det P'} Tr\left([*\mathcal{F}^w][K^r]^t + [*\mathcal{F}^r][K^w]^t \right)$

$= -\frac{1}{2} \sum_{a=1}^{m} Tr\left([\mathcal{F}^a][g]^{-1}[K^a][g]^{-1} \right)$

Standard chart :

$\langle \mathcal{F}, K \rangle_A = Tr[\mathcal{F}_A^w][g_3]^{-1}[K^w]^t + [\mathcal{F}_A^r][g_3][K_A^r]^t \det[g_3]^{-1}$

$\langle \mathcal{F}, K \rangle_A \varpi_4 = \sum_{a=1}^{m} *\mathcal{F}_A^a \wedge K_A^a = \sum_{a=1}^{m} *K_A^a \wedge \mathcal{F}_A^a$

$\langle X, [Y,Z] \rangle_A = \sum_{\{\alpha\beta\}} \left\langle X^{\alpha\beta}, [Y_{\alpha}, Z_{\beta}] \right\rangle_{T_1U} = \sum_{\{\alpha\beta\}} \left\langle \left[X^{\alpha\beta}, Y_{\alpha}\right], Z_{\beta} \right\rangle_{T_1U}$

Chern-Weil theorem : $Tr\left([\mathcal{F}_A^r]^t [\mathcal{F}_A^w] \right) = 0$

Propagation

Conservation of energy in the vacuum :

$\frac{\partial}{\partial t} \langle \mathcal{F}, \mathcal{F} \rangle + \langle \mathcal{F}, \mathcal{F} \rangle \frac{1}{2} \frac{1}{\det g} \partial_0 (\det g) = 0$

Propagation along a curve of tangent V :

$V = v + c\varepsilon_0$

$\pounds_V g = 0$

$\langle V, V \rangle = w^2 - c^2$

Evolution of a signal along a propagation curve :

$[\delta\mathcal{F}^a(\tau)] = \theta(\tau)[K(\tau)]^t [\delta\mathcal{F}^a(O)][K(\tau)]$

$\begin{bmatrix} [\delta\mathcal{F}^r(\tau)] & [\delta\mathcal{F}^w(\tau)] \end{bmatrix} = \theta(\tau) \begin{bmatrix} [\delta\mathcal{F}^r(O)] & [\delta\mathcal{F}^w(O)] \end{bmatrix} \left[L_{K(\tau)}\right]$

$\frac{d}{d\tau} \begin{bmatrix} [\delta\mathcal{F}^r(\tau)] & [\delta\mathcal{F}^w(\tau)] \end{bmatrix} = \begin{bmatrix} [\delta\mathcal{F}^r(O)] & [\delta\mathcal{F}^w(O)] \end{bmatrix} \left[L_{K(\tau)}\right] \left(\frac{1}{\theta} \frac{d\theta}{d\tau} I_3 + [D(\tau)] \right)$

$\frac{d}{d\tau} \begin{bmatrix} [\delta\mathcal{F}^r(\tau)] & [\delta\mathcal{F}^w(\tau)] \end{bmatrix} = \begin{bmatrix} [\delta\mathcal{F}^r(\tau)] & [\delta\mathcal{F}^w(\tau)] \end{bmatrix} \frac{1}{\theta} \left(\frac{1}{\theta} \frac{d\theta}{d\tau} I_3 + [D(\tau)] \right)$

$$[D(\tau)] = \begin{bmatrix} \left([\partial v]^t - (Tr[\partial v]) I_3\right) & 0 \\ 0 & -[\partial v] \end{bmatrix}$$

$$[\partial v] = \begin{bmatrix} \partial_1 v^1 & \partial_2 v^1 & \partial_3 v^1 \\ \partial_1 v^2 & \partial_2 v^2 & \partial_3 v^2 \\ \partial_1 v^3 & \partial_2 v^3 & \partial_3 v^3 \end{bmatrix}$$

Potentials :

$\left[\delta \grave{A}(\tau)\right] = \theta(\tau)\left[\delta \grave{A}(O)\right][K(\tau)]$

$[\delta G(\tau)] = \theta(\tau)[\delta G(O)][K(\tau)]$

$\delta G_0^a(\tau) = \theta(\tau)\,\delta G_0^a(O)$

$\delta \grave{A}_0^a(\tau) = \theta(\tau)\,\delta \grave{A}_0^a(O)$

Bosons

Photon :$\Delta\varphi : [0,T] \to TM^* :: \Delta\varphi(t) = \sum_{\alpha,\beta=0}^3 [K(t)]_\alpha^\beta \Delta\varphi_\beta d\xi^\alpha(q(t))$

Graviton : $\Delta\Gamma : [0,T] \to TM^* \otimes L_0 :: \Delta\Gamma(t) = \sum_{\alpha,\beta=0}^3 [K(t)]_\alpha^\beta \Delta\Gamma_\beta d\xi^\alpha(q(t)) \otimes \overrightarrow{\kappa}_a$ with a equal either $1,2,3$

Antigraviton : $\Delta\overline{\Gamma} : [0,T] \to TM^* \otimes P_0 :: \Delta\overline{\Gamma}(t) = \sum_{\alpha,\beta=0}^3 [K(t)]_\alpha^\beta \Delta\overline{\Gamma}_\beta d\xi^\alpha(q(t)) \otimes \overrightarrow{\kappa}_a$ with a equal either $4,5,6$

Other bosons : $\Delta\grave{A} : [0,T] \to TM^* \otimes T_1U :: \Delta\grave{A}(t) = \sum_{\alpha,\beta=0}^3 [K(t)]_\alpha^\beta \Delta\grave{A}_\beta^a d\xi^\alpha(q(t)) \otimes \overrightarrow{\theta}_a$

A.5.7 LAGRANGIANS

Complex variables

$\frac{\partial L}{\partial z} = \frac{1}{2}\left(\frac{\partial L}{\partial \operatorname{Re} z} + \frac{1}{i}\frac{\partial L}{\partial \operatorname{Im} z}\right); \frac{\partial L}{\partial \overline{z}} = \frac{1}{2}\left(\frac{\partial L}{\partial \operatorname{Re} z} - \frac{1}{i}\frac{\partial L}{\partial \operatorname{Im} z}\right)$

$\frac{\partial L}{\partial \operatorname{Re} z} = \frac{\partial L}{\partial z} + \frac{\partial L}{\partial \overline{z}}; \frac{\partial L}{\partial \operatorname{Im} z} = i\left(\frac{\partial L}{\partial z} - \frac{\partial L}{\partial \overline{z}}\right)$

$\frac{\partial L}{\partial \operatorname{Re} z}\operatorname{Re} u + \frac{\partial L}{\partial \operatorname{Im} z}\operatorname{Im} u = 2\operatorname{Re}\frac{\partial L}{\partial z}u; -\frac{\partial L}{\partial \operatorname{Re} z}\operatorname{Im} u + \frac{\partial L}{\partial \operatorname{Im} z}\operatorname{Re} u = -2\operatorname{Im}\frac{\partial L}{\partial z}u$

Equivariance and covariance

$L = L(\psi, \nabla_\alpha\psi, P_i^\alpha, \mathcal{F}_{G\alpha\beta}, \mathcal{F}_{A\alpha\beta}, V^\alpha)$

Lagrange equations

$\forall z^i : \frac{d\left(L \det P'\right)}{dz^i} - \sum_\beta \frac{d}{d\xi^\beta}\frac{d\left(L \det P'\right)}{dz_\beta^i} = 0$

Tetrad equation

$\forall \alpha, \beta = 0...3 : \sum_i \frac{dL}{dP_i^\alpha} P_i^\beta - L\delta_\beta^\alpha = 0$

Energy-Momentum tensor

$T : \mathfrak{X}\left(J^1E\right) \to \mathfrak{X}(TM^* \otimes TM \otimes E^*) :: T = \sum_{i\alpha\beta}\frac{\partial L}{\partial z_\alpha^i} z_\beta^i \partial\xi_\alpha \otimes d\xi^\beta \otimes e^i$

$\Pi_i = \sum_{\alpha\beta}\frac{\partial L}{\partial z_\alpha^i} z_\beta^i \partial\xi_\alpha \otimes d\xi^\beta \otimes e^i \in \mathfrak{X}(TM^* \otimes TM \otimes E^*)$

$\delta\ell = \int_\Omega (div(T(V)) + Tr(T)(V))\varpi_4$

$T = \sum_{\alpha\beta}\{\sum_{ij}\frac{\partial L}{\partial \nabla_\alpha \psi^{ij}}\delta_\beta\psi^{ij} + 2\sum_{a,\gamma}\left(\frac{\partial L}{\partial \mathcal{F}_{A\alpha\gamma}^a}\delta_\beta\grave{A}_\gamma^a + \frac{\partial L}{\partial \mathcal{F}_{G\alpha\gamma}^a}\delta_\beta G_\gamma^a\right)\}\partial\xi_\alpha \otimes d\xi^\beta$

A.5.8 CONTINUOUS MODELS

Lagrangians

Interactions Fields / Fields

$$\int_\Omega \left(\sum_{\alpha\beta} C_G \left(\sum_{a=1}^3 \mathcal{F}^a_{G\alpha\beta}\mathcal{F}^{a\alpha\beta}_G - \sum_{a=4}^6 \mathcal{F}^a_{G\alpha\beta}\mathcal{F}^{a\alpha\beta}_G\right) + C_A \sum_{a=1}^m \mathcal{F}^a_{A\alpha\beta}\mathcal{F}^{a\alpha\beta}_A\right)\varpi_4(m)$$

$$\sum_{\alpha\beta}\left\{C_G \left(\sum_{a=1}^3 \mathcal{F}^a_{G\alpha\beta}\mathcal{F}^{a\alpha\beta}_G - \sum_{a=4}^6 \mathcal{F}^a_{G\alpha\beta}\mathcal{F}^{a\alpha\beta}_G\right) + C_A \sum_{a=1}^m \mathcal{F}^a_{A\alpha\beta}\mathcal{F}^{a\alpha\beta}_A + C_{EM}\mathcal{F}_{EM\alpha\beta}\mathcal{F}^{\alpha\beta}_{EM}\right\}$$

$$= \sum_{\alpha\beta} 4C_G \left\langle \mathcal{F}^{\alpha\beta}_G, \mathcal{F}_{G\alpha\beta}\right\rangle_{Cl} + C_A \left\langle \mathcal{F}^{\alpha\beta}_A, \mathcal{F}_{A\alpha\beta}\right\rangle_{T_1U} + C_{EM}\left\langle \mathcal{F}_{EM\alpha\beta}, \mathcal{F}^{\alpha\beta}_{EM}\right\rangle_{T_1U(1)}$$

$$= 8C_G \left\langle \mathcal{F}_G, \mathcal{F}_G\right\rangle_G + 2C_A \left\langle \mathcal{F}_A, \mathcal{F}_A\right\rangle_U + 2C_{EM}\left\langle \mathcal{F}_{EM\alpha\beta}, \mathcal{F}^{\alpha\beta}_{EM}\right\rangle_{T_1U(1)}$$

Particles

$$\delta E = C_I \frac{1}{M_p}\frac{1}{i}\left\langle \psi, \nabla_V \psi\right\rangle = -C_I \frac{1}{2} M_p \left\{k_0^t \operatorname{Re} \mathbf{Ad}_{\sigma^{-1}}\left(\upsilon(X_r, X_w) + \widehat{G}\right) + k_c^t \left(Ad_\varkappa \widehat{\grave{A}}\right)\right\}$$

$$[\nabla_\alpha \psi] = \vartheta(\sigma, 1)\left(\left[\gamma C\left(\nabla^G_\alpha \sigma\right)\right][\psi_0] + [\psi_0]\left[\grave{A}_\alpha\right]\right)$$

$$\nabla^G_\alpha \sigma = \mathbf{Ad}_{\sigma^{-1}}\left(\upsilon(X_{r\alpha}, X_{w\alpha}) + G_\alpha\right)$$

$$\upsilon(X_{r\alpha}, X_{w\alpha}) = \partial_\alpha \sigma \cdot \sigma^{-1}$$

Model with matter fields

$$\int_\Omega C_I \frac{1}{M}\frac{1}{i}\left\langle \psi(r(m), w(m)), \nabla_V \psi r(m), w(m)\right\rangle \mu(m)\varpi_4(m)$$

Model with individual particles

$$\psi_p(t) = \psi_p(q_p(t)) = \vartheta(\sigma_p(r_p(t), w_p(t)), 1)\psi_{0p}$$

$$\sum_{p=1}^N \int_0^T C_I \frac{1}{M_p}\frac{1}{i}\left\langle \psi(r_p(t), w_p(t)), \nabla_V \psi(r_p(t), w_p(t))\right\rangle dt$$

Particles

$$\sum_{\beta=0}^3 V^\beta \left\langle \psi, \nabla_\beta \psi\right\rangle = 0$$

$$\left\langle \psi_0, \left[\gamma C\left(\left[\sigma^{-1}\cdot \frac{\partial\sigma}{\partial r_a}, \nabla^G_V \sigma\right]\right)\right][\psi_0]\right\rangle = \sum_{\beta,j=0}^3 P^\beta_j \left[\frac{\partial\sigma}{\partial r_a}\cdot\sigma^{-1}, U\right]^j \left\langle \psi, \nabla_\beta \psi\right\rangle$$

$$\left\langle \psi_0, \left[\gamma C\left(\left[\sigma^{-1}\cdot \frac{\partial\sigma}{\partial w_a}, \nabla^G_V \sigma\right]\right)\right][\psi_0]\right\rangle = \sum_{\beta,j=0}^3 P^\beta_j \left[\frac{\partial\sigma}{\partial w_a}\cdot\sigma^{-1}, U\right]^j \left\langle \psi, \nabla_\beta \psi\right\rangle$$

$$k_0^t \operatorname{Re}\left([D(-Z)]\left[\frac{dZ}{dt}\right] + [Ad(-Z)]\left[\widehat{G}\right]\right) + k_c^t\left[\widehat{\grave{A}}\right] = 0$$

$$\frac{1}{c}[D(Z)][j(k_0)]\left\{[D(-Z)]\left[\frac{dZ}{dt}\right] + [Ad(-Z)]\left[\widehat{G}\right]\right\}$$

$$= -2\operatorname{Re}\left\{[D(-Z)]\left(i\left([u]^t[u] - 1\right) + j(u) + ij(u)j(u)\right)\right\}\sum_{\beta=1}^3\left([Q]^\beta\right)^t\left\{k_0^t \operatorname{Re}\left(\nabla^G_\beta \sigma\right) + k_c^t\left[\grave{A}_\beta\right]\right\}$$

$$\operatorname{Re}\left([D(Z)][j(k_0)][D(-Z)]\right)\frac{dw}{dt} + \operatorname{Im}\left([D(Z)][j(k_0)][D(-Z)]\right)\frac{dr}{dt}$$

$$= -\operatorname{Im}\left\{[D(Z)][j(k_0)][Ad(-Z)]\left[\widehat{G}\right]\right\}$$

Bonded particle :

$$\frac{dr}{dt} = \left(a_r + \frac{1}{2}j(r)\right)\left(2q\grave{A}_0 k_0 - G_{r0}\right)$$

Geodesics

$$\operatorname{Re}\left\{[D(-Z)]\left(i\left([u]^t[u] - 1\right) + j(u) + ij(u)j(u)\right)\right\}\sum_{\beta=1}^3\left([Q]^\beta\right)^t\left\{k_0^t \operatorname{Re}\left(\nabla^G_\beta \sigma\right) + k_c^t\left[\grave{A}_\beta\right]\right\}$$

$$= 0$$

Currents

Currents for the fields

$\phi_G = \sum_{\beta=0}^{3} \left[\mathcal{F}_G^{\alpha\beta}, G_\beta\right]_{T_1 Spin(3,1)} \otimes \partial\xi_\alpha \in T_1 Spin\left(3,1\right) \otimes TM$

$\phi_A = \sum_{\beta=0}^{3} \left[\mathcal{F}_A^{\alpha\beta}, \grave{A}_\beta\right]_{T_1 U} \otimes \partial\xi_\alpha \in T_1 U \otimes TM$

$[\phi_G]_{a=1...6}^{\beta=0..3} : \phi_G = \sum_{\beta=0}^{3} [\phi_G]_a^\beta \overrightarrow{\kappa}_a \otimes \partial\xi_\beta$

$[\phi_A]_{a=1...m}^{\beta=0..3} : \phi_A = \sum_{\beta=0}^{3} [\phi_A]_a^\beta \overrightarrow{\theta}_a \otimes \partial\xi_\beta$

$[\phi_A] = \sum_{b,c=1}^{m} C_{bc}^a [g]^{-1} \left[\mathcal{F}_A^b\right] [g]^{-1} \left[\grave{A}^c\right]^t$

$[\phi_G] = \sum_{a,b,c=1}^{6} \epsilon\left(a,b,c\right) [g]^{-1} \left[\mathcal{F}_G^b\right] [g]^{-1} [G^c]^t$

Current for the particles

$J_G = -\frac{C_I}{16 C_G} \mu M \mathbf{Ad}_\sigma \upsilon\left(k_0, 0\right) \otimes V \in T_1 Spin\left(3,1\right) \otimes TM$

$J_A = -\frac{C_I}{16 C_A} \mu M \sum_{a=1}^{m} k_c^a \overrightarrow{\theta}_a \otimes V \in T_1 U \otimes TM$

$J_{EM} = \frac{C_I}{8 C_{EM}} \mu q M V \in TM$

$[J_G]_{a=1..6}^{\beta=0..3} : J_G = \sum_{a=1}^{6} \sum_{\beta=0}^{3.} [J_G]_a^\beta \partial\xi^\beta \otimes \overrightarrow{\kappa}_a$

$[J_A]_{a=1..m}^{\beta=0..3} : J_A = \sum_{a=1}^{6} \sum_{\beta=0}^{3.} [J_A]_a^\beta \partial\xi^\beta \otimes \overrightarrow{\theta}_a$

$J_G = \frac{C_I}{16 C_G} M \upsilon\left(-\left[A\left(w\right)\right]\left[C\left(r\right)\right] k_0, \left[B\left(w\right)\right]\left[C\left(r\right)\right] k_0\right) \otimes V$

$C_I \frac{1}{i} \frac{1}{M} \left\langle \psi, \nabla_V \psi \right\rangle = C_I \frac{dK}{dt} + 8\left(4 C_G \mathbf{G}\left(J_G\right) + C_A \grave{\mathbf{A}}\left(J_A\right) + C_{EM} \grave{\mathbf{A}}_{EM}\left(J_{EM}\right)\right)$

Fields equations

$\phi_{EM} = 0; \Delta \grave{A}_{EM} = -\delta \mathcal{F}_{EM} = 2 J_{EM}^*$

$J_A = \phi_A; d\left(*\mathcal{F}_A\right) = 0$

$J_G = \phi_G; d\left(*\mathcal{F}_G\right) = 0$

Codifferential equations

$\sum_{\beta=1}^{3} \partial_\beta \left[*\mathcal{F}_G^r\right]_\beta = 0$

$\partial_0 \left[*\mathcal{F}_G^r\right] = \sum_{\beta=1}^{3} \left(\partial_\beta \left[*\mathcal{F}_G^w\right]\right) \left[j\left(\varepsilon_\beta\right)\right]$

$*\mathcal{F}\left(m\right) = d\phi^*\left(m\right)$

Current equations

$[k_G]_{1\times 3} = -\frac{C_I}{16 C_G} M \left\{\mathbf{Ad}_\sigma \upsilon\left(k_0, 0\right)\right\}$

$[J_G] = [V][k_G]$

$[*\mathcal{F}^r][G]^t - \left([*\mathcal{F}^r][G]^t\right)^t = c\left[j\left(k_G\right)\right]\left(\det P'\right)$

$[*\mathcal{F}^w] + 2[*\mathcal{F}^r] j\left([G]^{-1} G_0\right) = \frac{\det P'}{\det G}\left\{2 j\left(k_G\right) j\left(c G_0 + [G][v]\right) + [k_G]\left(c[G_0] + [G][v]\right)\right\}[G]$

$k_E = \mu \frac{C_I}{4 C_{EM}} q M$

$\gamma = 0..3 : \sum_{\beta=0}^{3} \{g_{\gamma\gamma} \partial_\beta \left([*\mathcal{F}_{EM}]_\gamma^\beta\right)\} \det P = -k_E \left\{[V]^t [g]\right\}_\gamma$

Energy

$T = -\frac{1}{2} C_I \mu M V \otimes k_0^t \operatorname{Re} D\left(-Z\right) dZ + 4 \sum_{\alpha\beta\gamma} \left(4 C_G \left\langle \mathcal{F}_G^{\alpha\gamma}, \partial_\beta G_\gamma \right\rangle_{Cl} + C_A \left\langle \mathcal{F}_A^{\alpha\gamma}, \partial_\beta \grave{A}_\gamma \right\rangle_{T_1 U}\right) \partial\xi_\alpha \otimes d\xi^\beta$

On shell :

$\langle \mathcal{F}_G, \mathcal{F}_G \rangle_G = -\mathbf{G}(\phi_G)$
$\langle \mathcal{F}_A, \mathcal{F}_A \rangle_U = -\grave{\mathbf{A}}(\phi_A)$
$\langle \mathcal{F}_{EM}, \mathcal{F}_{EM} \rangle_{EM} = \grave{\mathbf{A}}_{EM}(\delta \mathcal{F}_{EM})$
$L_{System} = L_{Fields} = -8C_G \mathbf{G}(J_G) - 2C_A \grave{\mathbf{A}}(J_A) - 2C_{EM} \grave{\mathbf{A}}_{EM}(J_{EM}) = -\frac{1}{4} C_I \frac{dK}{dt}$

Tetrad equation :

$\forall \alpha, \beta = 0...3 : C_I \frac{1}{i} \frac{1}{M} V^\alpha \langle \psi, \nabla_\beta \psi \rangle = 4 [X]^\alpha_\beta - 2\delta^\alpha_\beta Tr[X]$
$[X]^\alpha_\beta = -\sum_{\gamma=0}^{3} \{4C_G \langle \mathcal{F}_G^{\alpha\gamma}, \mathcal{F}_{G\beta\gamma} \rangle_{Cl} + C_A \langle \mathcal{F}_A^{\alpha\gamma}, \mathcal{F}_{A\beta\gamma} \rangle_{T_1 U}\}$
$[Y_{RR}] = C_G \operatorname{Re} [\mathcal{F}_G^r]^t [\mathcal{F}_G^r] + C_A [\mathcal{F}_A^r]^t [\mathcal{F}_A^r]$
$[Y_{WW}] = C_G \operatorname{Re}\left([\mathcal{F}_G^w]^t [\mathcal{F}_G^w]\right) + C_A [\mathcal{F}_A^w]^t [\mathcal{F}_A^w]$
$[Y_{RW}] = C_G \operatorname{Re} [\mathcal{F}_G^r]^t [\mathcal{F}_G^w] + C_A [\mathcal{F}_A^r]^t [\mathcal{F}_A^w]$
$Tr\{[Y_{RR}][g_3]\} = Tr\left([Y_{WW}][g_3]^{-1}\right) = 0$
$[g_3][Y_{RW}] = [Y_{RW}]^t [g_3]$
$[Y_{WW}] = -[g_3][Y_{RR}][g_3] \det [g_3]^{-1}$