Common structures in scientific theories

Jean Claude Dutailly

Common structures in scientific theories

Jean Claude Dutailly
Paris

10 June 2014

Abstract

The existence of common structures in physics and computer sciences theories has become an active field of research. In the present paper the scope is enlarged to any theory which is formalized in quantitative models meeting some precise but general conditions. The first section is dedicated to characterize the main components of any scientific theory, using the concepts of epistemology, and puts in prominent positions models, association of objects with properties which are represented by quantitative variables. In a second section we show that a large class of models can be represented in Hilbert spaces. In this framework we precise the concepts of observables, with its relation with measures, and develop several tools to characterize the evolution of systems, interacting systems and phases transitions.

Science has acquired a unique status in our societies. On one hand 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 ? On the other hand many of its assertions are controversial, when they impact our daily lifes (from the climate warming to almost any drug), but not least in the scientific commmunity itself. The latter is natural and even sound - controverse 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 ? A less travelled path is to revisit the concepts and principles which are at the fundation of our scientific theories. In order to be efficient we must be bold, and do not hesitate to dig deep. And without surprise we reach the fringes of philosophy. After all, for thousands of years philosophers have been the architects of knowledge. Unfortunately today they seem to tend more to be of the regulator kind, without much input in the matter.

Actually this kind of shake down happened to Mathematics a century ago. Mathematicians were very proud of their numerous theories in Algebra, Geom-

etry, Analysis, and hoped to bundle them together in a unified, fully logical, theory. The project lead to some surprises, such as theorems which are true but cannot be proven (Gödel), and to the development of mathematical logic and a set of new tools which enlarges the scope of what was so far seen as rational computation. These tools were soon linked to computational logic, with Türing, Church, Kleene and the $\lambda-$calculus, which found naturally many applications in Computer Sciences. Meanwhile, considering that many structures used in different fields have similar features, mathematicians (Eilenberg, Mac Lane) developped around 1945 the Category Theory, and it happens that it can be used as a common framework to represent logic and computational sciences. So, over all, the initial goal had not been achieved, the status of Mathematics stays ambiguous, but with a better understanding of the issues and more powerful tools to deal with them. Moreover some of the concepts developped in the Quantum Theory of Fields (Feynman diagrams), Quantum Gravity (strings and branes theories), and Quantum Computing (Heyting algebras) appear also to resort to Category Theory.

The purpose of this paper is to contribute to this research : is there a common background to scientific theories ? And if it is so, could it be used for a better understanding of pending scientific issues ? It can be done with Category Theory, and this is an active field of research (see J.Baez and Coecke for more) with a strong streak of Quantum Mechanics (QM). However we will not use this path. QM itself is at the core of many issues and it does not seem appropriate to give a special role to its concepts. Moreover it happens that the enquiry can be easily extended outside theoretical Physics, and addresses some classic but crucial problems in engineering and Economics, such as phase transitions. And indeed including Economics and even other social sciences gives a fresh vision of what we mean by science. Category Theory is powerful, but it is one of the most abstract branch of Mathematics, not necessarily known to many readers in these fields. Using more classic methods, under a large but more restricted set of conditions, it is possible to get rigorous mathematical proofs of key theorems, which have practical consequences.

To understand the issues we have to do with some Philosophy. The philosophers of the present day are not very helpful, but their predecessors have left a legacy which can be put to fruition. So the first section is dedicated to a reminder of what philosophy has to say about science, aimed at scientists, physicists, biologists, economists or any other who is more versed to the vocabulary and ideas of their field than to the controverses of philosophers. The purpose is here to set the ground, to show what are the specificities of scientific knowledge, how scientific theories are built and improved, the relation between experimentation and intuition. Without claiming any appartenance to a philosophical school let us say that the ideas which are expressed are in line with what is called conventionalism, and its most famous representant Henri Poincaré[1].

[1]Even if Henri Poincaré would have strongly reproved the use of infinite dimensional vector spaces.

From this introduction, models appear as a key component of any scientific theory, at least of those which use a formalism. Economists are familiar with 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, but this process is not neutral. It happens that in most of the models the mathematical formalism has important consequences, in that it can be shaped around Hilbert spaces structures : the state of systems can be represented by vectors of Hilbert spaces, it is measured through specifications of observables, related to operators on the Hilbert space, and its values are eigen vectors of this operator. The formalism helps to find the laws of the evolution of systems, notably in finding the rules in the transitions between phases. It can be extended to interacting systems, and shows that homogeneous systems have necessarily a finite number of phases. These features are precised in mathematical theorems, whose proofs have been given in a previous paper. So in the second section the focus is on the meaning and the practical consequences of the results.

1 WHAT IS SCIENCE?

What is knowledge ? How does it progress? The branch of philosophy which deals with these topics is epistemology. 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 precised the role of formalism in scientific theory, and Popper introduced, with the concept of falsifiability, a key element in the relation between experiment and formal theories. But since the middle of the XX° century epistemology seems to have drifted away from science, and philosophers tend to think that actually, philosophy and science have little to share. This feeling is shared by many scientists (Stephen Weinberg in "Dreams of a Final Theory"). This is a pity as modern sciences need more than ever a demanding investigation of their foundations.

Without pretending to create a new epistemology, and using all the basic work done by philosophers, I will try to draw a shematic view of epistemology, using a format and words which may be more familiar to the scientific reader.

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. In Science measures follow precise protocols, so it is generally assumed that a measure is justified.

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 can be transmitted and translated 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 fields addressing similar topics.

1.1 Assertions

1.1.1 Circumstantial assertion

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

two tags :

- the first tag identifies the specific case (the circumstances) about which the assertion is made
- the second tag identifies the author of the assertion

and two main components :

- the first component defines the subject : this is a crucial part, and it is expressed with concepts, which are themselves generally agreed upon in the community
- the second component is the value of the assertion. One has a **quantitative assertion** if the value is the measure of a phenomenon, expressed by a figure or a code, taken according to generally agreed procedures. One has a **logical assertion** if it can take one of two of the logical values : true or false.

The assertion can be **justified** - its author claims that he is certain of the value of his assertion - or not - the author has a doubt about his assertion but considers it is plausible. Usually, at least in sciences, a quantitative assertion which is supported by a measure is deemed justified, because the measure follows agreed procedures. A non justified assertion is a **hypothesis**.

Examples of circumstantial individual assertions :

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

Knowledge, and specially scientific knowledge, is more than individual circumstantial assertions : it is a method to build narratives from assertions. It

proceeds by enlargement, by going from individuals to a community, from circumstantial to universal, and by linking together assertions.

A similar assertion made by different people with respect to the same case can be shared : then it becomes the assertion of the community (it is tagged by the community). It is justified if all agree it is so, then it is an evidence. This is an essential criterium in the reliability of any measure. If the assertion is not justified this is a hypothesis, a common belief.

Evidences and common beliefs can be linked together. It is convenient to represent the links by the logical operators $\wedge$ (and), $\vee$ (or), $\neg$ (not). Such a string of assertions denotes the simultaneous occurence of each assertion, shared by the community. It is justified if each assertion is justified.

From a string S of assertions it is sometimes possible to deduce another assertion B : one says that S is an **explanation** of B and writes : $S \Rightarrow B$

Let us consider some examples :

"the volume of the vase is 1000 cm^3"$\wedge$"the liquid is pure water"$\wedge$"the density of pure water is 1"$\Rightarrow$"the mass of the liquid is 1kg"

The first three justified assertions give an explanation of the fourth assertion. It can be checked but, until this is done, it is still an hypothesis, a value which is plausible because it rests upon another argument (the definition of density, by a law).

"dinosaurs have disapeared around 65 m years BE"$\wedge$"a big meteorite has impacted Earth around 65 m years BE"$\Rightarrow$"the extinction of dinosaurs has been caused by the impact of a meteorite"

The first two assertions are justified, and they lead to the third assertion, which is a plausible explanation, as it cannot be checked.

"Alice had an incurable cancer"$\wedge$"Alice has prayed for her cure"$\Rightarrow$"Alice has been cured"

We have three evidences (they can be justified), we have a plausible explanation, but not everybody would agree that it is a justified explanation.

Mystery books are full of such narratives : at the end the detective comes with a set of justified assertions to unveil the assassin. It can or cannot be justified in court (notably if the accused admits that he is the assassin), but usually it stays a plausible explanation.

Circumstantial assertions can lead at best to plausible explanations, that is to plausible assertions, which can or cannot be checked. Moreover even a justified assertion does not imply that, in another occurence, a similar assertion would be justified.

There is no general rule to build an explanation : cirumstantial assertions do not provide a general method of inference. Usually they are built from the generalization of similar occurences, or by the use of more general laws. Sometimes they are a guess, which can be very fruitful if it leads to laws.

To go further one needs a feature which is called necessitation by philosophers, and is cleary represented by the symbol $\Rightarrow$. And this requiers to go from the circumstantial to the universal. In the following one assumes that the first step (from individual to shared) has been made.

1.1.2 Universal assertions

There are several kinds of **universal assertions**.

The first do not refer to any special occurence : "all physical objects are in a single container, the universe", "the universe is a 4 dimensional manifold", "God exists", ...They are logical assertions, they cannot be justified or invalidated directly because they do not refer to any circumstance. We call them **fundamental hypotheses**.

The second is used to define generic classes of objects, which are characterized by their properties : they are **definitions**.

They can be logical assertions such as :

"all insects have three pairs of legs", "material bodies travel along a world line in the 4 dimensional universe", "democratic states have an elected government",...

When applied to a specific occurence of the objects the assertion is true or false. There can be several similar assertions attached to the same object : one can require more from a democratic state than to have an elected government, but anyway this assertion must be true for any state to be deemed democratic.

They can be used to attach a quantity, that we call a **variable**, to a generic class of objects :

"for any gas there is a temperature T", "the gene G of a chromosome", "all fermions have a non null mass M", " a good in a market has a price p",...

The assertion tells that for any occurence of the generic class of objects it is possible to assign a value to the variable, through a measure done according to a precise procedure. The variable is represented as a mathematical object with a precise definition : it can be a code, chosen in a table, a scalar, a vector,...

The assertions of this second kind 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 occurence. So they are justifiable assertions.

They have a restricted predictive power : if an occurence of the object in a class has been identified then the assertion tells that there is some variable which can be measured. 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).

Assertions of the third kind have a fixed value (if they are logical) for any occurence, or at least one occurence, or provide a relation between variables (which assumes that they have been defined elsewhere) :

"for any gas in a vase the temperature T, the volume V and the pressure P are related by PV=kT for some constant k", "the speed of light is constant" (whenever it can be measured), "any dominant allele is transmitted to the descendance", "there is a gene which encodes the protein", "all sinners go to hell",...

These assertions enable us to introduce probability : because the circumstances can be repeated one can estimate probability laws for the value of the circumstantial assertions. The Mendel's heredity laws provide a simple example of this kind.

1.2 Scientific laws

Universal assertions can be linked as above. For quantitative assertions it is assumed that the measure can be defined and take any value (usually in some large domain), then the linkage provides a template such as :

S = "for any gas there is a temperature T" $\wedge$ "for any gas in a vase there is a volume V" $\wedge$ "for any gas there is a pressure P"

S'= "for any dominant allele A" $\wedge$ "A is present in the male chromosome" $\vee$ "A is present in the female chromome"

Then assertions of the third kind can be represented as consequence of a string of assertions of the first or second kinds :

$S \Rightarrow$"there is a relation PV = kT"

$S' \Rightarrow$"A is present in the chromosome of the descendance"

The meaning of $\Rightarrow$ is now clear : whatever the occurence, the value of the outcome (the right side) is defined by the value of the string (the left side). Because of the universality of these assertions, one can consider **predictions** : given a universal assertion, it is conceivable to predict its outcome in any given occurence.

This outcome can or cannot be checked. If it cannot be checked, then it is a hypothesis, universal as well as circumstantial. If the outcome can be checked in any circumstance it is still a universal hypothesis, but it suffices of a single circumstance to invalidate the assertion : it is **falsifiable**. So they are hypotheses, but with a strong plausibility.

This is this falsifiability which is often seen (notably by Popper) as a criterium of scientificity. We go from plausible explanation to **scientific laws**.

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 following explanation :

"any believer who is ill" $\wedge$ "any believer who prays to be cured" $\wedge$ "God wills" $\Rightarrow$ "the believer is cured"

For any occurence, 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 occurence, so that one has the assertion of the third kind :

"any believer who is ill" $\wedge$ "any believer who prays to be cured" $\wedge$ "if God wills" $\wedge$ "the believer is cured"

which could be justified. 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 :

"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.

The border between assertions of the second and third kinds can be slim, but is significant. The assertion "material bodies travel along a world line in the 4 dimensional universe" can be seen as a law (which could possibly be invalidated) but then we need to define explicitely what is a material body (as opposed notably to force fields).

1.2.1 Is there a scientific method ?

It is commonly believed that one distinctive feature of the scientic 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 scientifc 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 criterium 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 fortuituous : what distinguishes a scientific law from a justified assertion is its universality, notably that its predictions can be checked for any occurence.

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.

1.3 Theories

Scientific laws are an improvement over circumstantial explanations, because they have the character of necessity. 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. This is done by organizing laws related to a common field in **theories**. Each theory is a logically coherent set of generic objects, fundamental hypotheses and basic scientific laws, and a rule of inference which enables to generate scientific laws which can be implemented for any specific occurence related to the field.

So the first requisite for a theory is to define a collection of **objects**, characterized by their properties, such that they can be identified in any situation which is in the scope of the field.

For instance : in Classical Mechanics we have solid bodies, forces,... in Chemistry we have compounds, elements, energy,... in Economics we have market, added value, capital,... in Political Sciences we have judicial systems, executive power, ...

To these objects are associated variables which provide the link between the measures and the properties, and the necessary tools for the predictions and checking the laws.

For any specific situation which is in the scope of the field then we can identify a set of objects. A case represented by a generic set of objects is a **system** in the theory**.** To each object can be attached variables, which can be or not of interest for the study, and a system with a given collection of variables related to objects of the system is a **model**.

For instance : the model of a star system in classical mechanics with the position of planets, a mix of compounds in a given state in Chemistry with its composition, a market with the sales of different products in Economics, a state with its organization of powers in Political Sciences,...

Because of their universality, assertions of the second kind can be implemented for any model of the theory : they guarantee that measures can be done, with defined protocols, in any specific occurence of the model, for the objects which are identified in the system. When they are considered together in a system, it is assumed that the objects will behave in a specific way, and this behavior entails that the variables will take some specific values. And indeed such a behavior is just the realization of an occurence for a universal assertion, the application of a scientific law. So it is common to illustrate scientific laws in a model. But we need to distinguish the first step, that is the definition of the model (a system comprised of objects and their associated variables), from the second (specific values resulting from a law). The falsifiability test requires that one can consider all possible values of the variables before checking the prediction, so the set of possible values has a clear logical as physical meaning.

The second requisite for a theory is a set of fundamental hypotheses and basic scientific laws. Actually the latter can be expressed as assertions deemed true for a large set of models. Exemple : the Newton's gravitation laws are expressed as the forces between bodies with their mass and distances in a general system.

The third requisite for a theory is a **rule of inference**. Starting from fundamental hypotheses and basic laws, we shall be able to determine scientific laws for any model, which therefore will give predictions for any specific occurence of the model. And by consequence can provide the falsifiability check. The rule of inference requires a formalism, and depends of this formalism.

Whenever the basic laws are logical the formalism is that of the logic of predicates, and its usual rules of inference. A predicate related to a specific

model is proven true by using the fundamental hypotheses, basic laws which are expressed as axioms, and the rule of inference.

When it is possible to use a mathematical formalism the laws take the form of equations, and the mathematical calculus is used to prove more specific laws for each model. For instance the laws for the trajectories of planets in a star system are deduced from the gravitational laws, the laws of Mechanics, and the laws of Galilean Geometry. Each occurence of states of the system corresponds to a given collection of values to the variables, they are usually seen as the initial conditions and the prediction is then related to the following evolution of the system.

But there are other formalisms. The most illuminating example is the atomic representation used in chemistry. A set of symbols such as :

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

tells us almost everything which is useful to understand and work with most of chemical experiments.

So, to sum up, a scientific theory is comprised of :

- a collection of objects, related to the field encompassed by the theory;
- to each object is attached a collection of properties, which can be logical, or represented by variables,
- a collection of fundamental hypotheses and basic scientific laws
- a formalism which enables to deduce by inference other scientific laws which are valid for any model of systems

The universality and the falsifiability of the laws provide the scientific credence to the theory.

Most of these ideas have been formalized in the Category language by logicians, a theory is a category and a model is a functor, and we could go further in this path, but we will not need it here.

1.4 Which theory ?

For a given field there can be many different theories. To be deemed scientific their laws must pass the falsifiability test, but, as we have seen, until proven false, an assertion stays a plausible hypothesis. 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 ?

1.4.1 Simplicity

The first criterium 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 is a combination of 4 bases. The electric and magnetic fields have been unified by the Maxwell's laws, and the unifications of all force fields including gravitation is the Graal of physicists. Similarly there should be as few fundamental hypotheses and basic laws as possible. The Galilean system was not more accurate or legitimate (motion is relative, so the assertions that Earth circles the Sun or that Sun moves around the Earth are both valid) than the Ptolemaic system, but it provided a general theory to compute the trajectories of bodies around a star and paved the way to the Newton's gravitation law.

1.4.2 Positivism

The second criterium 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 developped 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, since the development of mathematical logic, we know that in any system of predicates (that is axioms) which is powerful enough to be efficient (account for arithmetics) there are theorems which cannot be proven. The inference rules by themselves do not suffice to tell which are the "right axioms" which should be added to make an efficient theory of mathematics. These basic axioms have been actually put forward even before the advent of mathematical logic : they are the product of many enquiries in Algebra, Geometry, Analysis,... and of the intuition of generations of mathematicians. 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. So Mathematics, as we know it, can be seen as a scientific theory by itself.

In natural sciences it is a sound requirement that there is a strong, unified background, explaining and reflecting the unity of the physical world. But in the different fields theories usually do not proceed from the most elementary laws. The atomic representation used in Chemistry precedes quantum field theories of particles. Biology acknowledges the role of chemical reactions, but its basic concepts are not embedded in chemistry. We do not have in Physics a theory which would be general and powerful enough to account for everything. And

anyway in most practical cases specific theories suffice. They use a larger set of assumptions, which are simplified cases of general laws (Galilean Geometry replacing Relativist Geometry, Newton's laws substituted to General Relativity) or phenomenological laws based on experimental data. In doing this the main motivation of scientists is efficiency : they do not claim the independance of their fields, but acknowledge the necessity of simpler theories for their work. However one cannot ignore that this move from one level to the other may cover a part of mystery. We still do not understand what is life. We do not have any determinist model of irreversible elementary process (such as the desintegration of a particle), which are essential in thermodynamics.

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 [2]. 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 occurences, past or related to a few number of cases. The huge number of factors involved, most of which cannot be controlled, weakens any prediction, 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 who claim to have knowledge, and it is a sound democratic principle that they should be kept accountable to the people who fund their work.

1.4.3 Conservative pragmatism

The third criterium 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.

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

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 occurences 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 criterium - 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 if to provide answers, not to explain a mystery by a riddle. And when the new enigma requires itself more powerful tools the race may turn into a justification in itself.

1.5 Truth and reality

In this picture what are the relations of science with truth and reality ?

A justified assertion can be accepted as truth in a Court of justice. But not that 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 occurences. 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] .

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 occurences of Human civilizations, with common needs and constraints. However this does not mean that we know what is reality : what we can achieve is the most accurate and plausible representation of reality, but it will stay temporary, subject to revision, and adjusted to the capability of our minds. Because this representation is made through a formalization, the language which is used acquires a special importance. This is a matter of much controverses but it is clear that major steps in the theories would have been impossible without

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

prior progresses in the formalism which is used : mechanics with differential and integral calculus, general relativity with differential geometry, QTF with group representations, Economics with accounting. The use of more powerful mathematical tools, and similarly of computational techniques, increases our capacity to check predictions, but also to build the theories.

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 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 ? As we have seen above, and clearly in the case of mathematics, 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 occurences. 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 create 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.

1.6 Fundamental principles

The previous review leads to state some general principles which have far reaching consequences.

The first principle is the Principle of relativity, which can be stated as "Scientific laws do not depend on the observer". This is a logical consequence of the universality of scientific laws : it should be checked for any occurence, as long as the proper protocols are followed, whatever the people who do the experiment (the observers), whenever and wherever they are located. So it is necessary that these protocols specify what are the parameters or variables which must checked, and their format. 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, 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.

Universality of scientific laws implies that experiments are reproducible, time after time, which requires either that the circumstances stay the same, or can be reproduced identically. This can be achieved only to some degree, controlled by checking all the parameters which could influence the results. It is assumed that the parameters which are not directly involved in the law which is tested are not significant, or keep a steady value, in time as well as in the domains which are exterior to the area which is studied. Universality implies some continuity of the phenomena, in the physical as in the social world. This is obvious for phenomenological laws, which are infered from repetition of occurences. Even probabilist laws are seen as the imperfect representation of more complex processes. Measurable changes, notably discontinuous ones, happen because, at a deeper level, processes have been building their infinitesimal effects. An earthquake is the result of the slow motion of tectonic plates. Thus continuity is a strong argument in favor of the unity of the fields which are accessible to scientific knowledge. If a theory can accept fundamental hypotheses, science cannot accept, at least in its broadest vision, uncontrolled factors, meaning variables whose value cannot be predicted or measured prior to an experiment. One can believe in miracle cures, but Science cannot rely on God's will : at some level any variable which appears in its laws must account for adjustable, or predictable, measurable data. And the most fundamental scientific laws should be continuous and determinist.

But the organization of scientific knowledge leads to more significant results, which are the topic of the next section.

2 HILBERT SPACES AND REPRESENTATIONS OF SYSTEMS

These results actually are about general, quantitative models, that is systems (generic collections of objects defined in the theory) and a set of variables related to the objects of the system, expressed as quantities with a precise mathematical definition. They do not involve the laws of the theory, that is the relations which can be assumed or deduced between the variables, they stay at a prior step, which is the specification of the model. But, as we will see, for a very large class of models, whatever the field, the mathematical formalism of the theory entails by itself a new, powerful, general mathematical formalism for the model. Because these results are mathematically proven, they are valid for any theory which uses mathematics as formalism and rule of inference, without any reference to the objects, fundamental hypotheses, or laws considered in the

theory. So that most of the models used in a scientific theory can be represented in a common framework, with important practical consequences.

The mathematical proofs have been given in a previous paper ("Hilbert spaces in modelling of systems"), to which the reader is refered for all the technical details, and I will focus here on the meaning and consequences of the results.

2.1 Hilbert spaces

Let be a system represented by a model, with a fixed finite number N of variables $(X_k)_{k=1}^N$ then the **state** of the system can be defined by the value of these variables, that we will denote collectively X. The first and fundamental result is the following :

Theorem 1 *For any system represented by a model, with a fixed finite number N of variables* $(X_k)_{k=1}^N$ *such that :*

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

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

iii) 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 :* $X'_k = f_k(X_k)$

Then :

there is a separable, infinite dimensional, Hilbert space H, defined up to isomorphism, such that $S = O_1 \times ... \times O_N$ *can be embedded as an open subset* $\Omega \subset H$ *which contains 0 and a convex subset.*

each state X is associated to a vector ψ *of H and there is a linear isometry :* $\Upsilon : S \to \Omega :: \psi = \Upsilon(X)$

The main condition is that the variables belong to a vector space. So it excludes models where some variables are qualitative : they take discreet values chosen in a table. We will see later how to deal with them. But the value of the variables can be restricted to some open domain of V_k (for instance one can require that they are positive scalars). At least one of the vector space V_k, and then their product $V = V_1 \times ... \times V_k$, must be infinite dimensional. It means that what is studied is not one occurence of the system (for instance the state at a given time) but the behavior of the state over an infinite set of occurences (its evolution over a period). Take a body whose trajectory is followed in a fixed frame by a vector X(t), then the state of the system is the map X and not the value X(t) at some given time t.

The vector spaces V_k must be separable real Fréchet spaces. A Fréchet space is a common structure (more general than a complete normed vector space), but it must be separable, which is a bit more restrictive. The condition iii) allows the

possibility to define the variables by different, related, quantities, over distinct ranges or using other units.

Let us give some usual examples where the conditions are met :

i) the variables X_k are maps : $X_k : R \to E$ where R is an open subset of $\mathbb{R}$ and E a finite dimensional vector space, and $\int_R \|X_k(t)\|^2 dt < \infty$ with any norm $\|\|$on E, then V_k is itself a separable Hilbert space and H can be identified with V and ψ with X. This addresses :

- all models in analytical Mechanics
- most of the models in fluid mechanics (the variable X can depend on other parameters than t)
- seismic studies (the variables X_k are the waves)
- almost all the models in Economics, representing the evolution of a system : values of bonds, currencies or equities, macroeconomic models, markets studies such as the consumption $X_k(r)$ of products k = 1...N with respect to the income r of a household....

ii) X_k are scalar continuous functions on a compact domain of a topological space

iii) X_k are complex p integrable maps on $\mathbb{R}^n$: $V_k = L^p(\mathbb{R}^n, dx, \mathbb{C})$ with $1 \leq p < \infty$, which addresses most of the models in electromagnetism

A scientific law will usually be expressed by some relation between the variables X_k, but such relations are not involved here : the variables are assumed to be independant. Similarly a variable X_k and its derivative $\frac{dX_k}{dt}$ are considered as independant variables. So we are here a step before the expression of laws. A law in the model implies that the values of X which will be observed belong to some subset of V, and the vector ψ belongs to some subset of H.

So we have the striking result that most models, whatever their field, their objects and hypotheses, have a common mathematical representation. It seems abstract, but this representation provides powerful tools.

Each variable X_k is itself associated to a Hilbert space H_k and $H = H_1 \times ... \times H_N$

On the real Hilbert space H there is a scalar products denoted $\langle\rangle$, which is a bilinear symmetric form and Hilbertian bases $(\widetilde{\varepsilon}_n)_{n\in\mathbb{N}}$ such that :

$\langle\widetilde{\varepsilon}_n, \widetilde{\varepsilon}_m\rangle = \delta_{mn}, \psi = \sum_{n\in\mathbb{N}} \psi_n \widetilde{\varepsilon}_n, \|\psi\|_F^2 = \langle\psi,\psi\rangle = \sum_{n\in\mathbb{N}} \psi_n^2$

The norm $\|\|_F$ provides the distance between two states : $\|\psi - \psi'\|_F$

The vector space V has also bases $(e_n)_{n\in\mathbb{N}}$, not necessarily orthonormal, such that : $\forall X \in V : X = \sum_{n\in J} X_n e_n$ where only a finite number of components $n \in J \subset \mathbb{N}$ is non null. But the basis itself has a countably infinite number of vectors.

For any basis $(e_n)_{n\in\mathbb{N}}$ of V there are unique families $(\varepsilon_n)_{n\in\mathbb{N}}, (\phi_n)_{n\in\mathbb{N}}$ of independant vectors of H (which depend on $(e_n)_{n\in\mathbb{N}}$ and are not necessarily Hilbertian) such that :

$\forall m, n \in \mathbb{N} : \langle\phi_n, \varepsilon_m\rangle_H = \delta_{nm}$

and the linear bijective map Υ is such that X and $\Upsilon(X)$ have the same components respectively in $(e_n)_{n\in\mathbb{N}}, (\varepsilon_n)_{n\in\mathbb{N}}$:

$\forall n \in \mathbb{N} : \varepsilon_n = \Upsilon(e_n)$
$\forall X \in O : X = \sum_{n\in\mathbb{N}} \langle \phi_n, \Upsilon(X)\rangle e_n \to \Upsilon(X) = \sum_{n\in\mathbb{N}} \langle \phi_n, \Upsilon(X)\rangle \varepsilon_n \in \Omega$

Models may include variables which take discreet values, defined in a table. When these variables are used to distinguish different behaviors for a given system, then it can be assumed that the variable X takes significantly different values. And the simplest way to represent this fact is to assume that the subset O of possible values of X is disconnected (in the topological meaning : it is the union of disjoint open subsets). For two connected components O_1, O_2 it is easy to prove that there is a continuous function $f : V \to [0,1]$ such that $f(\Upsilon(X)) = 1$ in O_1 and $f(\Upsilon(X)) = 0$ in O_2. The function f can be added to the collection of variables and the theorem 1 holds.

2.2 Observables

The vector ψ, as well as X, is defined in a basis with an infinite number of vectors. But on a given system one can take only a finite number of measures. So usually one can have only an estimate of X. Notice that this is not a matter of the precision of the measures, but of the **specification** of X as a map. For instance let us consider the trajectory X(t) of a bullet shot by a gun. The simplest specification is a straight line, which can be refined to a parabola. A marksman would consider wind, humidity,.. and use tables, based on tests. These tables are computed by estimating the function X from experimental data, and using more complicated types of functions X(t). It is clear that the quality of the estimate increases with the number of available data, but we will not focus here on the statistical problem, but on the step which comes before : the choice of a specification Y. Because X is a vector in the conditions of theorem 1, one can represent this process of specification by a map $\Phi : V \to V :: Y = \Phi(X)$ that we will call an **observable**, with the natural properties :

Definition 2 *i) an observable is a linear map :* $\Phi \in L(V;V)$ *(not necessarily continuous)*

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

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

The estimation of the state X of the system, for a specific occurence, will be done by a statistical method on the specified map $Y = \Phi(X)$. Because $\Phi(X)$ belongs to a finite dimensional vector subspace this estimation can be done with a finite number of data. The choice of Φ is free, under these constraints, and is up to the scientist who studies the system.

The most natural choice of observable would be to take a basis $(e_n)_{n\in\mathbb{N}}$ of V, select a finite number J of its vectors, and define the observable Y_J by the map

: $Y_J\left(\sum_{n\in\mathbb{N}} X_n e_n\right) = \sum_{n\in J} X_n e_n$. One keeps only the components with respect to the vectors $(e_n)_{n\in J}$ .. We will call such maps a **primary observable**. Y_J depends both on the choice of a basis and a finite set J.

To any observable one can associate a linear map on H by : $\widehat{\Phi} : H \to H$ :: $\widehat{\Phi} = \Upsilon \circ \Phi \circ \Upsilon^{-1}$

If $\widehat{\Phi}$ is a normal operator, that is if : $\widehat{\Phi} \circ \widehat{\Phi}^* = \widehat{\Phi}^* \circ \widehat{\Phi}$ where $\widehat{\Phi}^*$ is the adjoint of $\widehat{\Phi}$, then we say that Φ is a **normal observable**. Any primary observable is normal.

We have the following result :

Theorem 3 *For any normal observable Φ of a system meeting the conditions of the theorem 1 :*

i) Φ is a compact, continuous map $\Phi \in \mathcal{L}(V;V)$, and a finite linear combination of primary observables : $\Phi = \sum_{p=1}^{m} \lambda_p Y_{J_p}$ where $(J_p)_{p=1}^{m}$ are disjoint finite subsets of $\mathbb{N}$ for some basis $(e_n)_{n\in\mathbb{N}}$ of V

ii) $\widehat{\Phi}$ is a compact, self-adjoint, Hilbert-Schmidt and trace class operator

iii) if the system is in the state $X = \sum_{n\in\mathbb{N}} \langle \phi_n, \Upsilon(X)\rangle_H e_n$ the value of the observable is : $\Phi(X) = \sum_{n\in\mathbb{N}} \left\langle \phi_n, \widehat{\Phi}(\Upsilon(X))\right\rangle_H e_n$

iv) there is always a primary observable which is at least as efficient as Φ (the variance due to the specification is smaller).

The operator $\widehat{Y}_J$ associated to a primary observable is the orthogonal projection on the vector subspace spanned by the vectors $(\varepsilon_n)_{n\in J}$ associated to $(e_n)_{n\in J}$ so that $Y = Y_J(X)$ is necessarily an eigen vector of Y_J, and for a general normal observable $\Phi(X)$ is a linear combination of eigen vectors, each with the eigen value λ_p.

So the best choice is always to take a primary observable, and this is done usually by taking a family of maps defined through a finite number of parameters. However the specification process is rarely done by screening all potential, uncountably infinite, choices. One can see it as a random process, where all the possible choices have the same probability. More generally we have the following :

Theorem 4 *For any normal 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 :*

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

So even before proceeding to the measure, we have some hint as what the result will be, and even of the probability that this result occurs. Notice that

there is no assumption about the behavior of the system : the randomization comes only from the choice of a specification. Actually this result is intuitive for primary observables : by taking only a part of the components, the error is proportional to the size of the part which has been dropped. These results provide a strong framework to understand the relation beween measures, that is data as extracted from experiments, and models, that is the formalisation of a theory.

$\widehat{\Phi}$ 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 ψ has also a component in the eigen space corresponding to the null eigen value, which is not observed but exists.

One can extend the definition of primary observables to projections on vector spaces, finite or infinite dimensional. With linear combinations and composition of such operators we get, for any basis $(e_n)_{n\in\mathbb{N}}$ of V, a commutative von Neumann algebra over H, and the algebras associated to different bases are unitary isomorphic. This result is of interest because commutative von Neumann algebras are classified (this is quite technical). If the system is viewed as defined by all possible observables, then it is defined by the von Neumann algebra itself, and one sees that there are not so many possible models. There is a extensive litterature on von Neumann algebras, in the framework of Quantum Mechanics (see notably Bratelli) but we will not follow this venue here.

2.3 Principle of relativity

Let be two observers studying the same system. The first one use a model with variables X belonging to an open subset O of the vector space V, and the second the variables X' belonging to an open subset O' of the same vector space V. As seen above, variables are measured according to procedures, which should include how two observers can compare their data. According to the principle of relativity there is *a* bijective map $U : V \to V$ such that X and X'=U(X) represent the same state of the system. U is given by the definition of the variables, so this is a known map. We will assume that U is continuous. If the models meet the conditions of the theorem 1 then we have two Hilbert spaces H, H', which are isomorphic through V, and because they are defined up to an isometry one can consider that they are the same.

We have the following :

Theorem 5 *Whenever the map U is continuous, 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* Υ *is the linear map :* $\Upsilon : V \to H$ *associated to X.*

As a consequence U is necessary a linear map.
Then for any observables Φ, Φ' we have :
$\widehat{\Phi}' = \Upsilon \circ \Phi \circ U \circ \Upsilon^{-1} = \Upsilon \circ \Phi \circ \Upsilon^{-1} \circ \widehat{U} = \widehat{\Phi} \circ \widehat{U}$ with $\widehat{U} = \Upsilon \circ U \circ \Upsilon^{-1}$

This result is important, because it helps to find the Hilbert space H : in many cases it belongs to a category linked to the structural definitions of the variables.

If the map U is parametrized by a group, that is if there is a map : $U : G \to \mathcal{L}(V;V)$ 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 $\left(\widehat{U}, H\right)$ is a unitary representation of the group G. Unitary representations are well known, so whenever we have such result, it is possible to identify H itself. In particular any topological group G endowed with a Haar measure has at least a unitary representation on a Hilbert space. If G is a Lie group and the map $U : G \to \mathcal{L}(V;V)$ is 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 $T_1 G$ of G.

As a special case, if there is a map : $\widehat{U} : \mathbb{R}_+ \to \mathcal{L}(H,H)$ such that :
$\widehat{U}(t+t') = \widehat{U}(t) \circ \widehat{U}(t')$
$\widehat{U}(0) = Id$
$\lim_{\theta \to 0} \left\| \widehat{U}(\theta) - Id \right\| = 0$
it can be extended to $\mathbb{R}$. $\left(\widehat{U}, H\right)$ is a unitary representation of the abelian group $(\mathbb{R}, +)$. We have a one parameter group, and $\widehat{U}$ is differentiable. There is an infinitesimal generator $S = \frac{d}{ds}\widehat{U}(s)|_{t=0} \in \mathcal{L}(H;H)$ such that : $\widehat{U}(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} S^n = \exp tS$

S is anti-hermitian : $S = -S^*$, normal and has a spectral resolution P : $S = \int_{Sp(S)} sP(s)$. S is not compact, usually its spectrum is continuous.

These properties are of special interest for any model which involves the evolution of the system over time.

2.4 Evolution of a system

Many models deal with the evolution of a system, and actually this is the most usual way to introduce variables belonging to infinite dimensional vector spaces. The parameter "time" t in the argument of the map X (it can be any argument with the same properties, and there can be other arguments beside t, the results below holds in these cases), can have different meanings.

A) t is a parameter used to measure the duration of a phenomenon, usually the time elapsed since some specific event.

B) t is just a parameter used to identify a state of the system : t gives its temporal location.

As we have seen previously universality requires stability, notably of the environment of the system, as it could be appreciated by all the parameters

which are not included in the model. In the first case something has happened at a time 0, and then the system follows an evolution lead by continuous laws. The disturbances, if any, are accounted for in the initial conditions. In the second the laws are similar all over the period : nothing significant differentiates a subperiod from another. The environment stays the same. So the differences between the two depends also of how comprehensive the model is. In theoretical physics, when one ambitions to address everything, we are in the case B : the universe stays the same (at least for any human time scale). In engineering the time accounts for the continuous weakening resulting from infinitesimal stress and we are in the case A. In Economics it is common to take as the origin a recession or a recovery.

We address the two cases in a more precise way, with two different conditions.

In both cases we assume that the following conditions are met :

Conditions 1 :

i) *the variables* $(X_k)_{k=1}^N$ *are maps :* $(X_k)_{k=1}^N :: R \to E$ *where* R *is an open subset of* $\mathbb{R}$ *and* E *a normed vector space*

ii) *The map* $X = (X_k)_{k=1}^N$ *belongs to an open subset* O *of an infinite dimensional Fréchet space* V

iii) $\forall t \in R$ *the evaluation map :* $\mathcal{E}(t) : O \to E : \mathcal{E}(t) X = X(t)$ *is continuous*

The condition iii) is met in all usual cases (such as the integrables functions). We assume that $0 \in R$ (it can be any fixed point in R).

Usually the purpose of such models is to compute and check the law of the evolution of the system, that is to restrict the map X to some subset of O (usually X should meet differential equations). If there is the possibility that two maps X,X', as considered in the model, can take the same values for a non null period of time and take different values for all other times, then clearly we would consider that the model is undetermined : some variables are missing, and should be accounted for to explain the evolution of the system. This condition can be precised in the following theorem :

Theorem 6 *If the conditions 1 above are met and*

if for any X,X' the set $\varpi = \{t \in R : X(t) = X'(t)\}$ *has a non null Lebesgue measure then* $X = X'$

then :

there is a Hilbert space F, vector subspace of E, an open $\widetilde{O} \subset F$ *such that :*

$\forall X \in O \subset V, t \in R : X(0) \in \widetilde{O} \subset F; X(t) \in F$

$\forall t \in R$ *the map :* $\Theta(t) = \mathcal{E}(t) \circ \mathcal{E}(0)^{-1} \in \mathcal{L}(F; F)$ *is unitary*

$\forall t \in R$ *the map :* $\widehat{\mathcal{E}}(t) = \mathcal{E}(t) \circ \mathcal{E}(0)^{-1} \circ \Upsilon^{-1} \in \mathcal{L}(H; F)$ *is an isometry*

Of course this theorem, which assumes only very general properties on the formalization of the system, cannot provide the solutions X. But it provides some valuable results :

- for any $u \in \widetilde{O}$ then $\forall t : X(t) = \Theta(t)(u)$ is well defined and X(t) depends continuously on u : one says that the problem is well posed

- for any value $u \in F, t_0 > 0 \in R$ there is a unique map $X \in V$ such that $X(t_0) = u$ defined by $X = \Upsilon \circ \mathcal{E}(t_0) \circ \mathcal{E}(0)^{-1} \circ \Upsilon^{-1}(u)$: there is a unique map X which goes to a given point. More precisely if there are two such maps they would take the same values almost everywhere in R.

So the model is fully determinist. There is a map $\Theta(t)$ which gives the evolution of the system, the search for the law of evolution can be restricted to a Hilbert subspace F of E, and a unitary operator $\Theta(t)$ on F. E is usually a finite m dimensional vector space, and if we have observations for n dates t_p the statistical problem is to adjust a Hilbert space structure to a m×n matrix [x] . There are statistical methods, based on positive kernels, to deal with this kind of problems (see Berlinet).

Remarks :

i) The proposition 13 of the note "Hilbert spaces in modelling systems" is proven with a more restrictive condition. It can be enlarged, to get the theorem above, with the relation of equivalence $\mathcal{R} : X \sim X' \Leftrightarrow X(t) = X'(t)$ almost everywhere and by taking the quotient $V/\mathcal{R}$.

ii) Notice that the theorem (as well as the previous ones) does not require any continuity conditions on the maps X.

The cases of type B can be specified by the condition that the variables $X'(t) = X(t+\theta)$ and $X(t)$ represent the same state of the system, and we have the stronger result:

Theorem 7 *If the conditions 1 above are met and*

if $R = \mathbb{R}$ and for any fixed $\theta \in R$, the variables $X'_k(t) = X_k(t+\theta)$ and $X_k(t)$ represent the same state of the system

then :

there is a Hilbert space F, an open $\widetilde{O} \subset F$, a continuous anti-hermitian map $\widetilde{S} \in \mathcal{L}(F;F)$ such that :

$\forall X \in O \subset V : X(0) \in \widetilde{O} \subset F$

$\forall t : X(t) = \left(\exp t\widetilde{S}\right)(X(0)) \in F$

The maps X are smooth and $\frac{d}{ds}X(s)|_{s=t} = \widetilde{S}X(t)$

There is a continuous map $S \in \mathcal{L}(V;V)$ such that :

$\Theta(t) = \mathcal{E}(t) \circ \mathcal{E}(0)^{-1} = \exp tS$

$\forall t \in \mathbb{R} : X(t) = (\exp tS \circ X)(0) = \left(\sum_{n=0}^{\infty} \frac{t^n}{n!} S^n X\right)(0)$

and the operator $\widehat{S} = \Upsilon \circ S \circ \Upsilon^{-1}$ associated to S is anti-hermitian

Actually it is easy to see that this is a special case of the previous theorem, and Θ is then defined by the exponential of a unique operator. So, not only the model is determinist, its evolution is some kind of generalized exponential law. This is intuitive : if the system is in a steady environment, it should

have a simple evolution, with some constant law. This law is given by a single anti-hermitian map $\widetilde{S}$ which is directly related to the derivative $\frac{d}{ds}X(s)$, which provides a simpler method of statistical estimation. Notice that, even if X was not assumed to be continuous, smoothness is a necessary result. 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.

These results are of special interest for the studies of systems which encounter, in their evolution, states which are significantly different.

2.4.1 Phases transitions

There is a large class of problems where the maps X belong to the same family but the states X(t), for some periods, take significantly different values in the same vector space E : the system meets a phase transition. The conditions in which these transitions happen are of special interest. Common cases in Physics are change of phases for solid or liquid bodies, the desintegration of a particle, an earthquake,..., in Economics a crisis or a recession, in Finances a rupture in the markets,...

The questions which arise are then : what are the conditions, about the initial conditions or the maps X, for the occurence of such an event ? Can we forecast the time at which such event takes place ?

The states of the system are represented by vectors of E, and the phases can be characterized as connected components of the set E : two different phases have no common point. If the variables X are continuous with respect to t, and R is connected, then the set of values $\{X(t), t \in R\}$ is also connected : there cannot be different phases. So phases transitions imply that the maps X are not continuous and we are not in the conditions of the theorem 7. Totally discontinuous maps exist, but they are strange mathematical objects. Usually discontinuities happen at isolated points (even in Brownian motion) : 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, so 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 θ_p, which is finite over any finite period, and we have a series of phases separated by transitions occurring at precise times. So one can safely say that any scientific sensible model will depict either continuous evolutions (theorem 7) or a finite number of phases, each one corresponding to a continuous evolution, separated by precise instantaneous transitions.

It is legitimate to represent the transitions by a probability law. A sensible assumption is that the probability for a change of phase depends on the proximity of the state of the system from each phase. If we are in the conditions of the theorem 6 it is possible to address practically the problem.

Let us consider two phases, characterized by disjoint subsets E_1, E_2. Their choice is somewhat arbitrary, and anyway would be adjusted from previous data. The Hilbert vector space F can be considered as a subspace of E, so we have two disjoint subsets F_1, F_2. If F_1, F_2 are *closed convex subsets* of F the distance of any point x of F to one of the set F_i is defined by the projection $\pi_i : F \to E_i$: there is a unique $y = \pi_i(x) \in E_i$ such that $\|x - y\|_F$ is minimum. The map π_i is continuous, $\pi_i^2 = \pi_i$ and $\pi_i(x) = x$ when $x \in F_i$. So when we are in the phase F_1 we can relate the probability of a transition 1 -> 2 to $\|X(t) - \pi_2(X(t))\|_F$. And more generally the probability of any transition can be related to the quantity $r(t) = \|X(t) - \pi_1(X(t))\|_F + \|X(t) - \pi_2(X(t))\|_F$.

The result holds if F_1, F_2 are closed *vector subspaces* of F such that $F_1 \cap F_2 = \{0\}$. 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 F_1.

If we assume that the probability of a transition at a time t is a function f of r(t) then a very simple, non parametric, estimator of f can be built as follows. If we have observations over a past period [0,T], from the set of data $\{r(t), t \in [0,T]\}$ one can easily compute the function : $G : \mathbb{R} \to [0,T]$ where G(ρ) is the total duration of the periods when $r(t) \geq \rho$. This is a decrasing curve, from T to 0 when ρ goes from 0 to Max(r(t)). The probability of a transition at any given time when $r(t) \geq \rho$ is $f(\rho)$. One can compute the number of transitions $n(\rho)$ which have occured when $r(t) \geq \rho$, then the estimation $\widehat{f}(\rho)$ of $f(\rho)$ is $\widehat{f}(\rho) = \frac{n(\rho)}{G(\rho)}$.

This provides a general method to estimate the probabilty of a change of phase (such as an earthquake, a financial crisis,...). Practically this requires at first to find F from a batch of data. Then the definition of the phases is easily done by taking vector subspaces.

2.5 Interacting systems

It is common to have two systems which interact with each other, represented by similar models : atoms with their electronic clouds, companies competing in the same markets, tectonic plates... Each system is represented by variables X_1, X_2 following the general conditions of the theorem 1. If we want to represent the two interacting systems together, we would need to add variables Z_1, Z_2 which represent the action of one system on the other : the forces of one atom on the other,...So we would have two different models with the variables $(X_1, Z_1), (X_2, Z_2)$. However usually we do not or cannot measure the interactions. This is more obvious when one has to deal with many similar systems. It would be nice if we could represent the two systems, put together, in a single model, preserving the essential of the interactions, without the need to explicit these interactions. As we have seen the Hilbert space structure is a common

structure for any model, and summarizes the main features of the system. So it is legitimate to look for a Hilbert space structure which combines the Hilbert spaces of each system. We will formalize these requirements by the following conditions :

Conditions 2 :

The model representing the two interacting systems S_1, S_2 *, meeting the requirements of theorem 1, comprises variables Y, belonging to some vector space V', which must meet the following requirements :*

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

ii) Φ *must be such that whenever the systems* S_1, S_2 *are in the states* ψ_1, ψ_2 *then* S *is in the state* ψ' *and*

$\Upsilon'^{-1}(\psi') = \Phi\left(\Upsilon_1^{-1}(\psi_1), \Upsilon_2^{-1}(\psi_2)\right)$

iii) The positive kernel K' of (V', Υ') *must be such that :*

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

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

If the condition i) is met and each system S_1, S_2 , meets the requirements of theorem 1, then the new model meets also the same requirements, and there are a Hilbert space H' and a map $\Upsilon' : V' \to H'$ which associates a vector ψ' to any state of the interacting system, represented by $Y = \Phi(X_1, X_2)$. The condition ii) expresses the compatibility of Φ with $\Upsilon', \Upsilon_1, \Upsilon_2$ and the condition iii) the compatibility of the scalar products. So they are natural, legitimate requirements, whenever one accounts for the Hilbert space structures.

And we have the following :

Theorem 8 *Whenever two systems* S_1, S_2 *interact, there is a model* S *encompassing the two systems and meeting the conditions 2 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.*

This solution is an extension of what is done for models with discreet variables. For instance in Economics if we have one model for which the data are broken down by age, and another one where the breakdown is by income, to account for the interactions the simplest way is to breakdown the data with respect simultaneously to age and income, which is just the tensorial product of the two variables.

If this solution is natural, it is more subtle than it seems and there is much confusion about its meaning.

A tensor is not just the tensorial product of two vectors (if it is so it is said to be separable), it can be the sum of such tensors. If $(e_n)_{n\in\mathbb{N}}, (f_n)_{n\in\mathbb{N}}$ are basis of V_1, V_2 then a basis of $V_1 \otimes V_2$ is $(e_n \otimes f_m)_{n,m\in\mathbb{N}}$ so that $T \in V_1 \otimes V_2$ reads : $T = \sum_{m,n\in\mathbb{N}} T_{mn} e_n \otimes f_m$. If the states of the systems are represented

by $X_1 = \sum_{n\in J_1} X_{1n}e_n, X_2 = \sum_{n\in J_1} X_{2n}f_n$ the state of S is represented by T, where the components T_{mn} are null but for a finite subset in $\mathbb{N}^2$, usually larger than $J_1 \times J_2$. This has two consequences.

i) The measure of T requires more data, and brings more information, than the simple measure of the states of each system : this is intuitive, because S encompasses all the interactions, thus more than just the variables X_1, X_2.

ii) There is no canonical map (basis independant) : $V_1 \otimes V_2 \to V_1 \times V_2$. So there is no simple and unique way to associate two vectors (X_1, X_2) to one tensor T : from the knowledge of the state of the system S there is no way to infere the state of each system S_1, S_2.

This seems paradoxical, as one could imagine that both systems can always be studied, and their states measured, even if they are interacting. But if interactions are suspected, then the environment of each system is not steady : any modification of one system can impact the state of the other, so separate models would be incomplete without accounting for the interactions. We have to keep in mind that, if a model is arbitrary, its use must be consistent : if the scientist intends to study the interactions, they must be present somewhere in the model, as variables for the computations as well as data to be collected. Whence interactions have been acknowledged, 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. 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 of the conditions 2 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 T is always separable). 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. In Quantum Theory this issue is called the entanglement problem, and has been the topic of many discussions, notably with the Bell's inequality.

However there is a way to deal with this issue, it is more illuminating when one considers observables.

2.5.1 Observables in interacting systems and the probabilist interpretation

Using the same definition as previously an observable for the interacting systems is a linear map $\Phi \in L(V_1 \otimes V_2; V_1 \otimes V_2)$ which has a finite dimensional range. It is reasonable to require that whenever S is represented by a separable tensor $T = X_1 \otimes X_2$ then the observable is the tensorial product of observables Φ_1, Φ_2

on $S_1, S_2 : \Phi(X_1 \otimes X_2) = \Phi_1(X_1) \otimes \Phi_2(X_2)$ and there is a unique map $\Phi = \Phi_1 \otimes \Phi_2$ with this property. So observables of the interacting systems are tensorial products of observables of each system.

All the results seen previously for observables hold for $\Phi_1 \otimes \Phi_2$. We have seen that usually the state of S is not represented by a decomposable tensor, so we cannot write : $\Phi(T) = \Phi_1(X_1) \otimes \Phi_2(X_2)$ but any tensor is the sum of decomposable tensors. This can always be done in any basis. However it makes sense to look for a decomposition which is more significant, that is :

$\Phi(T) = \sum_{p_1,p_2} A_{p_1 p_2} \Phi_1(\xi_{1p_1}) \otimes \Phi_2(\xi_{2p_2})$

where ξ_{1p}, ξ_{2p_1} represent states of S_1, S_2 which can be easily identified. Then $A_{p_1 p_2}$ can be seen (up to a constant) as the probability that the system S_1 is in the state ξ_{1p_1} and simultaneously the system S_2 is in the state ξ_{2p_2}. The probabilist interpretation is purely formal : the interacting system S has no random behaviour, but the result can be seen as if it was the superposition of states between which the two systems would hesitate. This is a common interpretation in Quantum Mechanics : $\xi_{1p} \otimes \xi_{2p_1}$ are identified as "pure states" and the actual state is a superposition of pure states, as it has been (in an awkward manner) popularized with the "Schrödinger's cat". The issue is that it is usually possible to measure independantly the states of the systems S_1, S_2 but, as long as they interact, only T matters and the decomposition of T is not granted.

This interpretation is simpler when one considers many similar interacting systems.

2.5.2 Homogeneous systems

An interesting case is many similar systems (that we will call microsystems) represented by the same model, interacting together. Each microsystem is labeled by s = 1...N (we assume that their number N is fixed) with variables $(X_s)_{s=1}^N$ satisfying the conditions of the theorem 1 : X_s belong to an open O of an infinite dimensional separable Fréchet space V. For each microsystem the Hilbert space H and the linear map Υ are the same, but the vectors ψ_s representing the states are different quantities. 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 :

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

The laws for each microsystem can be different. For instance they can depend on the size of the microsystem. So in the general case the label s 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 $(X_s)_{s=1}^N$.

However there are models where one can assume that the microsystems are identical, in the meaning that they have the same behavior. This requires that they have the same size or relevant property : they are indistinguishable. We

will say that these interacting systems are **homogeneous** and we will characterize this assumption by the property that any permutation of the N microsystems gives the same state of the total system. 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 separable tensors. Thus the theorem 5 holds and to U is associated a unitary operator $\widehat{U}$ on $\mathbf{H}_N$ such that $\left(\mathbf{H}_N, \widehat{U}\right)$ is a unitary representation of $\mathfrak{S}(N)$. To be consistent, for a given system, it must be an irreducible representation $\left(\mathbf{h},\widehat{U}\right)$ with $\mathbf{h} \subset \mathbf{H}_N$. Such representations are finite dimensional and in bijective correspondance with the classes of conjugacy $\lambda = \{(n_k)_{k=1}^p \in \mathbb{N}^p : 0 \leq n_p \leq ... \leq n_1 \leq N, n_1 + ...n_p = N\}$ of $\mathfrak{S}(N)$. For such a class $\mathbf{h}$ is generated by p distinct vectors $(\widetilde{\varepsilon}_j)_{j=1}^p$ of any Hilbertian basis of H, then a tensor of $\mathbf{h}$ reads :

$\forall \Psi \in \mathbf{h} : \Psi = \sum_{\sigma \in \mathfrak{S}(\lambda^c)} \Psi^\sigma \widehat{U}(\sigma) \left(\otimes_{n_1} \widetilde{\varepsilon}_1 \otimes_{n_2} \widetilde{\varepsilon}_2 ... \otimes_{n_p} \widetilde{\varepsilon}_p\right)$

where $\mathfrak{S}(\lambda^c)$ is the subgroup of the permutations $\mathfrak{S}(N)$ which do not leave invariant $\otimes_{n_1} \widetilde{\varepsilon}_1 \otimes_{n_2} \widetilde{\varepsilon}_2 ... \otimes_{n_p} \widetilde{\varepsilon}_p$. The dimension of $\mathbf{h}$ his given by the cardinality of $\mathfrak{S}(\lambda^c)$ that is : $\frac{N!}{n_1!...n_p!}$. All the vector spaces $\mathbf{h}$ of the same conjugacy class have the same dimension, thus they are isomorphic. But the vector spaces $\mathbf{h}$ of different conjugacy classes are not isomorphic.

So the states of the interacting system belong to a finite dimensional vector space $\mathbf{h}$, even if the states of the microsystems are represented in infinite dimensional vector spaces H. The interactions bring order, and reduce the scope of possible states for the whole system.

Any tensor of $\mathbf{h}$ is a linear combination of permutations of the separable tensor $\otimes_{n_1} \widetilde{\varepsilon}_1 \otimes_{n_2} \widetilde{\varepsilon}_2 ... \otimes_{n_p} \widetilde{\varepsilon}_p$ which can be seen as representing a configuration where n_k microsystems are in the same state $\widetilde{\varepsilon}_k$. If O is a convex subset then S belongs to a convex subset, and the basis of $\mathbf{h}$ 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 being in the state X_k. As exposed above the identification with a probabilist model is formal : there is no random behaviour for the physical system.

For homogeneous systems we can define observables as previously, which are estimates of the state of the full system, but one has usually simpler linear maps : $G : H_N \to F$ valued in a finite dimensional vector space F (it can be a scalar). As such they provide only a limited information about the state S, but are more easily measured. Because the system is homogeneous, the map must be symmetric and, as a consequence the associated map $\widehat{G} = G \circ \Upsilon_N^{-1}$ is such that :

$\widehat{G}(\Psi) = \sum_{\sigma \in \mathfrak{S}(\lambda^c)} \Psi^\sigma \widehat{G}\left(\widehat{U}(\sigma)\left(\otimes_{n_1} \widetilde{\varepsilon}_1 \otimes_{n_2} \widetilde{\varepsilon}_2 ... \otimes_{n_p} \widetilde{\varepsilon}_p\right)\right) = C_\lambda(\Psi) G_\lambda$

with :

$G_\lambda = \widehat{G}\left(\otimes_{n_1} \widetilde{\varepsilon}_1 \otimes_{n_2} \widetilde{\varepsilon}_2 ... \otimes_{n_p} \widetilde{\varepsilon}_p\right) \in F$ and $C_\lambda(\Psi) = \sum_{\sigma \in \mathfrak{S}(\lambda^c)} \Psi^\sigma \in \mathbb{R}$

G_λ depends only on the class of conjugacy λ, that is the distribution n_k of the microsystems among the states $\widetilde{\varepsilon}_k$: for a given class of system, defined by the same variables, G_λ defines a set of vectors which do not depend of the state Ψ, and could, at least theoretically, be computed for each value of λ. $C_\lambda(\Psi)$ in a scalar. As the dimension of F is usually small, with respect to the number of classes of conjugacy, such a global variable provides only a limited information about the actual λ, and never the values of Ψ^σ for each σ.

In the probabilist interpretation each microsystem is seen as behaving independantly, so one can attribute a probability π_i for any microsystem to be in the state $\widetilde{\varepsilon}_i$ and then compute the probability associated to any distribution of states characterized by a class of conjugacy λ. The value of a global variable G then provides an estimate for the probability of such distribution. This is the starting point for the studies based on the entropy, defined as $E = -\sum_{j=1}^{N} \pi_j \ln \pi_j$. This definition, which stems from Statisics and Information theories, can be justified, in a more general framework, by the previous results : the state of a system is characterized by the class of conjugacy, the set of classes of conjugacy can be partially ordered and a quantity such as $-\sum_{j=1}^{N} \frac{n_j}{N} \ln \frac{n_j}{N}$ can be seen as a proxy for the measure of the disorder in the distribution of states.

The evolution of a homogeneous system can be studied using the previous results. If each microsystem meets the conditions of theorem 6, then there is a Hilbert space F such that at each time t the state of the system $\Psi(t)$ belongs to a finite dimensional vector space $\mathbf{f}$(t) defined by a class of conjugacy $\lambda(t)$ and a set of p distinct vectors $(\varphi_j(t))_{j=1}^{p}$ of any Hilbertian basis of F, both depending on t. Moreover if the evolution is continuous but at isolated points, then to each class of conjugacy is associated a fixed vector space $\mathbf{f}_\lambda$, which can be seen as different phases, and the evolution of the system is among the different classes of conjugacy. In a probabilist picture, one can assume that the probability for the system to be in a phase $\lambda : \Pr(S(t) \in \mathbf{f}_\lambda)$ is a function of $\frac{\|\pi_\lambda S(t)\|^2}{\|S(t)\|^2}$, where $\pi_\lambda$is the projection on $\mathbf{f}_\lambda$,and it can be estimated from data on a past period, with the knowlege of both λ and $\frac{\|\pi_\lambda S(t)\|^2}{\|S(t)\|^2}$. So not only the interactions between systems bring order, it also regulates the evolutions of global systems between phases.

If the number N of microsystems varies the previous results can be extended, at least in part : the state Ψ belongs to a Fock space, which is an infinite dimensional Hilbert space. This is used in Quantum Theory of Fields.

2.6 CONCLUSION

i) In the previous results Physicists have recognized the basic axioms of Quantum Mechanics (QM) , and actually they have been a lead in their discovery. They justify, clarify and enlarge methods commonly used in QM, by removing

the theoretical issue of scale and the practical hassle of projective complex representations, and thus they can be safer. But of course the scope of QM is larger than these axioms.

The formalism introduced here can be useful in theoretical Economics, both at the micro and macro levels, and in particular in understanding the emergence of typical structures in economic systems, for instance dominant economic models of companies in a market.

The results can have a more direct, practical interest, in the study of transitions, by providing a general framework to characterize the phases and a formula to estimate the related probabilities. This addresses in particular the statistical analysis of large batches of empirical data, in Mechanics, Geology, Financial Markets, consumer studies,...

ii) The axioms of QM have been a source of puzzlement, as it seemed strange that the general laws of nature should obey such a precise, sophisticated, mathematical expression, as eigen values. What we can see here is that these axioms are neither limited to the physical world, nor to the natural sciences. Most of the quantitative models of any scientific theory can be represented in a common mathematical formalism. It stems from the way we express our scientific laws and from the Mathematics that we use. Mathematics do not come out of the blue, they are not just a set of logical statements, they are based on axioms which, by sometimes lengthy deductions, provide the tools that we use and the axioms are themselves the justification of these tools, afterwards. The mathematical structures involved here, Hilbert spaces, are the generalization of vector spaces and scalar products, which themselves have been inspired by Geometry and the observation of the physical world. So the existence of a common structure to scientific theories can be seen as a strong argument for positivism and the unity of the real world, but it also can be seen also as the indication of the existence of some deeper organization of our mind and human intelligence.

iii) In scientific activities a considerable interest has been given, with good reasons, to the precision of measures. However the uncertainty which results from the specification of the models has received scant attention, except perhaps in Economics where the issue is more obvious. The concept of observable provides a general way to address the problem. The unavoidable discrepancy between the sophistication of our mathematical tools and the limited capability of our measures leads to the introduction of probabilist laws, not as the recognition of some random behaviour of reality, but as an efficient way to build a manageable representation of what it is. This fact is well understood in Thermodynamics but, unfortunately, not in particles physics. The fact that it is difficult to observe a particle requires, more than for bodies at a larger scale, to rely on crude specifications, so that the related uncertainty can become significant. When one considers a particle targeted to two close slits in a screen, it is clear that complicated things may happen, and it is sensible to opt for several possible trajectories with some probability law, without assuming some bizarre behavior from the particle.

iv) It is a common belief that interactions between a great number of systems should create complexity, and complexity means disorder and ruptures. It seems to be usually false on the first count - homogeneous systems show a finite number of typical behaviours - and partially true on the second : ruptures occur as transitions between phases. Physicists are used to such phenomena in Thermodynamics, but this results should open new areas of studies in Social Sciences. However, if the laws of Thermodynamics are valid, the Second Principle and its Arrow of Time are perhaps not as general as it is usually asserted : the evolution of a system is either continuous, with a single phase, or a series of transitions between different phases, but it does not seem that one goes always towards more disorder.

v) Discontinuous and irreversible processes are common at our scale. We can see how they can be the result of transitions of phenomena occurring at a smaller scale, but we do not see why the transitions happen. This could be an argument for a discontinuous, random, organization of the real world, but, as we have noticed before, the universality of scientific laws commands continuity and determinism. In the extremely general framework used here, one answer can be that models with discontinuous processes are incomplete : as we have seen it is always possible to replace a discreet variable (such that the appartenance to a phase) by a continuous variable taking discreet values on the phases, so one can argue that such models should be completed by some super, continuous, variable signaling the phase. This is not contradictory with the Principle of least action of Physics : what it tells is that the action is stationary at equilibrium, but of course there could be more than one equilibrium, and actually this is the starting point for the Equilibrium breakdown theories.

vi) The results, and notably the relation with Hilbert spaces, have been obtained under precise conditions, which can seem restrictive. It was the price to pay for an easy way to a practical formalism. However it seems that most of them could hold in a larger context. Category Theory is a normal starting point for such an endeavour. A theory (as defined before) can be seen as a category, systems are products of categories, and variables are functors to a category of vector spaces. Then the key would be to prove that a theory has a representation in the category of Hilbert spaces. However categories lead quite often to convoluted developments. One could consider to take the couple object / variables as the building bricks, systems as products (accounting for the interactions between objects), the equivalent of the functors then being their representation in a formal system.[4]

[4]Jean Claude Dutailly : jc.dutailly@free.fr

3 BIBLIOGRAPHY

J.C. Baez, M.Stay *Physics, Topology, Logic and Computation: A Rosetta Stone* arXiv 0903.0340 (2009)

A.Berlinet, C.Thomas-Agnan *Reproducing kernel Hilbert spaces in probability and statistics* Kluwer Academic Publishers, Boston, MA (2004).

A.Bird *Philosophy of science* Rootledge (1998)

O.Bratelli, D.W.Robinson *Operators algebras and quantum statistical mechanics* Springer (2002)

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

J.C.Dutailly *Hilbert spaces in modelling of systems* hal-00974251 v1, 5 apr 2014

J.C.Dutailly *Mathematics for theoretical physics* arXiv:1209-5665v2 [math-ph] 4 feb 2014

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

Tepper L.Gill, G.R.Pantsulaia, W.W.Zachary *Constructive analysis in infinitely many variables* arXiv:1206-1764v2 [math-FA] 26 june 2012

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)

E.T.Jaynes *Where do we stand on maximum entropy ?* Paper for a MIT conference (1978)

H.Poincaré *La valeur de la science* Flammarion (rep.2014)

H.Poincaré *La Science et l'hypothèse* Flammarion (rep.2014)

A.Smola, A.Gretton, L.Song, B.Schölkop *A Hilbert Space Embedding for Distributions*

J. E. Palomar Tarancon *Conceptual systems, conceptual convergence theory and and algebras of analogies*

K.Popper *The logic of scientific discovery* Rootledge (1959)

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

S.Weinberg *Dreams of a Final Theory* Pantheon Books (1992)