



HAL
open science

QUANTUM MECHANICS REVISITED

Jean Claude Dutailly

► **To cite this version:**

| Jean Claude Dutailly. QUANTUM MECHANICS REVISITED. 2014. hal-00770220v2

HAL Id: hal-00770220

<https://hal.science/hal-00770220v2>

Preprint submitted on 20 Aug 2014 (v2), last revised 1 Jul 2015 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Quantum Mechanics Revisited

Jean Claude Dutailly
Paris (France)

August 20, 2014

Abstract

The purpose of the paper is to study the foundations of the main axioms of Quantum Mechanics. From a general study of the mathematical properties of the models used in Physics to represent systems, we prove that the states of a system can be represented in a Hilbert space, that a self-adjoint operator is associated to any observable, that the result of a measure must be the eigen value of the operator and appear with the usual probability. Furthermore an equivalent of the Wigner's theorem holds, which leads to the Schrödinger equation.

These results are based on well known mathematics, and do not involve any specific hypothesis in Physics. They validate and explain the methods currently used, which are made simpler and safer, and open new developments.

In the second edition of this paper important developments have been added about interacting systems and the transitions of phases.

After almost a century the interpretation of quantum mechanics stays a largely open subject. If, for most of the workers who use it everyday, this is not a matter of concern, the unending flow of papers on this topic shows that, for some people at least, this is an issue. Rightly so, because, whatever one's philosophical belief, one cannot feel comfortable with a successful scientific theory which, according to some of the most authorized voices in physics, is beyond our understanding. And a scientist cannot truly be convinced by the usual argument : "It works, so we have to accept it". The capability to provide experimentally verifiable predictions is not the only criterium for a scientific theory. A "black box" in the "cloud" which answers rightly to our questions is not a scientific theory, if we have no knowledge of the basis upon which it has been designed. A scientific theory should provide a set of concepts and a formalism which can be easily and indisputably understood and used by the workers in the field. And this leads to look for a theory which helps us to describe, understand and as far as it is possible, explain, the world we live in.

So where do we stand in Quantum Physics ? The question is a bit muddled, as any student discovers in the many "introductory books" on the subject, by a constant mixing of physical experiments and formal systems, where it is not always easy to understand if the formal system validates the experiments or the

converse. So he can be told that the position and the momentum of a particle cannot be simultaneously measured because their operators do not commute, but if he asks why it is not so at a macroscopic level, he faces a long explanation which sums up usually to "this is the quantum world", understand the magic kingdom.

Actually, "Quantum Physics" encompasses several theories, with three distinct areas :

i) The duality matter / wave : the indisputable fact that particles can behave like fields which propagate, and conversely force fields can behave like pointwise particles. This departure from the classical picture (say Newtonian mechanics and Maxwell's electromagnetism) requires a new formalism, that is certainly well rendered by the quantum mechanics proper, but goes far beyond. The string, brane, or quantum loop theories illustrate the need for a new model for the world of particles. The spin of a particle is a phenomenon which requires similarly a new theoretical foundation.

ii) The quantum mechanics (QM) which is presented in all the books on the subject (such as summarized by Weinberg) as a set of axioms :

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

- Observables are represented by hermitian operators

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

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

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

and, depending on the authors, the Schrödinger equation.

iii) The Quantum theory of fields (QTF) which accounts for the interaction between particles and fields, and is summarized in the standard model.

These three aspects are entangled, for historical, practical and pedagogical reasons, but are distinct. Whatever their success as a predictive tool, QM and QTF cannot alone explain the duality matter / field, and the fact that the basic axioms of QM should apply to any system is still a matter of puzzlement.

In this paper I will focus on Quantum Mechanics proper, that is the set of axioms which sustain the models developed for studying the atomic and subatomic world. My purpose is to look for a logical and physical basis for these axioms. So I will stay on the most general level of physics, meaning a system which could be any object of the study of a physicist and I will not consider the meaning of concepts such as particles, fields or spin.

Since the seminal von Neumann's book most of the work on the formalization of Quantum Mechanics have followed the path of an algebraic construct, whether in the general picture (Branthall, Araki and others) or in the quantum theory of fields (Halvorson and others). In this framework the focus is moved from the Hilbert space to the set of observables, and indeed a system is itself

defined through the algebra of its observables. This provides a more comfortable background to develop a mathematical theory, notably with respect to the always sensitive issues of continuity, and many results that are certainly useful. But this approach has a fundamental drawback : it leads further from an understanding of the physical foundations of the theory itself. To tell that a system should be represented by a von Neumann algebra does not explain more than why a state should be represented in a Hilbert space at the beginning. The sophistication of the mathematical wrapping does not improve the understanding of the foundations of Quantum Mechanics : in both cases the axioms are just that, they are granted. And actually they are more muddled. There is no use to repeat that "the experiments validate the theory" : as long as the theory does not tell us why and how (other than through a philosophical discourse) it does not work at a macroscopic scale, it is not validated, indeed it is invalidated daily.

The approach in this paper is, in some ways, the opposite. We will focus on the interaction between measures and formal representations of a system. And to do so we will stick mainly to the common presentation of Quantum Mechanics, which, besides its formal imperfections (which are not my concern), is closer to the physics as it is done every day. We will just try to understand what lies besides the narrative which begins with something like "Let be a system,..." and later goes on by "let be X,Y, Z the fields,...". As this is the universal presentation of any physical experiment, it should deserve more than a putative glance. And from the study of the relations between experiments and formal physical theory, in the most general context, we will **prove** the following :

- the state of a system can be represented in a Hilbert space
- if two systems interact it is possible to represent the states in the tensor product of the Hilbert spaces
- to each observable is associated a self-adjoint operator
- the results which can be obtained through an observable belong to the vector space of its operator, and they appear with a probability $|\langle\psi, \psi_k\rangle|^2$
- an analog to the Wigner's theorem holds for any gauge transformations between observers
- the Schrödinger equation holds in the General Relativity context, and the presence of a universal constant such as \hbar is necessary

Many theorems will be used in this paper. They concern a broad range of mathematical topics which can be not known from the reader. So it is convenient to refer these theorems to a compendium of mathematics which has been published on arXiv (2nd edition). They are cited as (JCD Th. XXX).

This is the second edition of the paper, with an important development on interacting systems.

1 HILBERT SPACE

1.1 Representation of a system

Let us start with common Analytic Mechanics. A system, meaning a delimited area of space comprising material bodies, is represented by scalar generalized coordinates $q = (q_1, \dots, q_N)$, its evolution by the derivatives $\dot{q} = (\dot{q}_1, \dots, \dot{q}_N)$. By extension q can be the coordinates of a point Q of some manifold M to account for additional constraints, and then the state of the system at a given time is fully represented by a point of the vector bundle $TM : W = (Q, V_Q)$. According to the Least Action Principle the derivatives \dot{q} can be exchanged with conjugate momenta, and the state of the system is then represented in the phase space, with a symplectic structure. But we will not use this addition and stay at the very first step, that is the representation of the system by (q, \dot{q}) .

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

But let us consider two facts :

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

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

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

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

ii) the state of the system is represented by a fixed finite number N of variables $X = (X_k)_{k=1}^N$, which can be maps defined on Ω , with their derivatives so that the state of the system is defined by a finite number of maps, which usually belong to infinite dimensional vector spaces.

And it is legitimate to substitute the maps to the coordinates in Ω . We still have infinite dimensional vector spaces, but we have some mathematical tools to deal with them.

This representation is clearly linked to the measure process : all the measures taken on the system shall account for the variables X , meaning for their mathematical format (vectors, scalars,...) and the maps are estimated from the data by a statistical process. It is also the starting point for the formulation of any theory (which assumes some relation between the variables) and the com-

putation of predictions from initial data. Thus the model is really at the core of the methods in Physics.

1.2 Fundamental theorem

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

Conditions 1 :

- i) The system is represented by a fixed finite number N of variables $(X_k)_{k=1}^N$*
- ii) Each variable belongs to an open subset O_k of a separable Fréchet real vector space V_k*
- iii) At least one of the vector spaces $(V_k)_{k=1}^N$ is infinite dimensional*
- iv) For any other model of the system using N variables $(X'_k)_{k=1}^N$ belonging to open subset O'_k of V_k , and for $X_k, X'_k \in O_k \cap O'_k$ there is a continuous map : $X'_k = F_k(X_k)$*

Remarks :

i) The variables are assumed to be independant, in the meaning that there is no given relation such that $\sum_k X_k = 1$. But the relations which are tested in the model do not matter here. Actually to check the validity of a model one considers all the variables, those which are given and those which can be computed, they are all subject to measures and this is the comparison, after the experiment, between computed values and measured values which provides the validation. So in this initial stage of specification of the model there is no distinction between the variables, which are on the same footing. Similarly there is no distinction between variables internal and external to the system : if the evolution of variable is determined by the observer or phenomena out of the system (it is external) its value must be measured to be accounted for in the model, so it is on the same footing as any other variable.

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

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

iii) A Fréchet space is a Hausdorff, complete, topological space endowed with a countable family of semi-norms (JCD 971). It is locally convex and metric. Are Fréchet spaces :

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

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

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

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

iv) The condition iv addresses the case when the variables are defined over connected domains. The change of variables is addressed later on.

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

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

Proof. i) Each value of the set \mathcal{S} of variables defines a state of the system,

denoted X, belonging to the product $O = \prod_{k=1}^N O_k \subset V = \prod_{k=1}^N V_k$. The couple (O,X), together with the property iv) defines the structure of a Fréchet manifold

M on the set \mathcal{S} , modelled on the Fréchet space $V = \prod_{k=1}^N V_k$. The coordinates are

the values $(x_k)_{k=1}^N$ of the functions X_k . This manifold is infinite dimensional. Each Fréchet space is metric, so V is a metric space, and M is metrizable.

ii) As M is a metrizable manifold, modelled on an infinite dimensional separable Fréchet space, the Henderson's theorem (Henderson - corollary 5, JCD 1386) states that it can be embedded as an open subset Ω of an infinite dimensional separable Hilbert space H, defined up to isomorphism. Moreover this

structure is smooth, the set $H - \Omega$ is homeomorphic to H , the border $\partial\Omega$ is homeomorphic to Ω and its closure $\overline{\Omega}$.

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

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

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

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

Notice that this is a real vector space.

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

1.3 Basis

Proposition 2 *For any basis $(e_i)_{i \in I}$ of V contained in O , there are unique families $(\varepsilon_i)_{i \in I}, (\phi_i)_{i \in I}$ of independant vectors of H , a linear isometry $\Upsilon : V \rightarrow H$ such that :*

$$\begin{aligned} \forall X \in O : \Upsilon(X) &= \sum_{i \in I} \langle \phi_i, \Upsilon(X) \rangle_H \varepsilon_i \in \Omega \\ \forall i \in I : \varepsilon_i &= \Upsilon(e_i) \\ \forall i, j \in I : \langle \phi_i, \varepsilon_j \rangle_H &= \delta_{ij} \\ \text{and } \Upsilon &\text{ is a compatible chart of } M. \end{aligned}$$

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

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

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

where the set $\mathbb{R}_0^I \subset \mathbb{R}^I$ is the subset of maps $I \rightarrow \mathbb{R}$ such that only a finite number of components x_i are non null.

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

The inner product on H defines a positive kernel :

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

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

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

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

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

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

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

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

H_0 can be defined similarly by the bijective map :

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

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

$\ell^2(I) \subset \mathbb{R}^I$ is the set of families $y = (y_i)_{i \in I} \subset \mathbb{R}^I$ such that : $\sup(\sum_{i \in J} (y_i)^2) < \infty$ for any countable subset J of I .

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

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

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

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

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

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

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

$$o(\psi) \in H_1^\perp, H_0 \subset H_1 \Rightarrow o(\psi) \in H_0^\perp$$

$$\Rightarrow \forall i \in I : \langle \varepsilon_i, o(\psi) \rangle_H = 0 = K_V(\Xi^{-1}(\varepsilon_i), \Xi^{-1}(o(\psi))) = K_V(e_i, \Xi^{-1}(o(\psi)))$$

$$\Rightarrow \Xi^{-1}(o(\psi)) = 0 \Rightarrow o(\psi) = 0$$

$$H_1^\perp = 0 \text{ thus } H_1 \text{ is dense in } H \text{ (JCD Th.1115), and as it is closed : } H_1 = H$$

$(\tilde{\varepsilon}_i)_{i \in I}$ is a hilbertian basis of H and

$$\forall \psi \in H : \psi = \sum_{i \in I} \langle \tilde{\varepsilon}_i, \psi \rangle_H \tilde{\varepsilon}_i \text{ with } \sum_{i \in I} |\langle \tilde{\varepsilon}_i, \psi \rangle_H|^2 < \infty$$

$$\Leftrightarrow (\langle \tilde{\varepsilon}_i, \psi \rangle_H)_{i \in I} \in \ell^2(I)$$

H_0 is the interior of H , it is the union of all open subsets contained in H , so $\Omega \subset H_0$

$H_0 = \text{Span}((\tilde{\varepsilon}_i)_{i \in I})$ thus the map :

$$\tilde{\pi}_H : H_0 \rightarrow \mathbb{R}_0^I :: \tilde{\pi}_H(\sum_{i \in I} \tilde{y}_i \tilde{\varepsilon}_i) = \tilde{y} = (\tilde{y}_i)_{i \in I}$$

is bijective and : $\tilde{\pi}_H(H_0) = \tilde{R}_0 \subset \mathbb{R}_0^I \subset \ell^2(I)$

Moreover : $\forall \psi \in H_0 : \tilde{\pi}_H(\psi) = (\langle \tilde{\varepsilon}_i, \psi \rangle_H)_{i \in I} \in \mathbb{R}_0^I$

Thus :

$\forall X \in O : \Xi(X) = \sum_{i \in I} \langle \tilde{\varepsilon}_i, \Xi(X) \rangle_H \tilde{\varepsilon}_i \in \Omega$

and $\tilde{\pi}_H(\Xi(X)) = (\langle \tilde{\varepsilon}_i, \Xi(X) \rangle_H)_{i \in I} \in \tilde{R}_0$

$\forall i \in I, e_i \in O \Rightarrow \Xi(e_i) = \varepsilon_i = \sum_{j \in I} \langle \tilde{\varepsilon}_j, \varepsilon_i \rangle_H \tilde{\varepsilon}_j$

and $\tilde{\pi}_H(\varepsilon_i) = (\langle \tilde{\varepsilon}_j, \varepsilon_i \rangle_H)_{j \in I} \in \tilde{R}_0$

iv) Let be : $\tilde{e}_i = \Xi^{-1}(\tilde{\varepsilon}_i) \in V_0$ and $\mathcal{L}_V \in GL(V_0; V_0) :: \mathcal{L}_V(e_i) = \tilde{e}_i$

We have the following diagram :

$$\begin{array}{ccccc}
 & \Xi & & \mathcal{L}_H^{-1} & \\
 e_i & \rightarrow & \varepsilon_i & \rightarrow & \tilde{\varepsilon}_i \\
 & \searrow & & & \downarrow \\
 & \mathcal{L}_V & \searrow & & \downarrow \Xi^{-1} \\
 & & & \searrow & \downarrow \\
 & & & & \tilde{e}_i
 \end{array}$$

$\langle \tilde{e}_i, \tilde{e}_j \rangle_V = \langle \Xi(\tilde{e}_i), \Xi(\tilde{e}_j) \rangle_H = \langle \tilde{\varepsilon}_i, \tilde{\varepsilon}_j \rangle_H = \delta_{ij}$

So $(\tilde{e}_i)_{i \in I}$ is an orthonormal basis of V_0 for the scalar product K_V

$\forall X \in V_0 : X = \sum_{i \in I} \tilde{x}_i \tilde{e}_i = \sum_{i \in I} \langle \tilde{e}_i, X \rangle_V \tilde{e}_i$ and $(\langle \tilde{e}_i, X \rangle_V)_{i \in I} \in \mathbb{R}_0^I$

The coordinates of $X \in O$ in the basis $(\tilde{e}_i)_{i \in I}$ are $(\langle \tilde{e}_i, X \rangle_V)_{i \in I} \in \mathbb{R}_0^I$

The coordinates of $\Xi(X) \in H_0$ in the basis $(\tilde{\varepsilon}_i)_{i \in I}$ are $(\langle \tilde{\varepsilon}_i, \Xi(X) \rangle_H)_{i \in I} \in \mathbb{R}_0^I$

$\langle \tilde{\varepsilon}_i, \Xi(X) \rangle_H = \langle \Xi(\tilde{e}_i), \Xi(X) \rangle_H = \langle \tilde{e}_i, X \rangle_V$

Define the maps :

$\tilde{\pi}_V : V_0 \rightarrow \mathbb{R}_0^I :: \tilde{\pi}_V(\sum_{i \in I} \tilde{x}_i \tilde{e}_i) = \tilde{x} = (\tilde{x}_i)_{i \in I}$

$\Upsilon : V_0 \rightarrow H_0 :: \Upsilon = \tilde{\pi}_H^{-1} \circ \tilde{\pi}_V^{-1}$

which associates to each vector of V the vector of H with the same components in the orthonormal bases, then :

$\forall X \in O : \Upsilon(X) = \Xi(X)$

and Υ is a bijective, linear map, which preserves the scalar product, so it is continuous and is an isometry.

v) There is a bijective linear map : $\mathcal{L}_H \in GL(H_0; H_0)$ such that : $\forall i \in I : \varepsilon_i = \mathcal{L}_H(\tilde{\varepsilon}_i)$.

$(\tilde{\varepsilon}_i)_{i \in I}$ is a basis of H_0 thus $\varepsilon_i = \sum_{j \in I} [\mathcal{L}_H]_i^j \tilde{\varepsilon}_j$ where only a finite number of coefficients $[\mathcal{L}_H]_i^j$ is non null.

Let us define : $\varpi_i : H_0 \rightarrow \mathbb{R} :: \varpi_i(\sum_{j \in I} \psi_j \varepsilon_j) = \psi_i$

This map is continuous at $\psi = 0$ on H_0 :

take $\psi \in H_0, \|\psi\| \rightarrow 0$ then $\psi = \sum_{i \in I} \langle \tilde{\varepsilon}_i, \psi \rangle_H \tilde{\varepsilon}_i$ and $\tilde{\psi}_j = \langle \tilde{\varepsilon}_j, \psi \rangle_H \rightarrow 0$ so

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

$\psi_i = \sum_{j \in J} [\mathcal{L}_H]_i^j \tilde{\psi}_j \Rightarrow |\psi_i| < \varepsilon \sum_{j \in I} \max |[\mathcal{L}_H]_i^j|$ and $(|[\mathcal{L}_H]_i^j|)_{j \in I}$ is

bounded $\Rightarrow |\psi_i| \rightarrow 0$

Thus ϖ_i is continuous and belongs to the topological dual H'_0 of H_0 . It can be extended as a continuous map $\overline{\varpi}_i \in H'$ according to the Hahn-Banach

theorem (JCD Th.958). Because H is a Hilbert space, there is a vector $\phi_i \in H$ such that : $\forall \psi \in H : \overline{\omega}_i(\psi) = \langle \phi_i, \psi \rangle_H$ so that :

$$\forall X \in O : \Upsilon(X) = \Xi(X) = \sum_{i \in I} \psi_i \varepsilon_i = \sum_{i \in I} \langle \phi_i, \psi \rangle_H \varepsilon_i = \sum_{i \in I} \langle \phi_i, \Xi(X) \rangle_H \varepsilon_i$$

$\forall i \in I :$

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

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

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

1.3.1 Remarks

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

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

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

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

1.4 Complex structure

The variables and vector spaces are real (the Henderson's theorem holds only for real structures) and H is a real Hilbert space. However it can be useful to endow H with the structure of a complex Hilbert space : the set does not change

but one distinguishes real and imaginary components, and the scalar product is given by a hermitian form. Notice that this is a convenience, not a necessity.

Proposition 3 *Any real separable infinite dimensional Hilbert space can be endowed with the structure of a complex separable Hilbert space*

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

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

Define :

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

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

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

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

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

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

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

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

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

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

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

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

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

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

γ is sesquilinear

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

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

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

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

Thus γ is definite positive ■

1.5 Decomposition of the Hilbert space

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

Proposition 4 *If the model is comprised of N continuous variables $(X_k)_{k=1}^N$, each belonging to a separable Fréchet vector space V_k , then the real Hilbert space H of states of the system is the Hilbert sum of N Hilbert space $H = \bigoplus_{k=1}^N H_k$ and any vector ψ representing a state of the system is uniquely the sum of N vectors ψ_k , each image of the value of one variable X_k in the state ψ*

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

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

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

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

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

$$K((X_1, \dots, X_N), (X'_1, \dots, X'_N)) = \sum_{k=1}^N K_k(X_k, X'_k) = \sum_{k=1}^N \langle \Upsilon(X_k), \Upsilon(X'_k) \rangle_{H_k}$$

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

1.6 Discrete variables

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

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

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

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

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

Proposition 5 *If the condition of the proposition 1 are met, and the set O of possible states of the system has two connected components O_1, O_2 then there is a continuous function $f : V \rightarrow [0, 1]$ such that $f(\Upsilon(X)) = 1$ in O_1 and $f(\Upsilon(X)) = 0$ in O_2*

Proof. The connected components O_1, O_2 of a topological space are closed, so O_1, O_2 are disjoint and both open and closed in V .

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

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

2 OBSERVABLES

The value of the variables in a model are estimated from a batch of data, by statistical methods. In a first step a specification is assumed for the functions (affine, polynomial,...) so that they depend on a limited number of parameters, and in a second step these parameters are adjusted to fit the data. We will not dwell on the statistical methods, which are presumed to be defined and fixed (such as the size of a sample), so the data will not be involved. We will focus on the first step : the specification. It sums up to replace X by another variable $\Phi(X)$ that we will call an **observable**. It still belongs to the same vector space V , so there is a map $\Phi : V \rightarrow V$. If X is the actual value of the variables (and so of the state) then $\Phi(X)$ is the function which is estimated by the statistical procedure : the data are used to compute the value of the parameters of $\Phi(X)$. We make three general assumptions about Φ :

Definition of an observable :

- i) an observable is a linear map : $\Phi \in L(V; V)$
- 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 $\Phi(O) \subset O$.

Then using the linear chart Υ given by any basis, to Φ one can associate a map : $\widehat{\Phi} : H \rightarrow H :: \widehat{\Phi} = \Upsilon \circ \Phi \circ \Upsilon^{-1}$ and $\widehat{\Phi}$ is an operator on H . And conversely.

2.1 Primary observables

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

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

So the procedure can involve simultaneously several variables.

Proposition 6 To any primary observable Y_J is associated uniquely a self-adjoint, compact, trace-class operator \widehat{Y}_J on $H : Y_J = \Upsilon^{-1} \circ \widehat{Y}_J \circ \Upsilon$ such that the measure $Y_J(X)$ of the primary observable Y_J , if the system is in the state $X \in O$, is

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

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

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

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

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

ii) The primary observable Y_J is the map :

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

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

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

$\Upsilon(Y_J(X)) = \Upsilon\left(\sum_{j \in J} x_j e_j\right) = \sum_{j \in J} \langle \phi_j, \Upsilon(Y_J(X)) \rangle_H \varepsilon_j$
 $= \sum_{j \in J} \langle \phi_j, \Upsilon(X) \rangle_H \varepsilon_j$
 iii) $\forall X \in O : \Upsilon(Y_J(X)) \in H_J$ where H_J is the vector subspace of H spanned by $(\varepsilon_j)_{j \in J}$. It is finite dimensional, thus it is closed in H . There is a unique map (JCD Th.1111) :

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

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

\widehat{Y}_J is a Hilbert-Schmidt operator : take the Hilbertian basis $\tilde{\varepsilon}_i$ in H :

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

\widehat{Y}_J is a trace class operator with trace $\dim H_J$

$$\sum_{i \in I} \langle \widehat{Y}_J(\tilde{\varepsilon}_i), \tilde{\varepsilon}_i \rangle = \sum_{ij \in J} \langle \phi_j, \tilde{\varepsilon}_i \rangle \langle \varepsilon_j, \tilde{\varepsilon}_i \rangle = \sum_{j \in J} \langle \phi_j, \varepsilon_j \rangle = \sum_{j \in J} \delta_{jj} = \dim H_J$$

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

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

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

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

■

2.2 von Neumann algebra

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

We will enlarge the family of primary observables in several steps.

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

Proposition 7 *The operators $\left\{ \widehat{Y}_J \right\}_{J \subset I}$ are self-adjoint and commute*

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

\widehat{Y}_J has for eigen values :

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

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

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

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

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

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

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

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

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

So the operators commute. ■

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

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

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

Let us define the scalar product on W :

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

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

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

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

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

Proposition 8 \overline{A} is a commutative von Neumann algebra of $\mathcal{L}(H, H)$

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

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

The elements of $A = \text{Span} \left\{ \widehat{Y}_J \right\}_{J \subset I}$ that is of finite linear combination of \widehat{Y}_J commute
 $Y, Z \in \overline{A} \Rightarrow \exists (Y_n)_{n \in \mathbb{N}}, (Z_n)_{n \in \mathbb{N}} \in A^{\mathbb{N}} : Y_n \rightarrow_{n \rightarrow \infty} Y, Z_n \rightarrow_{n \rightarrow \infty} Z$
The composition is a continuous operation.
 $Y_n \circ Z_n = Z_n \circ Y_n \Rightarrow \lim (Y_n \circ Z_n) = \lim (Z_n \circ Y_n) = \lim Y_n \circ \lim Z_n = \lim Z_n \circ \lim Y_n = Z \circ Y = Y \circ Z$
So \overline{A} is commutative.
 \overline{A} is identical to the bicommutant of its projections, that is to \overline{A}'' (JCD Th.1189) ■

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

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

2.3 Secondary observables

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

1. A spectral measure defined on a measurable space E with σ -algebra σ_E and acting on the Hilbert space H is a map : $P : \sigma_E \rightarrow \mathcal{L}(H; H)$ such that :
 - i) $P(\varpi)$ is a projection
 - ii) $P(E) = \text{Id}$
 - iii) $\forall \psi \in H$ the map: $\varpi \rightarrow \langle P(\varpi)\psi, \psi \rangle_H = \|P(\varpi)\psi\|^2$ is a finite positive measure on (E, σ_E) .

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

- i) $\chi(\varpi)$ is a closed vector subspace of H
- ii) $\chi(E) = H$
- iii) $\forall \varpi, \varpi' \in \sigma_E, \varpi \cap \varpi' = \emptyset : \chi(\varpi) \cap \chi(\varpi') = \{0\}$

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

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

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

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

And conversely (JCD Th.1252), for any continuous normal operator $\hat{\Phi}$ on H

$$\hat{\Phi} \in \mathcal{L}(H; H) : \hat{\Phi} \circ \hat{\Phi}^* = \hat{\Phi}^* \circ \hat{\Phi}$$

there is a unique spectral measure P on $(\mathbb{R}, \sigma_{\mathbb{R}})$ such that $\hat{\Phi} = \int_{Sp(\hat{\Phi})} sP(s)$

where $Sp(\hat{\Phi}) \subset \mathbb{R}$ is the spectrum of $\hat{\Phi}$.

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

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

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

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

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

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

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

3. We will define :

A secondary observable is a linear map $\Phi \in L(V; V)$ valued in a finite dimensional vector subspace of V , such that $\hat{\Phi} = \Upsilon \circ \Phi \circ \Upsilon^{-1}$ is a normal operator.

Proposition 9 *Any secondary observable Φ is a compact, continuous map, its associated map $\hat{\Phi} = \Upsilon \circ \Phi \circ \Upsilon^{-1}$ is a compact, self-adjoint, Hilbert-Schmidt and trace class operator.*

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

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

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

$\Phi, \hat{\Phi}$ are compact and thus continuous (JCD Th.912).

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

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

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

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

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

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

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

iii) The observable reads :

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

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

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

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

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

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

We have :

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

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

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

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

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

The probability law $\widehat{\mu}_\psi : \sigma_{\mathbb{R}} \rightarrow \mathbb{R}$ reads : $\widehat{\mu}_\psi(\varpi) = \Pr(\lambda_p \in \varpi) = \frac{\|\widehat{\pi}_{H_p}(\psi)\|^2}{\|\psi\|^2}$

To sum up :

Proposition 10 For any primary or secondary observable Φ , there is a basis $(e_i)_{i \in I}$ of V , a compact, self-adjoint, Hilbert-Schmidt and trace class operator $\widehat{\Phi}$ on the associated Hilbert space H such that :

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

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

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

2.4 Efficiency of an observable

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

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

The error with the choice of Φ is :

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

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

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

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

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

So :

Proposition 11 For any secondary observable there is always a primary observable which is at least as efficient.

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

3 PROBABILITY

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

Usually neither the map Φ nor the basis $(e_i)_{i \in I}$ are explicit, even if they do exist through the choice of a statistical estimator. So we can look at the discrepancy $X - \Phi(X)$ from a different point of view : for a given, fixed, value of the state X , what is the uncertainty which stems from the choice of Φ among a large class of observables ? This sums up to assess the risk linked to the choice of a specification for the estimation of X .

3.1 Primary observables

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

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

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

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

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

If we see the choice of a finite subset $J \in 2^I$ as an event in a probabilist point of view, for a given $\psi \neq 0 \in H$ the quantity $\hat{Y}_J(\psi)$ is a random variable, with a distribution law $\hat{\mu}_\psi$.

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

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

So we have :

Proposition 12 *For any primary observable Φ , the value $\Phi(X)$ which is measured is an eigen vector of the operator Φ , and the probability to measure a value $\Phi(X)$ if the system is in the state X is :*

$$\Pr(\Phi(X) | X) = \frac{\|\hat{\Phi}(Y(X))\|_H^2}{\|Y(X)\|_H^2}$$

3.2 Secondary observables

For a secondary observable, as defined previously :

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

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

The vectors decompose as :

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

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

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

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

and

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

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

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

And we have :

Proposition 13 *For any secondary observable Φ , the value $\Phi(X)$ which is observed if the system is in the state X is a linear combination of eigen vectors X_p of Φ for the eigen value λ_p : $\Phi(X) = \sum_{p=1}^n \lambda_p X_p$*

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

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

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

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

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

This result can be illustrated by a simple example. Let us take a single continuous variable x, which takes its values in \mathbb{R} . It is clear that any physical measure will at best give a rational number $Y(x) \in \mathbb{Q}$ up to some scale. There are only countably many rational numbers for uncountably many real scalars. So the probability to get $Y(x) \in \mathbb{Q}$ should be zero. The simple fact of the

measure gives an incommensurable weight to rational numbers, implying that each of them has some small, but non null, probability to appear. In this case I can be assimilated to \mathbb{Q} , the subsets J are any finite collection of rational numbers.

4 CHANGE OF VARIABLES

In the conditions 1 we have noticed that, in the model, the variables could be defined over different connected domains. Actually one can go further and consider the change of variables, which leads to a theorem similar to the well known Wigner's theorem.

4.1 Fundamental theorem for a change of variables

A given system can be represented by different, related, variables. The two main cases are :

i) The variables are the coordinates of a geometric quantity (a vector, a tensor,...) expressed in some basis. According to the general Principle of Relativity, the state of the system shall not depend on the observers (those measuring the coordinates). So the coordinates shall follow the rules which are specified by their mathematical definitions, and the new values of the coordinates shall represent the same state of the system. We will see another example with interacting, indistinguishable systems.

ii) The variables are maps, depending on arguments which are themselves coordinates of some event : $X_k = X_k(\xi_1, \dots, \xi_{p_k})$. Similarly these coordinates ξ can change according to some rules, while the variable X_k represents the same event. A simple example that we will develop later on is a simple function of the time $X_k(t)$ such that the time t can be expressed in different units, or with different origin : $X_k(t)$ and $X'_k(t) = X_k(t + \theta)$ represent the same quantity.

We will summarize these features in the following :

Conditions 2 :

i) *The same system is represented by the variables $X = (X_1, \dots, X_N)$ and $X' = (X'_1, \dots, X'_{N'})$ which belong to the Fréchet vectors spaces V and V' , both infinite dimensional.*

ii) *X and X' take value in open subsets $O \subset V, O' \subset V'$*

iii) *There is a continuous bijective map $U : V \rightarrow V'$ such that X and $X'=U(X)$ represent the same state of the system*

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

From the first theorem, there are manifolds M, M' representing the states of the system, which can be embedded as open subsets Ω, Ω' of separable Hilbert

spaces H, H' , defined up to an isomorphism. U being a continuous bijective map, we can assume that there is a unique Hilbert space H . It implies that V and V' must have the same, infinite, dimension and one can take $V = V'$ but the open O, O' can be different.

Proposition 14 *Whenever a change of variables on a system meets the conditions 2 above, 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 $\Upsilon : V \rightarrow H$ associated to X .*

Proof. X and X' define two maps of the same manifold $\Xi : O \rightarrow \Omega \subset H, \Xi' : O' \rightarrow \Omega' \subset H$

Let V_0, V'_0 be the largest vector subspaces which contains O, O' . Because U is bijective, the basis of V_0, V'_0 must have the same cardinality. Let $(e_i)_{i \in I}, (e'_i)_{i \in I}$ be bases of V_0, V'_0 , $\varepsilon_i = \Xi(e_i), \varepsilon'_i = \Xi'(e'_i)$. We can define two linear maps :

$$\Upsilon : O \rightarrow \Omega :: \forall X \in O : \Upsilon(X) = \Xi(X)$$

$$\Upsilon' : O' \rightarrow \Omega' :: \forall X' \in O' : \Upsilon'(X') = \Xi'(X')$$

which are bijective and valued in the vector subspaces

$$H_0 = \text{Span}((\varepsilon_i)_{i \in I}), H'_0 = \text{Span}((\varepsilon'_i)_{i \in I})$$

Υ, Υ' are isometries :

$$\forall X_1, X_2 \in V_0 : \langle \Upsilon(X_1), \Upsilon(X_2) \rangle_H = \langle X_1, X_2 \rangle_V$$

$$\forall X'_1, X'_2 \in V'_0 : \langle \Upsilon'(X'_1), \Upsilon'(X'_2) \rangle_H = \langle X'_1, X'_2 \rangle_V$$

Thus U preserves the scalar product K_V on V :

$$\forall X_1, X_2 \in O : \Upsilon'(U(X_1)) = \Upsilon(X_1), \Upsilon'(U(X_2)) = \Upsilon(X_2)$$

$$\Rightarrow \langle \Upsilon'(U(X_1)), \Upsilon'(U(X_2)) \rangle_H = \langle U(X_1), U(X_2) \rangle_V = \langle \Upsilon(X_1), \Upsilon(X_2) \rangle_H = \langle X_1, X_2 \rangle_V$$

Let us define :

$$\widehat{U} : H_0 \rightarrow H_0 :: \widehat{U} = \Upsilon \circ U \circ \Upsilon^{-1} \text{ so } \widehat{U}(\Upsilon(X)) = \Upsilon(U(X))$$

\widehat{U} preserves the scalar product on H_0 :

$$\langle \widehat{U}(\Upsilon(X_1)), \widehat{U}(\Upsilon(X_2)) \rangle_H = \langle \Upsilon(U(X_1)), \Upsilon(U(X_2)) \rangle_H = \langle U(X_1), U(X_2) \rangle_V = \langle X_1, X_2 \rangle_V = \langle \Upsilon(X_1), \Upsilon(X_2) \rangle_H$$

It is continuous on the dense vector subspace H_0 so it can be extended to H (JCD Th.1003).

As seen in Proposition 1 starting from the basis $(\varepsilon_i)_{i \in I}$ of H_0 one can define a hermitian basis $(\tilde{\varepsilon}_i)_{i \in I}$ of H , an orthonormal basis $(\tilde{e}_i)_{i \in I}$ of V for the scalar product K_V

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

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

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

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

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

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

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

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

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

4.1.1 Observables

As \widehat{U} is unitary, it cannot be self adjoint or trace class. So it differs from an observable.

For any primary or secondary observable Φ there is a self-adjoint, Hilbert-Schmidt and trace class operator $\widehat{\Phi}$ on the associated Hilbert space H such that : $\widehat{\Phi} = \Upsilon \circ \Phi \circ \Upsilon^{-1}$. For the new variable the observable is $\Phi' = \Phi \circ U$ and it is associated to the operator :

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

4.1.2 Change of units

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

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

which implies for any single variable X_k : $\alpha_k = 1$. So the variables in the model should be dimensionless quantities. This is in agreement with the elementary rule that any formal theory should not depend on the units which are used. But also implies that there should be some "universal system of units" (based on a single quantity) in which all quantities of the theory can be measured.

4.2 Application to groups of transformations

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

4.2.1 Group representation

If there is a map $U : G \rightarrow \mathcal{GL}(V; V)$ 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 :

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

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

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

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

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

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

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

Unitary representations of Lie groups are well known, so whenever we have such result, it is possible to replace (H, \widehat{U}) by an equivalent representation on mathematical objects which are more manageable. Notably any topological group G endowed with a Haar measure has at least a unitary representation (the left or the right regular representation, acting on the arguments) on a Hilbert space of functions (JCD Th.1905). When G is a compact Lie group then any continuous unitary representation is completely reducible in the direct sum of orthogonal finite dimensional irreducible (JCD Th.1915). Then H itself is a Hilbert sum of finite dimensional Hilbert spaces.

The observables are parametrized by g :

$$\widehat{\Phi}(g) = \Upsilon \circ \Phi(1) \circ U(g) \circ \Upsilon^{-1} = \widehat{\Phi}(1) \circ \widehat{U}(g) \in \mathcal{L}(H; H)$$

but $(\widehat{\Phi}, H)$ is no longer a representation of the group G ($\widehat{\Phi}(1)$ is not invertible).

4.2.2 Gauge theories

The usual geometric representations, based on frames defined through a point and a set of vectors, such as in Galilean Geometry and Special Relativity, have been generalized by the formalism of fiber bundles, which encompasses also General Relativity, and is the foundation of gauge theories. So we will use this general formalism here (see JCD part VI, for the mathematical background). Gauge theories use abundantly group transformations, so they are a domain of choice to implement the previous results.

Fiber bundles are structures which are built over a manifold : they are a generalisation of the usual tangent bundle, which provides a larger flexibility about the choice of the objects living over the manifold. So the starting point

is a manifold M , which is usually the universe, and as we will not be concerned with cosmology, some bounded area delimited in time and space, which can be large. M is the container of the system. That it has the structure of a manifold comes from the fact that to locate any physical even x in M we need 4 scalars $\xi = (\xi^i)_{i=0}^3$: 3 for the spatial location and 1 for the time location. A chart is any mean used to relate x to ξ . It can be represented by a map : $\xi = \varphi_M(x)$. Whenever we have two sets of compatible charts φ_M, φ'_M , meaning that there is a continuous (and we can require that it is smooth) map such that $\xi' = \varphi'_M(x) = \chi(\xi) = \chi(\varphi_M(x))$, they define a structure of manifold. The additional assumptions about M define the geometry which is used in the model.

In Galilean Geometry M has the structure of a 3 dimensional affine euclidean space, the space coordinates are expressed in an orthonormal basis with a given origin, and the time is an independant coordinate expressed in \mathbb{R} . The evolution of the system can be followed in any spatial frame, and the change of observer is done through the group of displacements, independantly from the time. In Relativist Geometry observers, as any material body, follow a world line, which entails that the definition of the system itself is linked to an observer of reference : his present at a time $t = 0$ is a 3 dimensional riemannian submanifold $N(0)$ which is transported through the evolution of the system in a foliation $N(t)$, and the chart is given by a field of geodesics orthogonal to $N(t)$. The rules for a change of observers follow locally the well known formula. According to the Principle of Relativity whenever a physical quantity is represented by maps $X(\xi), X'(\xi')$ we shall have $X(\xi) = X'(\chi(\xi))$: the equations shall be covariant. We will see more on this topic in the next section. Here we will focus on other changes of observers, related to the quantities defined in the fiber bundle structure.

Let $\mathcal{P} = P(M, G, \pi_P)$ be a principal fiber bundle with base M , Lie group G , projection π_P and trivialization the map :

$$\varphi_P : M \times G \rightarrow \mathcal{P} :: p = \varphi_P(m, g)$$

$$\text{Any element of } \mathcal{P} \text{ is projected on } M : \forall w \in W : \pi_P(\varphi_P(m, w)) = m$$

Each observer defines a gauge : $\mathbf{p}(m) = \varphi_P(m, 1)$ and a change of observer is a change of trivialization, defined by a map : $\chi : M \rightarrow G$ and :

$$p = \varphi_P(m, g) = \tilde{\varphi}_P(m, \chi(m) \cdot g) \Leftrightarrow \tilde{g} = \chi(m) \cdot g \text{ (} \chi(m) \text{ acts on the left)}$$

$\chi(m)$ can be identical over M (the change is said to be global) or depend on m (the change is local).

A change of trivialization induces a change of gauge :

$$\mathbf{p}(m) = \varphi_P(m, 1) \rightarrow \tilde{\mathbf{p}}(m) = \tilde{\varphi}_P(m, 1) = \mathbf{p}(m) \cdot \chi(m)^{-1}$$

The principal fiber bundle is assumed to define the frames of reference used by an observer. In Galilean Geometry the group is $SO(3)$. In relativist geometry G is either $SO(3,1)$ or $SO(1,3)$ depending on the signature of the metric. For elementary particles M is endowed with an additional fiber bundle structure (representing the inner degrees of freedom) with the group $SU(3) \times SU(2) \times U(1)$.

Let W be a vector space and $[W, \rho]$ a representation of the group G . Then we have an associated vector bundle $\mathcal{W} = P[W, \rho]$ which has locally the structure

of a vector space. It is convenient to define a holonomic basis $(\mathbf{w}_i(m))_{i=1}^n$ from a basis $(w_i)_{i=1}^n$ of V and a gauge \mathbf{p} by : $\mathbf{w}_i(m) = (\mathbf{p}(m), w_i)$ then any vector of $P[W, \rho]$ reads :

$$\begin{aligned} w_m &= (\mathbf{p}(m), w) = (\mathbf{p}(m), \sum_{i=1}^n w^i w_i) = \sum_{i=1}^n w^i \mathbf{w}_i(m) \\ &\sim (\varphi_P(m, g), \rho^{-1}(w)) \end{aligned}$$

A change of trivialization on \mathcal{P} induces a change of holonomic basis on any associated vector bundle :

$$\begin{aligned} \mathbf{w}_i(m) &= (\mathbf{p}(m), w_i) \rightarrow \\ \tilde{\mathbf{w}}_i(m) &= (\tilde{\mathbf{p}}(m), w_i) = (\mathbf{p}(m) \cdot \chi(m)^{-1}, w_i) \sim (\mathbf{p}(m), \chi(m)^{-1}(w_i)) = \\ &\chi(m)^{-1} \mathbf{w}_i(m) \end{aligned}$$

so that the components of a vector in the holonomic basis change as :

$$\begin{aligned} w_m &= \sum_{i=1}^n w^i \mathbf{w}_i(m) = \sum_{i=1}^n \tilde{w}^i \tilde{\mathbf{w}}_i(m) = \sum_{i=1}^n \tilde{w}^i \chi(m)^{-1} \mathbf{w}_i(m) \\ \Rightarrow \tilde{w}^i &= [\chi(m)]_j^i w^j \end{aligned}$$

A section σ of \mathcal{W} is defined by a map : $\sigma : M \rightarrow W$

$$\sigma(m) = (\mathbf{p}(m), \sigma(m))$$

so that the components of a section change in a change of gauge as :

$$\begin{aligned} \sigma(m) &= \sum_{i=1}^n \sigma^i(m) \mathbf{w}_i(m) = \sum_{i=1}^n \tilde{\sigma}^i(m) \tilde{\mathbf{w}}_i(m) \\ \Leftrightarrow \tilde{\sigma}^i(m) &= [\chi(m)]_j^i \sigma^j(m) \end{aligned}$$

and this is the relation that, according to the Principle of Relativity, any physical quantity represented by such a section must follow.

This can be extended to any variable which is a derivative of a section, in the framework of the r-jet formalism. The r-jet extension $J^r \mathcal{W}$ of \mathcal{W} , is itself a vector bundle $J^r \mathcal{W}(M, J_0^r(\mathbb{R}^m; W), \pi^r)$. Its elements have the coordinated expressions : $Z = (z_{\alpha_1 \dots \alpha_s}^i, i = 1 \dots n, s = 0, \dots, r)$. The index i refers to the component in the vector bundle, the indices $\alpha_1, \dots, \alpha_s$ to the partial derivatives with respect to the coordinates in M . Whenever σ is a r continuously differentiable section on \mathcal{W} , then $z_{\alpha_1 \dots \alpha_s}^i = \frac{\partial^s \sigma^i}{\partial \xi^{\alpha_1} \dots \partial \xi^{\alpha_s}}$ takes the value of the partial derivatives of the components σ^i of σ in a holonomic basis of \mathcal{W} and σ defines a section $J^r \sigma$ of $\mathfrak{X}(J^r \mathcal{W})$. But conversely for a section of $\mathfrak{X}(J^r \mathcal{W})$ the components $z_{\alpha_1 \dots \alpha_s}^i$ are independant variables.

The set of sections on \mathcal{W} , denoted $\mathfrak{X}(\mathcal{W})$, is an infinite dimensional vector space, as well as the space $\mathfrak{X}(J^r \mathcal{W})$. It is easy (take a global change of gauge) to see that these vector spaces are representations of the group G . Whenever there is a scalar product (bilinear symmetric or hermitian two form) $\langle \rangle$ on a vector space W , so that (W, ρ) is a unitary representation of the group G : $\langle \rho(g)v, \rho(g)v' \rangle = \langle v, v' \rangle$, the scalar product can be extended on the associated vector bundle $P[W, \rho]$:

$$\langle (\mathbf{p}(m), w), (\mathbf{p}(m), w') \rangle_{P[W, \rho]} = \langle w, w' \rangle_W$$

If this scalar product is definite positive, with any measure μ on the manifold M (usually the Lebesgue measure associated to a volume form as in the relativist context), one can define the spaces of integrables sections $L^q(M, \mu, \mathcal{W}), L^q(M, \mu, J^r \mathcal{W})$ of $\mathcal{W}, J^r \mathcal{W}$ (by taking the integral of the norm pointwise). In all cases they are separable Fréchet spaces, and for $q = 2$ they are Hilbert spaces, and unitary

representation of the group G . Notice that the signature of the scalar product is that of the product defined on \mathcal{W} , the metric on M (such as a lorentzian metric) is not involved.

According to the Principle of Relativity, any physical quantity represented in a vector bundle must then be equivariant in a change of gauge (corresponding to a change of observer) : its components must change according to an action which is parametrized by the group G . This condition imposes constraints in the specification of the lagrangian, but also helps us on the possible choice of the maps representing sections on the vector bundle. When there is a definite positive scalar product, it seems legitimate (at least when M is a relatively compact open in a larger manifold) to assume that sections representing physical quantities in these vector bundles are integrable, and notably that they belong to the spaces $L^2(M, \mu, \mathcal{W})$, $L^2(M, \mu, J^r\mathcal{W})$. Because they provide unitary representations of G , following the previous theorem, it seems natural to assume that the Hilbert space H is just L^2 itself. Then all the previous results (observables, probability,...) can be implemented directly on the space of variables. Moreover, when G is compact, then the representation is reducible into finite dimensional representations, which is a step towards quantization. Conversely this leads to find either a definite positive scalar product on \mathcal{W} , or to restrict the maps to some subspace for which the scalar product is positive.

In gauge theories, the force fields (gravitational field, electromagnetic, weak and strong fields) are represented by principal connections on a principal fiber bundle. Because of restrictions on the lagrangian (notably that the potential cannot appear explicitly) and to keep a vectorial quantity, the key variable used is the strength of the field, which is a 2-form \mathcal{F} on M valued in the Lie algebra T_1G of the group G , and can be seen as the exterior covariant derivative of the potential of the connection. For instance for the electromagnetic field \mathcal{F} is a scalar valued 2-form. This 2-form lives in the adjoint bundle to the principal bundle $P(M, G, \pi_P)$, which is the associated vector bundle $P[T_1G, Ad]$ where Ad is the adjoint map : $Ad : G \times T_1G \rightarrow \mathcal{L}(T_1G; T_1G)$. Bisymmetric, Ad invariant bilinear form on a Lie algebra are usually found through the Killing form, which is definite negative for compact and semi-simple Lie algebra (JCD Th.1737).

Thus, if following the previous theorem, we assume that \mathcal{F} belongs to L^2 , G is a compact group, and the representation is the sum of finite dimensional representations. Indeed for elementary particles this is the case ($SU(3) \times SU(2) \times U(1)$ is compact), and there are a finite number of fields, each corresponding to one representation.

For the gravitational field however we know the group, it is not compact, and the Killing form is degenerate. So either we look for another key variable (usually the scalar curvature) or we restrict \mathcal{F} to belong to a subspace for which the scalar product is definite positive. But we are here on the, difficult, topic of quantum gravity...

For material bodies the formalism of vector bundle can still be implemented : the trajectories are just path over a fiber bundle, parametrized by the time. The issue is to define the vector space W . Usually the translation and the rotation parts of the motion are considered separately, but the components of the motion are not independant in the relativist context : indeed the boost, in the matrix which defines the change of gauge, is related to the spatial speed. Moreover the instantaneous rotation is more physically related to the inertial momentum than to the geometric rotation and has no precise meaning for particles. So it is necessary to review the way to define the motion and momenta, and this leads, through Clifford algebras, to the spinors. It is then possible, by using the previous theorems, to retrieve a clear understanding of the spin and of the concept of particles fields in the QM context.

But this is not here the place to tell more on these subjects, which have been reviewed in my paper on Particles and fields (2014).

4.3 Fourier transform

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

$$\begin{aligned} & \text{If } \varphi \in L^2(G, \mu, \mathbb{C}) \cap L^1(G, \mu, \mathbb{C}) : \\ & \mathcal{F}(\varphi)(\vartheta) = \int_G \varphi(g) \overline{\vartheta(g)} \mu(g) \\ & \mathcal{F}^*(h)(g) = \int_{\widehat{G}} h(\vartheta) \vartheta(g) \nu(\vartheta) \text{ for a unique Haar measure } \nu \text{ on } \widehat{G} \text{ and } \mathcal{F}^* = \mathcal{F}^{-1} \end{aligned}$$

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

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

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

4.4 One parameter groups

An important case, related to the previous one, is when the variables X depend on a scalar real argument, and the model is such that $X(t), X'(t') = X(t + \theta)$, with any fixed θ , represent the same state. The associated operator is parametrized by a scalar and we have a map :

$$\begin{aligned}\widehat{U} : \mathbb{R}_+ &\rightarrow \mathcal{GL}(H, H) \text{ such that :} \\ \widehat{U}(t + t') &= \widehat{U}(t) \circ \widehat{U}(t') \\ \widehat{U}(0) &= Id\end{aligned}$$

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

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

Because $\widehat{U}(t)$ is unitary S is anti-hermitian :

$$\begin{aligned}\langle \widehat{U}(t) \psi, \widehat{U}(t) \psi' \rangle_H &= \langle \psi, \psi' \rangle_H \\ \Rightarrow \left\langle \frac{d}{dt} \widehat{U}(t) \psi, \widehat{U}(t) \psi' \right\rangle_H + \left\langle \widehat{U}(t) \psi, \frac{d}{dt} \widehat{U}(t) \psi' \right\rangle_H &= 0 \Rightarrow S = -S^*\end{aligned}$$

S is normal and has a spectral resolution P :

$$S = \int_{Sp(S)} s P(s)$$

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

We will see in the next section a striking application of this case.

5 THE EVOLUTION OF THE SYSTEM

In many models involving maps, the variables X_k are functions of the time t , which represents the evolution of the system. So this is a privileged argument of the functions, which is understood in one of two different ways :

A) t is a parameter used to identify a temporal location. It is one of the coordinates of the manifold M in which the system stays. This is the usual case in Analytical Mechanics or in Theories of Fields. It is clear that the geometry used in the model is then crucial.

B) t is a parameter used to measure the duration of a phenomenon, usually the time elapsed since some specific event. This is the case in more specific models, such as in Engineering. Then usually the geometry is galilean, time is independant from the spatial coordinates.

In the case A, if the geometry is Galilean then the time is independent from the spatial coordinates for any observer and one can consider a change of coordinates such as : $t' = t + \theta$ with any constant θ . The variables X, X' such that $X'(t') = X(t + \theta)$ represent the same system, we have a change of variable where the map U is parametrized by the abelian group $(\mathbb{R}, +)$ and, according to the previous theorem, we have a one parameter group. In relativist geometry the relation between coordinates ξ, ξ' is more complicated in a general change of observer, and does not reflect the simple action of a group. However, coming back to the definition of a system through a foliation as seen in the subsection above, one can see that we would have the same system by labelling the initial submanifold $N(0)$ by θ rather than 0. More generally in a change of coordinates $\xi' = \xi + \theta$ with a fixed 4 vector θ the maps $X(\xi)$ and $X'(\xi') = X(\xi + \theta)$ represent the same state of the system. In a fixed affine change of coordinates on any manifold the usual rules which involve the jacobian show that any quantity defined on the tangent bundle stays unchanged. This is the same reasoning which is used to prove that the coordinates cannot be explicitly present (as independent variables) in a lagrangian. So, using this specific change of coordinates one can implement the previous theorem also in the relativist context.

In the case B it is clear that the origin of time matters. So far we have not made any additional assumption about the model but, considering its evolution, we can go a bit further. First when one considers the evolution of a system, as it is conceived here, one should account for its initial state, that is for the value $X(0)$ of the variables at some origin 0 of time. It is part of the definition of the system itself, thus one could impose to X the condition that it belongs to V and moreover that $X(0)$ takes a given value. In other words there is a partition of the set of possible maps V into subsets of maps, each of them being characterized by the value of X at $t = 0$. Second, whatever the laws that the variables follow, it is legitimate to assume that their value $X(t)$ at t depends on their value $X(0)$ at the origin. In an ideal model the physicist expects that to each initial condition corresponds a unique solution. We will be a bit less demanding and we will summarize these features in the following conditions, which are related to the case B but, as we will see, apply also to the case A :

Conditions 3 :

The model representing the system and its evolution is such that :

i) V is an infinite dimensional Fréchet space V of maps : $X = (X_k)_{k=1}^N :: R \rightarrow E$ where R is an open subset of \mathbb{R} and E a normed vector space

ii) $\forall t \in R$ the evaluation map : $\mathcal{E}(t) : V \rightarrow E : \mathcal{E}(t) X = X(t)$ is continuous

iii) *The laws for the evolution of the system are such that the variables $(X_k)_{k=1}^N$, which define the possible states considered for the system (that we call the admissible states), must meet the conditions :*

- they belong to an open subset O of V

- the initial state of the system, defined at $t = 0 \in R$, belongs to an open subset A of E

- for any solutions X, X' belonging to O if the set $\varpi = \{t, X(t) = X(t')\}$ has a non null Lebesgue measure then $X = X'$.

The last condition means that the system is semi determinist : to the same initial conditions can correspond several different solutions, but if two solutions are equal on some interval then they are equal almost everywhere.

The condition ii) is rather technical and should be usually met. Practically it involves some relation between the semi-norms on V and the norm on E (this is why we need a norm on E) : when two variables X, X' are close in V , then their values $X(t), X'(t)$ must be close everywhere. More pecisely. $\mathcal{E}(t)$ is linear so the continuity can be checked at $X = 0$ and reads:

$\forall t \in R, \forall X \in O : \forall \varepsilon > 0, \exists \eta : d(X, 0)_V < \eta \Rightarrow \|X(t)\|_E < \varepsilon$ where d is the metric on V

In all usual cases (such as L^p spaces or spaces of differentiable functions) $d(X, 0)_V \rightarrow 0 \Rightarrow \forall t \in R : \|X(t)\|_E \rightarrow 0$ and the condition ii) is met, but this is not a general result.

Notice that :

- the variables X can depend on any other arguments besides t as previously
- E can be infinite dimensional but must be normed
- no continuity condition is imposed on X .

5.1 Fundamental theorems for the evolution of a system

If the model meets the conditions 3 then it meets the conditions of the proposition 1, that is : there is a separable, infinite dimensional, Hilbert space H , defined up to isomorphism, such that the states (admissible or not) \mathcal{S} belonging to O can be embedded as an open subset $\Omega \subset H$ which contains 0 and a convex subset. Moreover to any basis of V is associated a bijective linear map $\Upsilon : V \rightarrow H$.

Proposition 15 *If the conditions 3 are met, then there are :*

- i) a Hilbert space F , an open subset $\tilde{A} \subset F$
- ii) a map : $\Theta : R \rightarrow \mathcal{L}(F; F)$ such that $\Theta(t)$ is unitary and, for the admissible states $X \in O \subset V$:

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

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

- iii) an isometry : $\hat{\mathcal{E}} : R \rightarrow \mathcal{L}(H; F)$ such that for the admissible states $X \in O \subset V$:

$$\forall X \in O : \hat{\mathcal{E}}(t)\Upsilon(X) = X(t)$$

where H is the Hilbert space and Υ is the linear chart associated to X and any basis of V

Proof. i) Define the equivalence relation on V :

$$\mathcal{R} : X \sim X' \Leftrightarrow X(t) = X'(t) \text{ almost everywhere for } t \in R$$

and take the quotient space V/\mathcal{R} , then the set of admissible states is a set

\tilde{O} such that :

$$\tilde{O} \in O \subset V$$

$$\forall X \in \tilde{O} : X(0) \in A$$

$$\forall X, X' \in \tilde{O}, \forall t \in R : X(t) = X'(t) \Rightarrow X = X'$$

ii) Define :

$$\forall t \in R : \tilde{F}(t) = \{X(t), X \in \tilde{O}\} \text{ thus } \tilde{F}(0) = A$$

A is a subset of E . There are families of independant vectors belonging to A , and a largest family $(f_j)_{j \in J}$ of independant vectors. It generates a vector space $F(0)$ which is a vector subspace of E , containing A .

$$\forall u \in F(0) : \exists (x_j)_{j \in J} \in \mathbb{R}_0^J : u = \sum_{j \in J} x_j f_j$$

The map :

$$\tilde{\Theta}(t) : \tilde{F}(0) \rightarrow \tilde{F}(t) :: \tilde{\Theta}(t) u = \mathcal{E}(t) \circ \mathcal{E}(0)^{-1} u$$

is bijective and continuous

The set $F(t) = \tilde{\Theta}(t) F(0) \subset E$ is well defined by linearity :

$$\tilde{\Theta}(t) \left(\sum_{j \in J} x_j f_j \right) = \sum_{j \in J} x_j \tilde{\Theta}(t)(f_j)$$

The map $\tilde{\Theta}(t) : F(0) \rightarrow F(t)$ is linear, bijective, continuous on an open subset A , thus continuous, and the spaces $F(t)$ are isomorphic, vector subspaces of E , containing $\tilde{F}(t)$.

Define $(\varphi_j)_{j \in J}$ the largest family of independant vectors of $\{\tilde{\Theta}(t)(f_j), t \in R\}$.

This is a family of independant vectors of E , which generates a subspace \tilde{F} of E , containing each of the $F(t)$ and thus each of the $\tilde{F}(t)$. Moreover each of the φ_j is the image of a unique vector f_j for some $t_j \in R$.

The map $\tilde{\Theta}(t)$ is then a continuous linear map $\tilde{\Theta}(t) \in \mathcal{L}(\tilde{F}; \tilde{F})$

iii) The conditions of proposition 1 are met for O and V , so there are a Hilbert space H and a linear map $\Upsilon : O \rightarrow \Omega$

Each of the φ_j is the image of a unique vector f_j for some $t \in R$, and thus there is a uniquely defined family $(X_j)_{j \in J}$ of \tilde{O} such that $X_j(t_j) = \varphi_j$.

Define on \tilde{F} the bilinear symmetric definite positive form with coefficients :

$$\langle \varphi_j, \varphi_k \rangle_{\tilde{F}} = K_V \left(\mathcal{E}(t_j)^{-1} \varphi_j, \mathcal{E}(t_k)^{-1} \varphi_k \right) = \left\langle \Upsilon \mathcal{E}(t_j)^{-1} \varphi_j, \Upsilon \mathcal{E}(t_k)^{-1} \varphi_k \right\rangle_H = \langle X_j, X_k \rangle_H$$

By the Gram-Schmidt procedure we can build an orthonormal basis $(\tilde{\varphi}_j)_{j \in J}$ of \tilde{F} : $\tilde{F} = \text{Span}(\tilde{\varphi}_j)_{j \in J}$ and the Hilbert vector space : $F = \left\{ \sum_{j \in J} \tilde{x}_j \tilde{\varphi}_j, (\tilde{x}_j)_{j \in J} \in \ell^2(J) \right\}$

which is a vector space containing \tilde{F} (but is not necessarily contained in E).

iv) The map $\tilde{\Theta}(t) \in \mathcal{L}(\tilde{F}; \tilde{F})$ is a linear homomorphism, \tilde{F} is dense in F , thus $\tilde{\Theta}$ can be extended to a continuous operator $\Theta \in \mathcal{L}(F; F)$ (JCD Th.1003).

$\tilde{\Theta}(t)$ is unitary on \tilde{F} : $\langle u, v \rangle_{\tilde{F}} = K_V \left(\mathcal{E}(0)^{-1} u, \mathcal{E}(0)^{-1} v \right)$ so $\Theta(t)$ is unitary on F .

iv) Define the map :

$$\widehat{\mathcal{E}}(t) : \Omega \rightarrow F :: \widehat{\mathcal{E}}(t) \Upsilon(X) = X(t)$$

where $\Omega \subseteq H$ is the open associated to V and O .

For $X \in \widetilde{O}$:

$$\widehat{\mathcal{E}}(t) \Upsilon(X) = X(t) = \widetilde{\Theta}(t) X = \mathcal{E}(t) \circ \mathcal{E}(0)^{-1} X$$

$$\widehat{\mathcal{E}}(t) = \mathcal{E}(t) \circ \mathcal{E}(0)^{-1} \circ \Upsilon^{-1}$$

$\widehat{\mathcal{E}}(t)$ is linear, continuous, bijective on Ω , it is an isometry :

$$\left\langle \widehat{\mathcal{E}}(t) \psi, \widehat{\mathcal{E}}(t) \psi' \right\rangle_F = \langle X(t), X'(t) \rangle_F = \langle \Upsilon X, \Upsilon X' \rangle_H = \langle \psi, \psi' \rangle_H$$

v) $A = \widetilde{F}(0)$ is an open subset of $F(0)$, which is itself an open vector subspace of F . Thus A can be embedded as an open subset \widetilde{A} of F . ■

The existence of a Hilbert space structure on E was not obvious. As a consequence the model is determinist, up to the equivalence between maps almost everywhere equal. But the operator $\Theta(t)$ depends on t and not necessarily continuously, so the problem is not necessarily well posed. However 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 $m \times n$ matrix $[x]$.

We see now that in the case A we get a stronger result.

Proposition 16 *For any system represented by a model where :*

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

ii) *belonging to an open subset O of an infinite dimensional Fréchet space V*

iii) *$\forall t \in R$ the evaluation map : $E(t) : V \rightarrow E : E(t)X = X(t)$ is continuous*

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

then :

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

$$\mathcal{E}(t) = \mathcal{E}(0) \circ \exp tS$$

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

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

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

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

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

vii) *The maps X are smooth and :*

$$\frac{d}{ds} X(s) \Big|_{s=t} = \widetilde{S} X(t)$$

Proof. i) We have a change of variables U depending on a parameter $\theta \in \mathbb{R}$ which reads with the evaluation map : $\mathcal{E} : \mathbb{R} \times V \rightarrow F :: \mathcal{E}(t) X = X(t)$:

$\forall t, \theta \in \mathbb{R} : \mathcal{E}(t)(U(\theta)X) = \mathcal{E}(t+\theta)(X) \Leftrightarrow \mathcal{E}(t)U(\theta) = \mathcal{E}(t+\theta) = \mathcal{E}(\theta)U(t)$:

U defines a one parameter group of linear operators:

$$U(\theta + \theta')X(t) = X(t + \theta + \theta') = U(\theta) \circ U(\theta')X(t)$$

$$U(0)X(t) = X(t)$$

It is obviously continuous at $\theta = 0$ so it is continuous.

ii) The conditions of proposition 1 are met, so there are a Hilbert space H, a linear chart Υ , and $\widehat{U} : \mathbb{R} \rightarrow \mathcal{L}(H; H)$ such that $\widehat{U}(\theta)$ is linear, bijective, unitary :

$$\forall X \in O : \widehat{U}(\theta)(\Upsilon(X)) = \Upsilon(U(\theta)(X))$$

$$\widehat{U}(\theta + \theta') = \Upsilon \circ U(\theta + \theta') \circ \Upsilon^{-1} = \Upsilon \circ U(\theta) \circ U(\theta') \circ \Upsilon^{-1} = \Upsilon \circ U(\theta) \circ \Upsilon^{-1} \circ \Upsilon \circ U(\theta') \circ \Upsilon^{-1} = \widehat{U}(\theta) \circ \widehat{U}(\theta')$$

$$\widehat{U}(0) = \Upsilon \circ U(0) \circ \Upsilon^{-1} = Id$$

The map $\widehat{U} : \mathbb{R} \rightarrow \mathcal{L}(H; H)$ is uniformly continuous with respect to θ , it defines a one parameter group of unitary operators. So there is an anti-hermitian operator \widehat{S} with spectral resolution P such that :

$$\widehat{U}(\theta) = \sum_{n=0}^{\infty} \frac{\theta^n}{n!} \widehat{S}^n = \exp \theta \widehat{S}$$

$$\frac{d}{ds} \widehat{U}(s) |_{\theta=s} = \left(\exp \theta \widehat{S} \right) \circ \widehat{S}$$

$$\widehat{S} = \int_{Sp(S)} sP(s)$$

$$\left\| \widehat{U}(\theta) \right\| = 1 \leq \exp \left\| \theta \widehat{S} \right\|$$

iii) $S = \Upsilon^{-1} \circ \widehat{S} \circ \Upsilon$ is a continuous map on the largest vector subspace V_0 of V which contains O, which is a normed vector space with the norm induced by the positive kernel.

$$\|S\| \leq \|\Upsilon^{-1}\| \|\widehat{S}\| \|\Upsilon\| = \|\widehat{S}\| \text{ because } \Upsilon \text{ is an isometry.}$$

So the series $\sum_{n=0}^{\infty} \frac{\theta^n}{n!} S^n$ converges in V_0 and :

$$U(\theta) = \Upsilon^{-1} \circ \widehat{U}(\theta) \circ \Upsilon = \sum_{n=0}^{\infty} \frac{\theta^n}{n!} S^n = \exp \theta S$$

$$\forall \theta, t \in \mathbb{R} : U(\theta)X(t) = X(t + \theta) = (\exp \theta S)X(t)$$

$$\mathcal{E}(t) \exp \theta S = \mathcal{E}(t + \theta)$$

Exchange θ, t and take $\theta = 0$:

$$\mathcal{E}(\theta) \exp tS = \mathcal{E}(t + \theta)$$

$$\mathcal{E}(0) \exp tS = \mathcal{E}(t) \in \mathcal{L}(V; E)$$

which reads :

$$\forall t \in \mathbb{R} : U(t)X(0) = X(t) = (\exp tS)X(0)$$

(U, V_0) is a continuous representation of $(\mathbb{R}, +)$, U is smooth and X is smooth

:

$$\frac{d}{ds} U(s)X(0) |_{s=t} = \frac{d}{ds} X(s) |_{s=t} = SX(t)$$

$$\Leftrightarrow \frac{d}{ds} \mathcal{E}(s) |_{s=t} = S\mathcal{E}(t)$$

The same result holds whatever the size of O in V, so S is defined over V.

iv) The set : $F(t) = \{X(t), X \in V\}$ is a vector subspace of E.

Each map is fully defined by its value at one point :

$$\forall t \in \mathbb{R} : X(t) = (\exp tS \circ X)(0)$$

$$X(t) = X'(t) \Rightarrow \forall \theta : X(t + \theta) = X'(t + \theta) \Leftrightarrow X = X'$$

So the conditions 2 are met.

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

The map $\Theta(\theta) : F \rightarrow F$ defines a one parameter group, so it has an infinitesimal generator $\tilde{S} \in \mathcal{L}(F; F) : \Theta(\theta) = \exp \theta \tilde{S}$ and because $\Theta(\theta)$ is unitary \tilde{S} is anti-hermitian.

$$\frac{d}{ds} \Theta(s) X(0) |_{s=t} = \frac{d}{ds} X(s) |_{s=t} = \tilde{S} X(t) \quad \blacksquare$$

As a consequence such a model is necessarily determinist, and the system is represented by smooth maps whose evolution is given by a unique operator. And it is clear that the conditions 3 are then met, so the case A is actually a special case of B. Notice that, even if X was not assumed to be continuous, smoothness is a necessary result. This result can seem surprising, but actually the basic assumption about a translation in time reflects the assumption that the laws of evolution are smooth, and as the consequence the variables depend smoothly on the time. And conversely this implies that, whenever there is some discontinuity in the evolution of the system, the conditions above cannot hold : time has a specific meaning, related to a change in the environment.

5.2 Observables

When a system is studied through its evolution, the observables can be considered from two different points of view :

- in the movie way : the estimation of the parameters is done at the end of the period considered, from a batch of data corresponding to several times (which are not necessarily the same for all variables). So this is the map X which is estimated through an observable $X \rightarrow \Phi(X)$.

- in the picture way : the estimation is done at different times (the same for all the variables which are measured). So there are the values X(t) which are estimated. Then the estimation of X(t) is given by $\varphi(X(t)) = \varphi(\mathcal{E}(t)X)$, with φ a linear map from E to a finite dimensional vector space, which usually does not depend on t (the specification stays the same).

In the best scenario the two methods should give the same result, which reads :

$$\varphi(\mathcal{E}(t)X) = \mathcal{E}(t)(\Phi X) \Leftrightarrow \varphi = \mathcal{E}(t) \circ \Phi \circ \mathcal{E}(t)^{-1}$$

But usually, when it is possible, the first way gives a better statistical estimation.

5.3 Phases Transitions

There is a large class of problems which involve transitions in the evolution of a system. They do not involve the maps X, which belong to the same family as above, but the values X(t) which are taken over a period of time in some vector space E. There are distinct subsets of E, that we will call phases (to avoid any confusion with states which involves the map X), between which the state of the

system goes during its evolution, such as the transition solid / gas or between magnetic states. The questions which arise are then : what are the conditions, about the initial conditions or the maps X, for the occurrence of such an event ? Can we forecast the time at which such event takes place ?

Staying in the general model meeting the conditions 3, the first issue is the definition of the phases. The general idea is that they are significantly different states, which can be formalized by : the set $\{X(t), t \in R, X \in O\}$ is disconnected, it comprises two disjoint subsets E_1, E_2 closed in E .

If the maps $X : R \rightarrow F$ are continuous and R is an interval of \mathbb{R} (as we will assume) then the image $X(R)$ is connected, so the maps X cannot be continuous, and we cannot be in the case A, but we can be in the case B (a fact which, in itself, is interesting : a change of phase needs a change of period). This is a difficult but also very common issue : in the real life such discontinuous evolutions are the rule. However even if totally discontinuous maps exist, they are strange mathematical objects. Usually discontinuities are assumed to 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 $(\theta_\alpha)_{\alpha \in A}, A \subset \mathbb{N}$. If $X(0) \in E_1$ then the odd transition points $\theta_{2\alpha+1}$ mark a transition $E_1 \rightarrow E_2$ and the opposite for the even points $\theta_{2\alpha}$.

If the conditions 3 are met then Θ is continuous except in $(\theta_\alpha)_{\alpha \in A}$, the transition points do not depend on the initial state $X(0)$, but the phase on each segment does. Then it is legitimate to assume that there is some probability law which rules the occurrence of a transition. We will consider two cases.

The simplest assumption is that the probability of the occurrence of a transition at any time t is constant. Then it depends only on the cumulated lengths of the periods $T_1 = \sum_{\alpha=0} [\theta_{2\alpha}, \theta_{2\alpha+1}], T_2 = \sum_{\alpha=0} [\theta_{2\alpha+1}, \theta_{2\alpha+2}]$ respectively.

Let us assume that $X(0) \in E_1$ then the changes $E_1 \rightarrow E_2$ occur for $t = \theta_{2\alpha+1}$, the probability of transitions read :

$$\Pr(X(t + \varepsilon) \in E_2 | X(t) \in E_1) = \Pr(\exists \alpha \in \mathbb{N} : t + \varepsilon \in [\theta_{2\alpha+1}, \theta_{2\alpha+2}]) \\ = T_2 / (T_1 + T_2)$$

$$\Pr(X(t + \varepsilon) \in E_1 | X(t) \in E_2) = \Pr(\exists \alpha \in \mathbb{N} : t + \varepsilon \in [\theta_{2\alpha}, \theta_{2\alpha+1}]) \\ = T_1 / (T_1 + T_2)$$

$$\Pr(X(t) \in E_1) = T_1 / [R]; \Pr(X(t) \in E_2) = T_2 / [R]$$

The probability of a transition at t is : $T_2 / (T_1 + T_2) \times T_1 / (T_1 + T_2) + T_1 / (T_1 + T_2) \times T_2 / (T_1 + T_2) = 2T_1T_2 / (T_1 + T_2)^2$. It does not depend of the initial phase, and depends only on Θ .

However usually the probability of a transition depends on the values of the variables. The phases are themselves characterized by the value of $X(t)$, so a sensible assumption is that the probability of a transition increases with

the proximity of the other phase . Using the Hilbert space structure of F it is possible to address practically this case.

If E_1, E_2 are *closed convex subsets* of F , which is a Hilbert space, there is a unique map : $\pi_1 : F \rightarrow E_1$. The vector $\pi_1(x)$ is the unique $y \in E_1$ such that $\|x - y\|_F$ is minimum. The map π_1 is continuous and $\pi_1^2 = \pi_1$. And similarly for E_2 .

The quantity $\|X(t) - \pi_1(X(t))\|_F + \|X(t) - \pi_2(X(t))\|_F =$ the distance to the other subset than where $X(t)$ lies, so one can assume that the probability of a transition at t is : $f(\|X(t) - \pi_1(X(t))\|_F + \|X(t) - \pi_2(X(t))\|_F)$ where $f : \mathbb{R} \rightarrow [0, 1]$ is a probability density. The probability of a transition depends only on the state at t , but one cannot assume that the transitions points θ_α do not depend on X .

The result holds if E_1, E_2 are closed *vector subspaces* of F such that $E_1 \cap E_2 = \{0\}$. Then

$$X(t) = \pi_1(X(t)) + \pi_2(X(t))$$

$$\text{and } \|X(t)\|^2 = \|\pi_1(X(t))\|^2 + \|\pi_2(X(t))\|^2$$

$\frac{\|\pi_1(X(t))\|^2}{\|X(t)\|^2}$ can be interpreted as the probability that the system at t is in the phase E_1 .

One important application is forecasting a transition for a given map X . From the measure of $X(t)$ one can compute for each t the quantity $Y(t) = \|X(t) - \pi_1(X(t))\|_F + \|X(t) - \pi_2(X(t))\|_F$ and, if we know f , we have the probability of a transition at t . The practical problem is then to estimate f from the measure of Y over a past period $[0, T]$. A very simple, non parametric, estimator can be built as follows. From a set of data $\{Y(t), t \in [0, T]\}$ one can easily compute the function : $G : \mathbb{R} \rightarrow [0, T]$ where $G(y)$ is the total duration of the periods when $Y(t) \geq y$. This is a decreasing curve, from T to 0 when y goes from 0 to $\text{Max}(Y)$. The probability of a transition at any given time when $Y(t) \geq y$ is $p(y)$. One can compute the number of transitions $n(y)$ which have occurred when $Y(t) \geq y$, then the estimation $\hat{p}(y)$ of $p(y)$ is $\hat{p}(y) = \frac{n(y)}{G(y)}$.

6 INTERACTING SYSTEMS

6.1 Representation of interacting systems

In the propositions above no assumption has been done about the interaction with "exterior" variables. If the values of some variables are given (for instance to study the impact of external factors with the system) then they shall be fully integrated into the set of variables, at the same footing as the others.

A special case occurs when one considers two systems S_1, S_2 , which are similarly represented, meaning that that we have the same kind of variables, defined as identical mathematical objects and related significance. To account for the interactions between the two systems the models are of the form :

$$\begin{array}{ccc}
\lrcorner & S_1 & \ulcorner \\
X_1 & & Z_1 \\
V_1 & \times & W_1 \\
& \downarrow \Upsilon_1 & \\
& \psi_1 & \\
& H_1 & \\
& & \lrcorner \\
& & S_{1+2} \\
& & \ulcorner \\
& & X_1 & & X_2 \\
& & V_1 & \times & V_2 \\
& & \psi_1 & & \psi_2 \\
& & H_1 & \times & H_2
\end{array}$$

X_1, X_2 are the variables (as above X denotes collectively a set of variables) characteristic of the systems S_1, S_2 , and Z_1, Z_2 are variables representing the interactions. Usually these variables are difficult to measure and to handle. One can consider the system S_{1+2} with the direct product $X_1 \times X_2$, but doing so we obviously miss the interactions Z_1, Z_2 .

We see now how it is possible to build a simpler model which keeps the features of S_1, S_2 and accounts for their interactions.

We consider the models without interactions (so with only X_1, X_2) and we assume that they meet the conditions 1. For each model $S_k, k = 1, 2$ there are a linear map : $\Upsilon_k : V_k \rightarrow H_k :: \Upsilon_k(X_k) = \psi_k = \sum_{i \in I_k} \langle \phi_{ki}, \psi_k \rangle e_{ki}$
a positive kernel : $K_k : V_k \times V_k \rightarrow \mathbb{R}$

Let us denote S the new model. Its variables will be collectively denoted Y , valued in a Fréchet vector space V' . There will be another Hilbert space H' , and a linear map $\Upsilon' : V' \rightarrow H'$ similarly defined. As we have the choice of the model, we will impose some properties to Y and V' in order to underline both that they come from S_1, S_2 and that they are interacting.

Conditions 4 :

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

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

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

iii) The positive kernel is a defining feature of the models, so we want a positive kernel K' of (V', Υ') such that :

$$\forall X_1, X'_1 \in V_1, \forall X_2, X'_2 \in V_2 :$$

$$K'(\Phi(X_1, X_2), \Phi(X'_1, X'_2)) = K_1(X_1, X'_1) \times K_2(X_2, X'_2)$$

We will prove the following :

Proposition 17 *Whenever two systems S_1, S_2 interact, there is a model S encompassing the two systems and meeting the conditions 4 above. It is obtained by taking the tensor product of the variables specific to S_1, S_2 . Then the Hilbert space of S is the tensorial product of the Hilbert spaces associated to each system.*

Proof. First let us see the consequences of the conditions if they are met.

The map $\varphi : H_1 \times H_2 \rightarrow H' :: \varphi(\psi_1, \psi_2) = \Phi(\Upsilon_1^{-1}(\psi_1), \Upsilon_2^{-1}(\psi_2))$ is bilinear. So, by the universal property of the tensorial product, there is a unique map $\widehat{\varphi} : H_1 \otimes H_2 \rightarrow H'$ such that $\varphi = \widehat{\varphi} \circ \iota$ where $\iota : H_1 \times H_2 \rightarrow H_1 \otimes H_2$ is the tensorial product.

The condition iii) reads :

$$\begin{aligned} & \langle \Upsilon_1(X_1), \Upsilon_1(X'_1) \rangle_{H_1} \times \langle \Upsilon_2(X_2), \Upsilon_2(X'_2) \rangle_{H_2} \\ &= \langle (\Upsilon' \circ \Phi(\Upsilon_1(X_1), \Upsilon_2(X_2)), \Upsilon' \circ \Phi(\Upsilon_1(X'_1), \Upsilon_2(X'_2))) \rangle_{H'} \\ & \langle \psi_1, \psi'_1 \rangle_{H_1} \times \langle \psi_2, \psi'_2 \rangle_{H_2} = \langle \varphi(\psi_1, \psi_2), \varphi(\psi'_1, \psi'_2) \rangle_{H'} = \langle \widehat{\varphi}(\psi_1 \otimes \psi_2), \widehat{\varphi}(\psi'_1 \otimes \psi'_2) \rangle_{H'} \end{aligned}$$

The scalar products on H_1, H_2 extend in a scalar product on $H_1 \otimes H_2$, endowing the latter with the structure of a Hilbert space with :

$$\langle (\psi_1 \otimes \psi_2), (\psi'_1 \otimes \psi'_2) \rangle_{H_1 \otimes H_2} = \langle \psi_1, \psi'_1 \rangle_{H_1} \langle \psi_2, \psi'_2 \rangle_{H_2}$$

and then the reproducing kernel is the product of the reproducing kernels (JCD Th.1208).

So we must have : $\langle \widehat{\varphi}(\psi_1 \otimes \psi_2), \widehat{\varphi}(\psi'_1 \otimes \psi'_2) \rangle_{H'} = \langle \psi_1 \otimes \psi_2, \psi'_1 \otimes \psi'_2 \rangle_{H_1 \otimes H_2}$ and $\widehat{\varphi}$ must be an isometry : $H_1 \otimes H_2 \rightarrow H'$

So by taking $H' = H_1 \otimes H_2$ and $V' = V_1 \otimes V_2$ we meet the conditions. ■

The conditions above are a bit abstract, but are logical and legitimate in the view of the Hilbert spaces. They lead to a natural solution, which is not unique and makes sense only if the systems are defined by similar variables. The measure of the tensor S can be addressed as before, the observables being linear maps defined in the tensorial products $V_1 \otimes V_2, H_1 \otimes H_2$ and valued in finite dimensional vector subspaces of these tensor products.

6.1.1 Comments

A key point in this representation is the difference between the simple direct product : $V_1 \times V_2$ and the tensorial product $V_1 \otimes V_2$, an issue about which there is much confusion.

The knowledge of the states (X_1, X_2) of both systems requires two vectors of I components each, that is 2×1 scalars, and the knowledge of the state S requires a vector of I^2 components. So the measure of S requires more data, and brings more information, because it encompasses all the interactions. Moreover *a tensor is not necessarily the tensorial product of vectors* (if it is so it is said to be separable), it can be the sum of such tensors. There is no canonical map : $V_1 \otimes V_2 \rightarrow V_1 \times V_2$. So there is no simple and unique way to associate two vectors (X_1, X_2) to one tensor S. This seems paradoxical, as one could imagine that both systems can always be studied, and their states measured, even if they

are interacting. But 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 as in the section 1. The study of their interactions can be a topic of the models, but this is done in another picture and requires additional hypotheses about the laws of the interactions. Or, if we intend to account for both systems and their interactions in a single model, we need a representation which supports more information that can bring $V_1 \times V_2$. The tensorial product is one way to enrich the model, this is the most economical and, as far as one follows the guidelines i),ii),iii) above, the only one. The complication in introducing general tensors is the price that we have to pay to account for the interactions. This representation does not, in any way, imply anything about *how* the systems interact, or even if they interact at all (in this case S is always 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 Mechanics a common interpretation of this representation is to single out separable tensors $\Psi = \psi_1 \otimes \psi_2$, called "pure states", so that actual states would be a superposition of pure states (a concept popularized by the famous Schrödinger's cat). It is clear that in an interacting system the pure states are an abstraction, which actually would represent two non interacting systems, so their superposition is an artificial construction. It can be convenient in simple cases, where the states of each system can be clearly identified, or in complicated models to represent quantities which are defined over the whole system as we will see later. But it does not imply any mysterious feature, notably any probabilist behaviour, for the real systems.

6.2 Homogeneous systems

The previous result can be extended to N (a number that we will assumed to be fixed) similar systems (that we will call microsystem), represented by the same model, interacting together. For each microsystem, identified by a label s , the Hilbert space H and the linear map Υ are the same, the state S of the total system can be represented as a vector belonging to the tensorial product $\mathbf{V}_N = \otimes_{s=1}^N V$, associated to a tensor Ψ belonging to the tensorial product $\mathbf{H}_N = \otimes_{s=1}^N H$. The linear maps $\Upsilon \in \mathcal{L}(V; H)$ can be uniquely extended as maps $\Upsilon_N \in \mathcal{L}(\mathbf{V}_N; \mathbf{H}_N)$ such that (JCD Th.423) :

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

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

We have general properties on these tensorial products (JCD Th.1208).

If $(\tilde{\varepsilon}_i)_{i \in I}$ is a hilbertian basis of H then $E_{i_1 \dots i_N} = \tilde{\varepsilon}_{i_1} \otimes \dots \otimes \tilde{\varepsilon}_{i_N}$ is a hilbertian basis of $\otimes_{s=1}^N H$. The scalar product is defined by linear extension of

$$\langle \Psi, \Psi' \rangle_{\mathbf{H}_N} = \langle \psi_1, \psi'_1 \rangle_H \times \dots \times \langle \psi_N, \psi'_N \rangle_H$$

for separable tensors :

$$\Psi = \psi_1 \otimes \dots \otimes \psi_N$$

$$\Psi' = \psi'_1 \otimes \dots \otimes \psi'_N$$

The subspaces $\otimes_{s=1}^p H \otimes \tilde{\varepsilon}_i \otimes_{s=p+2}^N H$ are orthogonal and $\mathbf{H}_N \simeq \ell^2(I^N)$

From this, any operator on H can be extended on \mathbf{H}_N with similar properties : a self adjoint, unitary or compact operator extends uniquely as a self adjoint, unitary or compact operator (JCD Th.1211).

In the general case the label matters : the state $S = X_1 \otimes \dots \otimes X_N$ is deemed different from $S = X_{\sigma(1)} \otimes \dots \otimes X_{\sigma(N)}$ where $(X_{\sigma(p)})_{p=1}^N$ is a permutation of $(X_s)_{s=1}^N$. If they have all the same behaviour they are, for the observer, indistinguishable. Usually the behaviour is related to a parameter analogous to a size, so in such cases the microsystems are assumed to have the same size. We will say that these interacting systems are **homogeneous** and we will characterize this assumption as follows : any permutation of the N microsystems gives the same state of the total system.

Proposition 18 *The states Ψ of homogeneous systems belong to an open subset \mathbf{h} of the Hilbert space \mathbf{H}_N defined by :*

i) a class of conjugacy $\mathfrak{S}(\lambda)$ of the group of permutations $\mathfrak{S}(N)$, defined itself by a decomposition of N in p parts :

$$\lambda = \{0 \leq n_p \leq \dots \leq n_1 \leq N, n_1 + \dots + n_p = N\}.$$

ii) p distinct vectors $(\tilde{\varepsilon}_j)_{j=1}^p$ of a hermitian basis of H

The tensors representing the states of the system read :

$$\forall \Psi \in \mathbf{h} : \Psi = \sum_{\sigma \in \mathfrak{S}(\lambda^c)} \Psi^\sigma \hat{U}(\sigma) (\otimes_{n_1} \tilde{\varepsilon}_{j_1} \otimes_{n_2} \tilde{\varepsilon}_{j_2} \dots \otimes_{n_p} \tilde{\varepsilon}_{j_p})$$

where $\mathfrak{S}(\lambda^c)$ is the conjugacy class complementary to $\mathfrak{S}(\lambda)$ and \hat{U} is the permutation of indices on \mathbf{H}_N

Proof. In the representation of the general system the microsystems are identified by some label $s = 1 \dots N$. An exchange of labels $U(\sigma)$ is a change of variables, represented by an action of the group of permutations $\mathfrak{S}(N)$: U is defined uniquely by linear extension of $U(\sigma)(X_1 \otimes \dots \otimes X_N) = X_{\sigma(1)} \otimes \dots \otimes X_{\sigma(N)}$ on separable tensors.

We can implement the results proven previously. To U is associated a unitary operator \hat{U} on \mathbf{H}_N such that (\mathbf{H}_N, \hat{U}) is a unitary representation of $\mathfrak{S}(N)$.

The action of \hat{U} on H_N is defined uniquely by linear extension of

$$\hat{U}(\sigma)(\psi_1 \otimes \dots \otimes \psi_N) = \psi_{\sigma(1)} \otimes \dots \otimes \psi_{\sigma(N)} \text{ on separable tensors.}$$

$\Psi \in \mathbf{H}_N$ reads in a Hilbert basis $(\tilde{\varepsilon}_i)_{i \in I}$ of H :

$$\Psi = \sum_{i_1 \dots i_N \in I} \Psi^{i_1 \dots i_N} \tilde{\varepsilon}_{i_1} \otimes \dots \otimes \tilde{\varepsilon}_{i_N} \text{ and :}$$

$$\begin{aligned} \hat{U}(\sigma)\Psi &= \sum_{i_1 \dots i_N \in I} \Psi^{i_1 \dots i_N} \hat{U}(\sigma)(\tilde{\varepsilon}_{i_1} \otimes \dots \otimes \tilde{\varepsilon}_{i_N}) = \sum_{i_1 \dots i_N \in I} \Psi^{i_1 \dots i_N} \tilde{\varepsilon}_{\sigma(i_1)} \otimes \\ &\dots \tilde{\varepsilon}_{\sigma(i_N)} \\ &= \sum_{i_1 \dots i_N \in I} \Psi^{\sigma(i_1) \dots \sigma(i_N)} \tilde{\varepsilon}_{i_1} \otimes \dots \otimes \tilde{\varepsilon}_{i_N} \end{aligned}$$

$$\begin{aligned} \langle \widehat{U}(\sigma)\Psi, \widehat{U}(\sigma)\Psi' \rangle &= \langle \Psi, \Psi' \rangle \\ \Leftrightarrow \sum_{i_1 \dots i_N \in I} \Psi^{\sigma(i_1) \dots \sigma(i_N)} \Psi'^{\sigma(i_1) \dots \sigma(i_N)} &= \sum_{i_1 \dots i_N \in I} \Psi^{i_1 \dots i_N} \Psi'^{i_1 \dots i_N} \end{aligned}$$

Let $\widehat{Y}_{J_N} = \otimes_N \widehat{Y}_J$ be the extension of the projection from H to H_J for any finite subset J of I , with $\text{card}(J) \geq N$ defined as previously :

$$\widehat{Y}_{J_N} (\sum_{i_1 \dots i_N \in I} \Psi^{i_1 \dots i_N} \tilde{\varepsilon}_{i_1} \otimes \dots \otimes \tilde{\varepsilon}_{i_N}) = \sum_{i_1 \dots i_N \in J} \Psi^{i_1 \dots i_N} \tilde{\varepsilon}_{i_1} \otimes \dots \otimes \tilde{\varepsilon}_{i_N}$$

Then :

$$\begin{aligned} \forall \sigma \in \mathfrak{S}(N) : \widehat{U}(\sigma) \widehat{Y}_{J_N} (\sum_{i_1 \dots i_N \in I} \Psi^{i_1 \dots i_N} \tilde{\varepsilon}_{i_1} \otimes \dots \otimes \tilde{\varepsilon}_{i_N}) \\ = \sum_{i_1 \dots i_N \in J} \Psi^{\sigma(i_1) \dots \sigma(i_N)} \tilde{\varepsilon}_{i_1} \otimes \dots \otimes \tilde{\varepsilon}_{i_N} = \widehat{Y}_{J_N} \widehat{U}(\sigma) \Psi \end{aligned}$$

So if $\mathbf{h} \subset \mathbf{H}_N$ is invariant by $\widehat{U} : \forall \sigma \in \mathfrak{S}(N) : \widehat{U}(\sigma)\mathbf{h} \subset \mathbf{h}$ then $\widehat{Y}_{J_N}\mathbf{h}$ is invariant by \widehat{U} . If $(\mathbf{h}, \widehat{U})$ is an irreducible representation then the only invariant subspace are 0 and \mathbf{h} itself, so necessarily $\mathbf{h} \subset \widehat{Y}_{J_N}\mathbf{H}_N$ for $\text{card}(J) = N$. Which implies : $\mathbf{h} \subset \otimes_N H_J$ with $H_J = \widehat{Y}_J H$.

$\mathfrak{S}(N)$ is a finite, compact group. Its unitary representations are the sum of orthogonal, finite dimensional, unitary, irreducible representations. These representations are not equivalent. To represent the possible states of the system we can choose any representation, and to be consistent these representations should be defined up to an isomorphism, so it shall be one of the irreducible representation.

Let \mathbf{h} be a subspace of \mathbf{H}_N such that $(\mathbf{h}, \widehat{U})$ is a finite dimensional, unitary, irreducible representation. Then $\mathbf{h} \subset \otimes_N H_J$.

\mathbf{h} is a Hilbert space, thus it has a hilbertian basis, composed of separable tensors which are of the kind $\tilde{\varepsilon}_{j_1} \otimes \dots \otimes \tilde{\varepsilon}_{j_N}$ where $\tilde{\varepsilon}_{j_k}$ are chosen among the vectors of a hermitian basis $(\tilde{\varepsilon}_j)_{j \in J}$ of H_J

If $\tilde{\varepsilon}_{j_1} \otimes \dots \otimes \tilde{\varepsilon}_{j_N} \in H, \forall \sigma \in \mathfrak{S}(N) : \widehat{U}(\sigma)\tilde{\varepsilon}_{j_1} \otimes \dots \otimes \tilde{\varepsilon}_{j_N} = \tilde{\varepsilon}_{j_{\sigma(1)}} \otimes \dots \otimes \tilde{\varepsilon}_{j_{\sigma(N)}} \in \mathbf{h}$ and because the representation is irreducible the basis of \mathbf{h} is necessarily composed from a set of $p \leq N$ vectors $\tilde{\varepsilon}_j$ by action of $\widehat{U}(\sigma)$

Conversely :

There is a partition of $\mathfrak{S}(N)$ in conjugacy classes $\mathfrak{S}(\lambda)$ which are subgroups defined by a decomposition of N in p parts :

$\lambda = \{0 \leq n_p \leq \dots \leq n_1 \leq N, n_1 + \dots + n_p = N\}$. Notice that there is an order on the sets $\{\lambda\}$. Each element of a conjugacy class is then defined by a repartition of the integers $\{1, 2, \dots, N\}$ in p subsets of n_k items (this is a Young Tableau) (JCD 5.2.2).

For any hermitian basis $(\tilde{\varepsilon}_i)_{i \in I}$ of H , any subset J of cardinality N of I , any conjugacy class λ , any family of vectors $(\tilde{\varepsilon}_{j_k})_{k=1}^p$ chosen in $(\tilde{\varepsilon}_i)_{i \in J}$, the action of \widehat{U} on the tensor :

$$\Psi_\lambda = \otimes_{n_1} \tilde{\varepsilon}_{j_1} \otimes_{n_2} \tilde{\varepsilon}_{j_2} \dots \otimes_{n_p} \tilde{\varepsilon}_{j_p}, j_1 \leq j_2 \dots \leq j_p$$

gives the same tensor if $\sigma \in \mathfrak{S}(\lambda) : \widehat{U}(\sigma)\Psi_\lambda = \Psi_\lambda$

gives a different tensor if $\sigma \in \mathfrak{S}(\lambda^c)$ the conjugacy class complementary to

$$\mathfrak{S}(\lambda) : \mathfrak{S}(\lambda^c) = \mathfrak{C}_{\mathfrak{S}(N)}^{\mathfrak{S}(\lambda)}$$

so it provides an irreducible representation by :

$$\forall \Psi \in \mathbf{h} : \Psi = \sum_{\sigma \in \mathfrak{S}(\lambda^c)} \Psi^\sigma \widehat{U}(\sigma) (\otimes_{n_1} \tilde{\varepsilon}_{j_1} \otimes_{n_2} \tilde{\varepsilon}_{j_2} \dots \otimes_{n_p} \tilde{\varepsilon}_{j_p})$$

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

So, up to isomorphism, there is a bijective correspondance between the conjugacy classes λ of $\mathfrak{S}(N)$ and the unitary irreducible representation of $\mathfrak{S}(N)$.

\mathbf{h} is then isomorphic to $\odot_{n_1} H_J \odot_{n_2} H_J \dots \odot_{n_p} H_J$ where the symmetric tensorial product \odot and the symmetrizer S_n are defined uniquely by linear extension of :

$$S_n : \otimes_n H \rightarrow \odot_n H :: S_n(\psi_1 \otimes \dots \otimes \psi_k) = \sum_{\sigma \in \mathfrak{S}(n)} \psi_{\sigma(1)} \otimes \dots \otimes \psi_{\sigma(k)}$$

$$\odot_{k=1}^n \psi_k = S_n(\otimes_{k=1}^n \psi_k)$$

and the space of n order symmetric tensor on H_J is $\odot_n H_J$

The result extends to V_N by : $S = \Upsilon_N^{-1}(\Psi)$ ■

Remarks :

i) $\otimes_{n_1} \tilde{\varepsilon}_{j_1} \otimes_{n_2} \tilde{\varepsilon}_{j_2} \dots \otimes_{n_p} \tilde{\varepsilon}_{j_p}$ can be seen as representing a configuration where n_k microsystems are in the same state $\tilde{\varepsilon}_{j_k}$. The class of conjugacy, characterized by the integers n_p , correspond to the distribution of the microsystems between fixed states. Different choices of the hermitian basis $(\tilde{\varepsilon}_j)_{j \in I}$ and the subset J of I give an equivalent representation, so they can be arbitrary, but then the representations given by different conjugacy classes are not equivalent. So, for a given system, the set of states is characterized by a subset J of N elements in any basis of H , and by a class of conjugacy.

ii) If O is a convex subset then S belongs to a convex subset, and the basis can be chosen such that $\forall \Psi \in \mathbf{h}$ is a linear combination $(y_k)_{k=1}^q$ of the generating tensors with $y_k \in [0, 1]$, $\sum_{k=1}^q y_k = 1$. S can then be identified to the expected value of a random variable which would take one of the value $\otimes_{n_1} X_1 \otimes_{n_2} X_2 \dots \otimes_{n_p} X_p$, which corresponds to n_k microsystems having the state X_k . As exposed above the identification with a probabilist model is formal : there is no random behaviour involved for the physical system.

iii) In the probabilist picture one can assume that each microsystem behaves independantly, and has a probability π_j to be in the state represented by $\tilde{\varepsilon}_j$ and $\sum_{j=1}^N \pi_j = 1$. Then the probability that we have $(n_k)_{k=1}^p$ microstates in the states $(\tilde{\varepsilon}_k)_{k=1}^p$ is $\frac{N!}{n_1! \dots n_p!} (\pi_{j_1})^{n_1} \dots (\pi_{j_p})^{n_p}$.

iv) The set of symmetric tensor $\odot_n H_J$ is a closed vector subspace of $\otimes_n H_J$, this is a Hilbert space, $\dim \odot_n H_J = C_{N+n-1}^{N-1}$ with hilbertian basis $\frac{1}{\sqrt{n!}} \odot_{j \in J} \tilde{\varepsilon}_j = \frac{1}{\sqrt{n!}} S_N(\otimes_{j \in J} \tilde{\varepsilon}_j)$ (JCD 7.2.1,13.5.2).

A tensor is symmetric iff : $\Psi \in \odot_n H_J \Leftrightarrow S_n(\Psi) = n! \Psi$

v) for $\theta \in \mathfrak{S}(N)$: $\hat{U}(\theta)\Psi$ is usually different from Ψ

6.3 Global observables of homogeneous systems

The previous definitions of observables can be extended to homogeneous systems. However usually the values of the variables for each microsystem are not

known, and one uses global observables, which are linear map $G : \mathbf{V}_N \rightarrow F$ valued in some finite dimensional vector space F,

Proposition 19 *To any global observable G of a homogeneous system, which is in the state Ψ defined by a class of conjugation λ , is associated a linear map $\widehat{G} : \otimes_N H_J \rightarrow F :: \widehat{G} = G \circ \Upsilon_N^{-1}$*

$\widehat{G}(\Psi) = k\widehat{G}(\otimes_{n_1} \tilde{\varepsilon}_{j_1} \otimes_{n_2} \tilde{\varepsilon}_{j_2} \dots \otimes_{n_p} \tilde{\varepsilon}_{j_p})$ for a conjugacy class λ of $\mathfrak{S}(N)$, a family of vectors $(\tilde{\varepsilon}_{j_k})_{k=1}^p$ belonging to a hermitian basis $(\tilde{\varepsilon}_j)_{j \in J}$ of a N dimensional vector subspace H_J of H

k belongs to an open subset of \mathbb{R}

Proof. \widehat{G} must be symmetric for a homogeneous system : $\forall \Psi \in \mathbf{h}, \forall \sigma \in \mathfrak{S}(N) : \widehat{G} \circ \widehat{U}(\sigma) \Psi = \widehat{G} \Psi$

For a given system the space \mathbf{h} is characterized by the choice of N elements $(\tilde{\varepsilon}_j)_{j \in J}$ in any hilbertian basis of H and by a class of conjugacy λ . Then :

$$\Psi_\lambda = \otimes_{n_1} \tilde{\varepsilon}_{j_1} \otimes_{n_2} \tilde{\varepsilon}_{j_2} \dots \otimes_{n_p} \tilde{\varepsilon}_{j_p}, j_1 \leq j_2 \dots \leq j_p$$

$$\forall \Psi \in \mathbf{h} : \Psi = \sum_{\sigma \in \mathfrak{S}(\lambda^c)} \Psi^\sigma \widehat{U}(\sigma) (\Psi_\lambda)$$

$$\widehat{G}(\Psi) = \sum_{\sigma \in \mathfrak{S}(\lambda^c)} \Psi^\sigma \widehat{G}(\widehat{U}(\sigma) (\Psi_\lambda)) = C_\lambda(\Psi) \widehat{G}(\Psi_\lambda)$$

$$\text{with } C_\lambda(\Psi) = \sum_{\sigma \in \mathfrak{S}(\lambda^c)} \Psi^\sigma \in \mathbb{R}$$

So, for a given system the global variable $\widehat{G}(\Psi)$ is valued in a one dimensional vector space :

$\widehat{G}(\Psi) = k\widehat{G}(\otimes_{n_1} \tilde{\varepsilon}_{j_1} \otimes_{n_2} \tilde{\varepsilon}_{j_2} \dots \otimes_{n_p} \tilde{\varepsilon}_{j_p})$ where k belongs to an open subset of \mathbb{R} ■

The map \widehat{G} is part of the model, it can be theoretically computed and, as the choice of the basis $(\tilde{\varepsilon}_j)_{j \in I}$ and the subset J of I are arbitrary, the set $\left\{ \widehat{G}(\otimes_{n_1} \tilde{\varepsilon}_{j_1} \otimes_{n_2} \tilde{\varepsilon}_{j_2} \dots \otimes_{n_p} \tilde{\varepsilon}_{j_p}), \lambda, (\tilde{\varepsilon}_{j_k})_{k=1}^p \right\}$ is known, but the map \widehat{G} is not bijective so, to a given measure g of $\widehat{G}(\Psi)$, usually correspond different values of $k, \lambda, (\tilde{\varepsilon}_{j_k})_{k=1}^p$. The question which arises is then to estimate these quantities from a measure g .

As one can see there is no way to estimate Ψ^σ , as long as \widehat{G} is symmetric, as it should be for a homogeneous system. So one cannot attributes a vector to the states of the microsystems : they are identified by a label, which corresponds to a vector $\tilde{\varepsilon}_{j_k}$ of a hilbertian basis. This can be seen as a quantization of the states of the microsystems. $g = \widehat{G}(\Psi)$ itself is defined up to a scalar. As g and $\widehat{G}(\otimes_{n_1} \varepsilon_{j_1} \otimes_{n_2} \varepsilon_{j_2} \dots \otimes_{n_p} \varepsilon_{j_p})$ are both vectors, measured or computed, they are known, and actually the scalar shall be interpreted as related to the choice of units in representing the states $\tilde{\varepsilon}_{j_k}$. So, from a practical point of view, this is a vector $\gamma = \widehat{G}(\otimes_{n_1} \tilde{\varepsilon}_{j_1} \otimes_{n_2} \tilde{\varepsilon}_{j_2} \dots \otimes_{n_p} \tilde{\varepsilon}_{j_p})$ which is measured, and from it $\lambda, (\tilde{\varepsilon}_{j_k})_{k=1}^p$ are estimated.

In the probabilist picture the expected value of γ is :

$$\langle \gamma \rangle = z(\pi_1, \dots, \pi_N)$$

with

$$z(\pi_1, \dots, \pi_N) = \sum_{\lambda} \frac{N!}{n_1! \dots n_p!} \sum_{1 \leq j_1 \leq \dots \leq j_p \leq N} (\pi_{j_1})^{n_1} \dots (\pi_{j_p})^{n_p} \widehat{G}(\otimes_{n_1} \varepsilon_{j_1} \otimes_{n_2} \varepsilon_{j_2} \dots \otimes_{n_p} \varepsilon_{j_p})$$

We have a classic statistical problem. The usual method is based on the principle of Maximum Entropy popularized by Jaynes, that is to find $(\pi_j)_{j=1}^N$ meeting the constraints above and maximizing the entropy :

$$E = - \sum_{j=1}^N \pi_j \ln \pi_j$$

There is an abundant litterature on the subject, both in Statistics and in Physics about this principle and the meaning of entropy.

From a pure statistical point of view the problem can be seen as the estimation of the π_i from a statistic given by the measure of G. If the statistic G is sufficient, meaning that π_i depends only on γ , as F is finite dimensional whatever the number of microsystems, the Pitman-Koopman-Darmois theorem tells us that the probability law is exponential, then an estimation by the maximum likelihood gives the principle of Maximum Entropy.

In the usual interpretation of the probabilist picture, it is assumed that the state of each microsystem can be measured independantly. Then the entropy $E = - \sum_{j=1}^N \pi_j \ln \pi_j$ can be seen as a measure of the heterogeneity of the system. And it is asserted that the homogenisation of the states is the result of their interactions. But this is contradictory to the concept of homogeneous system : the microsystems, being undistinguishable, should be in the same state. In the framework presented here the interpretation is different : a tensor such as $\otimes_{n_1} \varepsilon_{j_1} \otimes_{n_2} \varepsilon_{j_2} \dots \otimes_{n_p} \varepsilon_{j_p}$ represents the state of the system, interactions included, the states of the individual microsystems have no definite value, the probability has no physical meaning : this is just a convenient way to visualize the state of the global system.

However these issues do not reduce the interest of entropy and conjugate variables. And actually the present framework gives a way to extend these concepts out of their traditional domains.

6.4 Evolution of homogeneous systems

The evolution of homogeneous systems raises many interesting issues. The assumptions are a combination of the previous conditions.

Proposition 20 *For a model representing the evolution of a homogeneous system comprised of a fixed number N of microsystems $s = 1 \dots N$ which are represented by the same model, with variables $(X_s)_{s=1}^N$ such that, for each microsystem :*

i) the variables X_s are maps : $X_s :: R \rightarrow E$ where R is an open subset of \mathbb{R} and E a normed vector space, belonging to an open subset O of an infinite dimensional Fréchet space V

ii) $\forall t \in R$ the evaluation map : $E(t) : O \rightarrow E : E(t) X_s = X_s(t)$ is continuous

iii) $\forall t \in R : X_s(t) = X'_s(t) \Rightarrow X_s = X'_s$

iv) The variable t can be defined and measured uniformly over the system

There is a map $S : R \rightarrow \otimes_N F$ such that $S(t)$ represents the state of the system at t . $S(t)$ takes its value in a vector space $f(t)$ such that $(\mathbf{f}(t), \widehat{U}_F)$, where \widehat{U}_F is the permutation on $\otimes_N F$, is an irreducible representation of $S(N)$

The crucial point is that the homogeneity is understood as the microsystems follow the same laws, but at a given time they do not have necessarily the same state.

The condition ii) refers to the issue in Relativist Physics of a time common to a system (which can be circumvented).

Proof. i) Implement the proposition 1 for each microsystem : there is a common Hilbert space H associated to V and a continuous linear map $\Upsilon : V \rightarrow H :: \psi_s = \Upsilon(X_s)$

ii) Implement the proposition 18 on the homogeneous system, that is for the whole of its evolution. The state of the system is associated to a tensor $\Psi \in \mathbf{h}$ where \mathbf{h} is defined by a hilbertian basis $(\tilde{\varepsilon}_i)_{i \in I}$ of H , a finite subset J of I , a conjugacy class λ and a family of p vectors $(\tilde{\varepsilon}_{jk})_{k=1}^p$ belonging to $(\tilde{\varepsilon}_i)_{i \in J}$. The vector space \mathbf{h} stays the same whatever t .

iii) Implement proposition 15 on the evolution of each microsystem : there is a common Hilbert space F , a map $\widehat{\mathcal{E}} : R \rightarrow \mathcal{L}(H; F)$ such that $\forall X_s \in O : \widehat{\mathcal{E}}(t) \Upsilon(X_s) = X_s(t)$ and $\forall t \in R$, $\widehat{\mathcal{E}}(t)$ is an isometry

Define $\forall i \in I : \varphi_i : R \rightarrow F :: \varphi_i(t) = \widehat{\mathcal{E}}(t) \tilde{\varepsilon}_i$

iv) $\widehat{\mathcal{E}}(t)$ can be uniquely extended in a continuous linear map :

$\widehat{\mathcal{E}}_N(t) : \otimes_N H \rightarrow \otimes_N F$ such that $\widehat{\mathcal{E}}_N(t) (\otimes_N \psi_s) = \otimes_N X_s(t)$

$\widehat{\mathcal{E}}_N(t) (\otimes_{s=1}^N \tilde{\varepsilon}_{i_s}) = \otimes_{s=1}^N \varphi_{i_s}(t)$

$\widehat{\mathcal{E}}_N(t)$ is an isometry, so $\forall t \in R : \{\otimes_{s=1}^N \varphi_{i_s}(t), i_s \in I\}$ is a hilbertian basis of $\otimes_N F$

v) Define as the state of the system at $t : S(t) = \widehat{\mathcal{E}}_N(t) (\Psi) \in \otimes_N F$

Define $\forall \sigma \in \mathfrak{S}(N) : \widehat{U}_F(\sigma) \in \mathcal{L}(\otimes_N F; \otimes_N F)$ by linear extension of :

$$\widehat{U}_F(\sigma) (\otimes_{s=1}^N f_s) = \otimes_{s=1}^N f_{\sigma(s)}$$

$$\widehat{U}_F(\sigma) (\otimes_{s=1}^N \varphi_{i_s}(t)) = \otimes_{s=1}^N \varphi_{\sigma(i_s)}(t) = \widehat{\mathcal{E}}_N(t) \widehat{U}(\sigma) (\otimes_{s=1}^N \tilde{\varepsilon}_{i_s})$$

$$\forall \Psi \in \mathbf{h} : \Psi = \sum_{\sigma \in \mathfrak{S}(\lambda^c)} \Psi^\sigma \widehat{U}(\sigma) (\otimes_{n_1} \tilde{\varepsilon}_{j_1} \otimes_{n_2} \tilde{\varepsilon}_{j_2} \dots \otimes_{n_p} \tilde{\varepsilon}_{j_p})$$

$$S(t) = \sum_{\sigma \in \mathfrak{S}(\lambda^c)} \Psi^\sigma \widehat{\mathcal{E}}_N(t) \circ \widehat{U}(\sigma) (\otimes_{n_1} \tilde{\varepsilon}_{j_1} \otimes_{n_2} \tilde{\varepsilon}_{j_2} \dots \otimes_{n_p} \tilde{\varepsilon}_{j_p})$$

$$S(t) = \sum_{\sigma \in \mathfrak{S}(\lambda^c)} \Psi^\sigma \widehat{U}_F(\sigma) \otimes_{n_1} \varphi_{j_1}(t) \otimes_{n_2} \varphi_{j_2}(t) \dots \otimes_{n_p} \varphi_{j_p}(t)$$

$$\forall \theta \in \mathfrak{S}(\lambda) : \widehat{U}_F(\theta) (\otimes_{n_1} \varphi_{j_1}(t) \otimes_{n_2} \varphi_{j_2}(t) \dots \otimes_{n_p} \varphi_{j_p}(t))$$

$$= \otimes_{n_1} \varphi_{j_1}(t) \otimes_{n_2} \varphi_{j_2}(t) \dots \otimes_{n_p} \varphi_{j_p}(t)$$

$$\forall \theta \in \mathfrak{S}(\lambda^c) : \widehat{U}_F(\theta) (\otimes_{n_1} \varphi_{j_1}(t) \otimes_{n_2} \varphi_{j_2}(t) \dots \otimes_{n_p} \varphi_{j_p}(t))$$

$$\neq (\otimes_{n_1} \varphi_{j_1}(t) \otimes_{n_2} \varphi_{j_2}(t) \dots \otimes_{n_p} \varphi_{j_p}(t))$$

and the tensors are linearly independant

So $\left\{ \widehat{U}_F(\sigma) (\otimes_{n_1} \varphi_{j_1}(t) \otimes_{n_2} \varphi_{j_2}(t) \dots \otimes_{n_p} \varphi_{j_p}(t)), \sigma \in \mathfrak{S}(\lambda^c) \right\}$ is an orthonormal basis of

$$\mathbf{f}(t) = \text{Span} \left\{ \widehat{U}_F(\sigma) (\otimes_{n_1} \varphi_{j_1}(t) \otimes_{n_2} \varphi_{j_2}(t) \dots \otimes_{n_p} \varphi_{j_p}(t)), \sigma \in \mathfrak{S}(\lambda^c) \right\}$$

$$\mathbf{f}(t) = \widehat{\mathcal{E}}_N(t) (\mathbf{h})$$

Let $\tilde{f}(t) \subset \mathbf{f}(t)$ be any subspace globally invariant by $\left\{ \widehat{U}_F(\theta), \theta \in \mathfrak{S}(N) \right\}$:

$$\widehat{U}_F(\theta) \tilde{f}(t) \in \tilde{f}(t)$$

$\widehat{\mathcal{E}}_N(t)$ is an isometry, thus a bijective map

$$\tilde{h} = \widehat{\mathcal{E}}_N(t)^{-1} \tilde{f}(t) \Leftrightarrow \tilde{f}(t) = \widehat{\mathcal{E}}_N(t) \tilde{h}$$

$$\widehat{U}_F(\theta) \widehat{\mathcal{E}}_N(t) \tilde{h} \in \widehat{\mathcal{E}}_N(t) \tilde{h}$$

$$\forall \Psi \in \mathbf{h} : \widehat{U}_F(\theta) \widehat{\mathcal{E}}_N(t) \Psi = \widehat{\mathcal{E}}_N(t) \widehat{U}(\theta) \Psi$$

$$\Rightarrow \widehat{\mathcal{E}}_N(t) \widehat{U}(\theta) \tilde{h} \in \widehat{\mathcal{E}}_N(t) \tilde{h}$$

$$\Rightarrow \widehat{U}(\theta) \tilde{h} \in \tilde{h}$$

So $(\mathbf{f}(t), \widehat{U}_F)$ is an irreducible representation of $\mathfrak{S}(N)$ ■

For each t the space $\mathbf{f}(t)$ is defined by a hilbertian basis $(f_i)_{i \in I}$ of F, a finite subset J of I, a conjugacy class $\lambda(t)$ and a family of p vectors $(f_{j_k}(t))_{k=1}^p$ belonging to $(f_i)_{i \in J}$. The set J is arbitrary but defined by \mathbf{h} , so it does not depend on t. For a given class of conjugacy different families of vectors $(f_{j_k}(t))_{k=1}^p$ generate equivalent representations and isomorphic spaces. So for a given system one can pick up a fixed ordered family $(f_j)_{j=1}^N$ of vectors in $(f_i)_{i \in I}$ such that for each class of conjugacy $\lambda = \{0 \leq n_p \leq \dots \leq n_1 \leq N, n_1 + \dots + n_p = N\}$ there is a unique vector space \mathbf{f}_λ defined by $\otimes_{n_1} f_1 \otimes_{n_2} f_2 \dots \otimes_{n_p} f_p$. Then if $S(t) \in \mathbf{f}_\lambda$:

$$S(t) = \sum_{\sigma \in \mathfrak{S}(\lambda^c)} S^\sigma(t) \widehat{U}_F(\sigma) (\otimes_{n_1} f_1 \otimes_{n_2} f_2 \dots \otimes_{n_p} f_p)$$

and at all time $S(t) \in \otimes_N F_J$.

The vector spaces \mathbf{f}_λ are orthogonal. With the orthogonal projection π_λ on \mathbf{f}_λ :

$$\forall t \in R : S(t) = \sum_\lambda \pi_\lambda S(t)$$

$$\|S(t)\|^2 = \sum_\lambda \|\pi_\lambda S(t)\|^2$$

The distance between S(t) and a given \mathbf{f}_λ is well defined and :

$$\|S(t) - \pi_\lambda S(t)\|^2 = \|S(t)\|^2 - \|\pi_\lambda S(t)\|^2$$

Whenever S, and thus Θ , is continuous, the space \mathbf{f}_λ stays the same. As we have seen previously one can assume that, in all practical cases, Θ is continuous but for a countable set $\{t_k, k = 1, 2, \dots\}$ of isolated points. Then the different spaces \mathbf{f}_λ can be seen as phases, each of them associated with a class of conjugacy λ . And there are as many possible phases as classes of conjugacy. So, in a probabilist picture, one can assume that the probability for the system to be in a phase λ : $\Pr(S(t) \in \mathbf{f}_\lambda)$ is a function of $\frac{\|\pi_\lambda S(t)\|^2}{\|S(t)\|^2}$. It can be estimated as seen previously from data on a past period, with the knowlege of both λ and $\frac{\|\pi_\lambda S(t)\|^2}{\|S(t)\|^2}$.

7 CORRESPONDANCE WITH QUANTUM MECHANICS

It is useful to compare the results proven in the present paper to the axioms of QM as they are usually expressed.

7.1 Hilbert space

QM : 1. *The states of a physical system can be represented by "rays" in a complex Hilbert space H . Rays meaning that two vectors which differ by the product by a complex number of module 1 shall be considered as representing the same state.*

In proposition 1 we have proven that in a model meeting precise conditions the states of the system can be represented as vectors in an infinite dimensional, separable, real Hilbert space. We have seen that it is always possible to endow the Hilbert space with a complex structure, but this is not a necessity. Moreover the Hilbert space is defined up to an isometry, so notably up to the product by a fixed complex scalar of module 1.

In Quantum Mechanics some states of a system cannot be achieved (through a preparation for instance) as a combination of other states, and thus superselection rules are required to sort out these specific states. Here there is a simple explanation : because the set H_0 is not the whole of H it can happen that a linear combination of states is not inside H_0 . The remedy is to enlarge the model to account for other physical phenomena, if it appears that these states have a physical meaning.

In QM there is a distinction between pure states, which correspond to actual measures, and mixed states which are linear combination of pure states, usually not actually observed. There has been a great effort to give a physical meaning to these mixed states. Here the concept of pure states appears only in the tensors representing interacting systems, with the usual, but clear, explanation.

Actually the main difference comes from the precise conditions in the proposition 1. The variables must be maps, but also must belong to a vector space. Thus for instance it does not apply to the model of a solid body represented by its trajectory $x(t)$ and its speed $v(t)$: the variable $x(t)$ is a map : $x : \mathbb{R} \rightarrow M$ valued in a manifold (an affine space in galilean geometry). So it is necessary to adapt the model, using the fiber bundle formalism, and this leads to a deep redefinition of the concept of motion (including rotation) and to the spinors. And as it has been abundantly said, the state is defined by maps over the evolution of the system, and not pointwise.

In the present picture there is no concept of wave function : the vector in the Hilbert space does not depend on any coordinate, except for the evolution of systems, where, in precise conditions, this vector depends on the time. The

Hilbert space can be represented by spaces of functions (for instance when a group representation is involved) but there is no given association between the arguments of these functions and the arguments of the maps in the model (such as a location). However it should happen quite often that the space V of maps is itself a Hilbert space, then the wave function can be assimilated to X with its arguments.

7.2 Observables

QM : 2. *To any physical measure Φ , called an observable, that can be done on the system, is associated a continuous, linear, self-adjoint operator $\hat{\Phi}$ on H .*

We have proven that this operator is also compact and trace-class. The main result is that we have here a clear understanding of the concept of observable, which is usually muddled in QM. In the axiomatic presentation of QM through von Neumann algebras, the observables can be identified with our extended primary observables, defined from any basis on V and taking J possibly infinite. It is clear that the model is well defined by a unique basis, thus one can say that the von Neumann algebra is commutative.

In QM a great emphasize is given to the commutation of observables, linked to the physical possibility to measure simultaneously two variables. This concept does not play any role here, for the strong reason that we consider maps with a domain over the whole extension, spatial and temporal, of the system. In our picture the variables and their properties are the model, they are listed explicitly and it is assumed that there is some way to estimate their value, without any consideration of the time when the measures are done. So the question of simultaneous measures does not arise, and the product of observables itself has no clear meaning and no use. If a variable is added, we have another model, the variable gets the same status as the others, and it is assumed that it can be measured.

Actually the importance granted to the simultaneity of measures, magnified by Dirac, is somewhat strange. It is clear that some measures cannot be done, at the atomic scale, without disturbing the state of the system that is studied, but this does not preclude to use the corresponding variables in a model, or give them a special status. Before the invention of radar the artillerymen used efficient models even if they were not able to measure the speed of their shells. And in a collider it is assumed that the speed and the location of particles are known when they collide.

7.3 Measure

QM : 3. *The result of any physical measure is one of the eigen-values λ of the associated operator $\hat{\Phi}$. After the measure the system is in the state represented by the corresponding eigen vector ψ_λ*

This is one of the most puzzling axiom. We have here a clear interpretation of this result, with primary observables, and there is always a primary observable which is at least as efficient than a secondary observable.

In our picture the process of measure has no impact on the state of the system, which is fully determined before and after the measure, and stays the same. Whenever a primary observable is measured, one always get the same result in the same process. Thus there is no mystery such has the collapse of a wave function.

7.4 Probability

QM : 4. *The probability that the measure is λ is equal to $|\langle \psi_\lambda, \psi \rangle|^2$ (with normalized eigen vectors). If a system is in a state represented by a normalized vector ψ , and an experiment is done to test whether it is in one of the states $(\psi_n)_{n=1}^N$ which constitutes an orthonormal set of vectors, then the probability of finding the system in the state ψ_n is $|\langle \psi_n, \psi \rangle|^2$.*

The first part is addressed by proposition 12. The second part has no direct equivalent in our picture but can be interpreted as follows : a measure of the primary observable has shown that $\psi \in H_J$ then the probability that it belongs to $H_{J'}$ for any subset $J' \subset J$ is $\|\widehat{Y}_{J'}(\psi)\|^2$. It is a computation of conditional probabilities :

Proof. The probability that $\psi \in H_K$ for any subset $K \subset I$ is $\|\widehat{Y}_K(\psi)\|^2$. The probability that $\psi \in H_{J'}$ knowing that $\psi \in H_J$ is :

$$\Pr(\psi \in H_{J'} | \psi \in H_J) = \frac{\Pr(\psi \in H_{J'} \wedge \psi \in H_J)}{\Pr(\psi \in H_J)} = \frac{\Pr(\psi \in H_{J'})}{\Pr(\psi \in H_J)} = \frac{\|\widehat{Y}_{J'}(\psi)\|^2}{\|\widehat{Y}_J(\psi)\|^2} = \|\widehat{Y}_{J'}(\psi)\|^2 \text{ because } \widehat{Y}_{J'}(\psi) = \psi \text{ and } \|\psi\| = 1 \blacksquare$$

7.5 Interacting systems

QM : 5. *When two systems interacts, the vectors representing the states belong to the tensorial product of the Hilbert states.*

This is the topic of proposition 17. We have seen how it can be extended to N systems, and the consequences that entails for homogeneous systems. If the number of microsystems is not fixed, the formalism of Fock spaces can be used but would require a mathematical apparatus that is beyond the scope of this paper.

There is a fierce debate about the issue of locality in physics, mainly related to the entanglement of states for interacting particles. It should be clear that the formal system that we have built is global : more so, it is its main asset.

While most of the physical theories are local, with the tools which have been presented we can deal with variables which are global, and get some strong results without many assumptions regarding the local laws.

7.6 Wigner's theorem

QM : 6. *If the same state is represented by two rays R, R' , then there is an operator \widehat{U} , unitary or antiunitary, on the Hilbert space H such that if the state ψ is in the ray R then $\widehat{U}\psi$ is in the ray R' .*

This is the topic of the proposition 14. The operator is necessarily unitary, which is actually usually the case.

7.7 Schrödinger equation

QM : 7. *The vector representing the state of a system which evolves with time follows the equation : $i\hbar\frac{\partial\psi}{\partial t} = \widehat{H}\psi$ where \widehat{H}*

This is actually the topic of the proposition 16 and the result holds for the variables X in specific conditions, including in the General Relativity context (with charts built as described previously). The imaginary i does not appear because the Hilbert space is real. As for Planck's constant of course it cannot appear in a formal model. However as said before all quantities must be dimensionless, as it is obvious in the equivalent expression $\psi(t) = \exp\frac{t}{i\hbar}\widehat{H}\psi(0)$. Thus it is necessary either to involve some constant, or that all quantities (including the time t) are expressed in a universal system of units. This is commonly done by using the Planck's system of units.

7.8 The scale issue

The results presented here hold whenever the model meets the conditions 1. So it is valid whatever the scale. But it is clear that the conditions are not met in many models used in classic physics, notably in Analytic Mechanics (the variables q are not vectorial quantities). Moreover actually in the other cases it can often be assumed that the variables belong themselves to Hilbert spaces. The results about observables and eigen values are then obvious, and those about the evolution of the system, for interacting systems or for gauge theories keep all their interest.

The "Quantic World", with its strange properties does not come from specific physical laws, which would appear below some scale, but from the physical properties of the atomic world themselves. And of course these cannot be addressed in the simple study of formal models.

So the results presented here, which are purely mathematical, give a consistent and satisfying explanation of some of the basic axioms of Quantum Mechanics, without the need for any exotic assumptions. They validate, and in many ways make simpler and safer, the use of techniques used for many years.

Moreover, as it is easy to check, most of these results do not involve any physics at all : they hold for any scientific theory which is expressed in a mathematical formalism. How this can be used is the topic of another paper, which deals with Epistemology (Common structures in scientific theories).

jc.dutailly@free.fr

BIBLIOGRAPHY

- H.Araki *Mathematical theory of quantum fields* Oxford Science Publications (2000)
- A.Berlinet, C.Thomas-Agnan *Reproducing kernel, Hilbert spaces in probability and statistics* Springer (2004)
- O.Bratelli, D.W.Robinson *Operators algebras and quantum statistical mechanics* Springer (2002)
- P.A.M.Dirac *The principles of Quantum Mechanics* Oxford Science Publications (1958)
- J.C.Dutailly *Mathematics for theoretical physics* arXiv:1209-5665v2 [math-ph] 4 feb 2014
- J.C.Dutailly *Particles and fields* CNRS (<http://hal.archives-ouvertes.fr/hal-0933043>, version 1 19 janvier 2014)
- J.C.Dutailly *Common structures in scientific theories* CNRS (<http://hal.archives-ouvertes.fr/hal-01003869>, version 1 10 juin 2014)
- Tepper L.Gill, G.R.Pantsulaia, W.W.Zachary *Constructive analysis in infinitely many variables* arXiv 1206-1764v2 [math-FA] 26 june 2012
- H.Halvorson *Algebraic quantum fields theory* arXiv:math-ph/0602036v1 14 feb 2006
- D.W.Henderson *Infinite dimensional manifolds are open subsets of Hilbert spaces* (1969) Internet paper
- F.Laloë *Comprenons-nous vraiment la mécanique quantique ?* CNRS Editions (2011)
- J.von Neumann *Mathematical Foundations of Quantum Mechanics*, Beyer, R. T., trans., Princeton Univ. Press. 1996 edition
- K.Popper *Quantum theory and the schism in physics* Routledge (1982)
- S.Weinberg *The quantum theory of fields* Cambridge University Press (1995)