



HAL
open science

HILBERT SPACES IN MODELLING OF SYSTEMS

Jean Claude Dutailly

► **To cite this version:**

| Jean Claude Dutailly. HILBERT SPACES IN MODELLING OF SYSTEMS. 2014. <hal-00974251>

HAL Id: hal-00974251

<https://hal.science/hal-00974251v1>

Preprint submitted on 5 Apr 2014

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.



HAL Authorization

HILBERT SPACES IN MODELLING OF SYSTEMS

Jean Claude Dutailly
Paris

April 4, 2014

Abstract

This paper studies the quantitative models which are used to represent systems, whatever the field (Physics, Economics, Finances,...). We prove that models which meet some very general conditions can be associated to Hilbert vector spaces : the state of the system is then identified with a unique vector. The measure of this state is done through observables, whose properties are studied and we show that they can take only values which are eigen vectors of the observables, and a probability law can be associated to the occurrence of these measures. The effect of a change of variables in a model is related to a unitary operator. This leads to general laws for the evolution of systems, and to powerful tools for the study of transitions of phases in a system.

Interacting systems can be modelled in tensorial product of spaces, and the states of homogeneous systems are associated to classes of conjugacy of the symmetric group. These properties are preserved for the evolution of systems.

Almost all formal theories use mathematical models, in which phenomena occurring in a system are represented by variables : position or velocity of bodies, fields, population, economic indicators, value of stocks,...These variables have two purposes :

- they enter into relations, which are the precise formalization of the laws of the theory. As such their mathematical characteristics (scalars, vectors, tensors, matrix,..) and properties (continuity, differentiability,..) are essential to any computation, in order to forecast the result of experiments and check the validity of the laws

- they are used to organize the collect of experimental data, through precise formats, sampling and statistical methods.

For instance a model in electromagnetism would study the trajectory $x(t)$ of a body with mass m and charge q , under an electromagnetic field (E,B) in some frame. The variables are the functions x , E , B . The quantities m,q are parameters.

In Economics a model would study the relation between the output Y of a branch of industry with respect to the labor L and the capital K , over a period of time $[0, T]$. The variables are the functions of t : Y, L, K .

The state of the system is fully known if we know the variables, and the set \mathcal{S} of all possible states is defined by the model itself, as it has been chosen by its author. Due to the potential relations between the variables, one will observe only a part of this set, however, in any experimental procedure, this concentration is a fact to be checked, and the whole of the space of configuration should be considered a priori. In this paper we will not be concerned by the potential relations between the variables, but we will focus on the set of possible states.

If the model meets some very general properties, essentially if the variables are vectors of an infinite dimensional vector space, the set \mathcal{S} acquires a specific mathematical structure : we will prove in the first section that it is in correspondance with an open subset of a Hilbert space H . So each state is associated to a vector ψ of a most convenient mathematical structure, and there is a simple correspondance between the value X of the variables and the vector ψ in the Hilbert space H .

It is theoretically possible to measure the value of any variable, and thus the state of the system. However there are limitations, beyond the usual imprecisions of any measure, due to the discrepancy between variables defined in infinite dimensional spaces and the available data, which are in finite number. So the value of the state is estimated from a finite batch of data by a statistical method using a simplified version of X (a specification of the function among the space V such as polynomial or affine map). In the second section we define observables, quantities which can be measured on the space \mathcal{S} to estimate the value of the state. We show that to any observable is associated a self-adjoint operator on the Hilbert space, which restricts the values that can be actually observed . Moreover the measure of any observable is affected by an imprecision which can be represented by a probability law, that can be considered as the measure of the imprecision coming from the choice of a specification.

A system can usually be represented equivalently by different sets of variables X, X' . Whenever there are two such related sets of variables $X' = U(X)$, we show in the 4th section that there is a unitary operator \hat{U} relating the vectors ψ, ψ' associated to the states represented by each set of variables. This result has far reaching consequences as, whenever the change of variable depends on a group, the Hilbert space shall be a unitary representation (H, \hat{U}) of the group.

Models commonly involve variables which depend on the time, and represent the evolution of the system. Under some general conditions we prove that the set of values taken at each time by the variables can be endowed by a structure of Hilbert space. This result is specially useful to study systems whose evolution shows distinct phases and we give a general method to build indicators to estimate the probability of a transition.

In the last section we consider interacting systems. The usual solution is to introduce a different model for each system, with specific variables represent-

ing the interactions. However we show that, under general specifications, it is possible to represent the two systems in the same model, by taking the tensorial product of the variables. We discuss the meaning of this representation, specially with respect to the common probabilist interpretation which considers that each microsystem can occupy randomly different states. If the system is homogeneous, comprised of undistinguishable microsystems obeying the same laws, then the Hilbert space associated to the system is specific, and in bijective correspondance with the classes of conjugacy of the symmetric group. This leads to the introduction of the concept of entropy. The results can be extended for the temporal evolution of systems, and to the study of transitions of systems between different phases.

The proofs use many theorems and definitions in various mathematical fields, which may be new to the reader. So the most expedient solution has been to refer to the extensive summary of mathematical results that I have written and is available on the same server. The references are denoted as (JCD Th.XX).

1 HILBERT SPACE

1.1 Hilbert space

We need first to precise what are the characteristics of the models which are studied in this paper. They are summarized in :

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 as $X_k = R(X_l)$. For instance X_k cannot be a ratio such that $\sum_k X_k = 1$. But the relations which are tested in the model do not matter, only the definition of the variables. Moreover the derivative $\frac{dX_k}{dt}$ of a variable X_k is considered as an independant variable.

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. For instance a model which studies the consumption $X_k(R)$ of products $k = 1 \dots N$ with respect to the income R of a household : the variables are then the function X_k and not the scalar value of $X_k(R)$.

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.

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 \mathcal{S} can be embedded as an open subset $\Omega \subset H$ which contains 0 and a convex subset.*

Proof. i) Each value of the set \mathcal{S} of variables defines a state of the system, denoted X , belonging to the product $O = \prod_{k=1}^N O_k \subset V = \prod_{k=1}^N V_k$. The couple (O, X) , together with the property iv) defines a 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 Fréchet space, the Henderson's theorem (Henderson - corollary 5, JCD 1386) states that it can be embedded as a open subset Ω of an infinite dimensional separable Hilbert space H , defined up to isomorphism. Moreover this structure is smooth, the set $H - \Omega$ is homeomorphic to H , the border $\partial\Omega$ is homeomorphic to Ω and its closure $\bar{\Omega}$.

iii) Let us denote $\langle \cdot \rangle_H$ the scalar product on H (this is a bilinear symmetric positive definite form). The map $\psi : \Omega \rightarrow \mathbb{R} :: \langle \psi, \psi \rangle_H$ 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.

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 .

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 .

1.2 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 :

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

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

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

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

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

$H_1^\perp = 0$ thus H_1 is dense in H (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 :

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

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 :

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

$$\forall i \in I, e_i \in O \Rightarrow \Xi(e_i) = \varepsilon_i = \sum_{j \in I} \langle \tilde{\varepsilon}_j, \varepsilon_i \rangle_H \tilde{\varepsilon}_j \text{ and } \tilde{\pi}_H(\varepsilon_i) = (\langle \tilde{\varepsilon}_j, \varepsilon_i \rangle_H)_{j \in I} \in \tilde{R}_0$$

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 :

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

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

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

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

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 :

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

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

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.

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

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

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

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

$\psi_i = \sum_{j \in J} [\mathcal{L}_H]_i^j \tilde{\psi}_j \Rightarrow |\psi_i| < \varepsilon \sum_{j \in I} \max |[\mathcal{L}_H]_i^j|$ and $\left(\left| [\mathcal{L}_H]_i^j \right| \right)_{j \in I}$ 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 $\bar{\varpi}_i \in H'$ 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 : $\forall \psi \in H : \bar{\varpi}_i(\psi) = \langle \phi_i, \psi \rangle_H$ so that :

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

$\forall i \in I :$

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

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

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 . ■

1.2.1 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 .

1.3 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 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 :

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

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

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

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

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

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

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

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

Thus γ is definite positive ■

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 : $\hat{\Phi} : H \rightarrow H :: \hat{\Phi} = \Upsilon \circ \Phi \circ \Upsilon^{-1}$ and $\hat{\Phi}$ is an operator on H . And conversely.

2.1 Primary observables

The simplest specification 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 4 To any primary observable Y_J is associated uniquely a self-adjoint, compact, trace-class operator \hat{Y}_J on $H : Y_J = \Upsilon^{-1} \circ \hat{Y}_J \circ \Upsilon$ such that the measure $Y_J(X)$ of the primary observable Y_J , if the system is in the state $X \in O$, is

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

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 $(\varepsilon_j)_{j \in J}$.

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 5 *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 $(\widehat{\varepsilon}_i)_{i \in J}$ of H_J starting with the vectors $(\varepsilon_i)_{i \in J}$ and an orthonormal basis $(\widehat{\varepsilon}_i)_{i \in J^c}$ of H_{J^c} starting with the vectors $(\varepsilon_i)_{i \in J^c}$

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

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

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

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

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

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

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

So the operators commute. ■

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

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

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

Let us define the scalar product on W :

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

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

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

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

3. Let us define A as the algebra generated by any finite linear combination or products of elements \widehat{Y}_J, J finite or infinite, and \overline{A} as the closure of A in

$$\mathcal{L}(H; H) : \overline{A} = \text{Span} \left\{ \widehat{Y}_J \right\}_{J \subset I}$$

Proposition 6 \overline{A} is a commutative von Neumann algebra of $\mathcal{L}(H, H)$

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

Because its generators are projections, \overline{A} is a von Neumann algebra (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 closure is with respect to the strong topology, that is in norm. 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.

However we will not pursue this avenue and look for more general maps Φ using spectral theory.

2.3 Secondary observables

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 : $\widehat{\pi}_{\chi(\varpi)}$

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

2. An application of standard theorems on spectral measures (JCD Th.1243, 1245) tells that, for any bounded measurable function $f : E \rightarrow \mathbb{R}$, the spectral

integral : $\int_E f(\xi) \widehat{\pi}_{\chi(\xi)}$ defines a continuous operator $\widehat{\Phi}_f$ on H . $\widehat{\Phi}_f$ is such that

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

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

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

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

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

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

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

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

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

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

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

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

3. We will define :

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

Proposition 7 Any secondary observable Φ is a compact, continuous map, its associated map $\widehat{\Phi} = \Upsilon \circ \Phi \circ \Upsilon^{-1}$ is a compact, self-adjoint, Hilbert-Schmidt and trace class operator.

$$\Phi = \sum_{p=1}^n \lambda_p Y_{J_p} \text{ where } (J_p)_{p=1}^n \text{ are disjoint finite subset of } I$$

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

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

$\Phi, \widehat{\Phi}$ are compact and thus continuous (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 $H_c = (\oplus_{p=1}^n H_p)^\perp$
 $\forall \psi \in H : \psi = \psi_c + \sum_{p=1}^n \psi_p$ 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 $J^c = \complement_I (\oplus_{p=1}^n J_p)$ for the basis of $V_c = \text{Span}((e_i)_{i \in J^c})$

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

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

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

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

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

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

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

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

To sum up :

Proposition 8 *For any primary or secondary observable Φ , there is a basis $(e_i)_{i \in I}$ of V, a self-adjoint, Hilbert-Schmidt and trace class operator $\widehat{\Phi}$ on the associated Hilbert space H such that :*

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

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

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 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 :

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

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

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

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

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

So :

Proposition 9 *For any secondary observable there is always a primary observable which is at least as efficient.*

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

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

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

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 $\widehat{Y}_J(\psi)$ is a random variable, with a distribution law $\widehat{\mu}_\psi$.

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

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

So we have :

Proposition 10 *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 :*

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

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 :

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

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

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

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 $\Pr(J_p|X) = \frac{\|\hat{\gamma}_p \psi\|^2}{\|\psi\|^2} = \frac{\|\psi_p\|^2}{\|\psi\|^2}$

And we have :

Proposition 11 *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 $\hat{\Phi}$ 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 $\frac{\|\psi_p\|^2}{\|\psi\|^2}$. But this gives a probabilistic behaviour to the system (X becoming a random variable) which is not present in its definition.

4 CHANGE OF VARIABLES

4.1 Fundamental theorem

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. This is the usual case in Physics, and, 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 assumptions :

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 have $V = V'$ but the open O, O' can be different.

Proposition 12 *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 subvector spaces $H_0 = \text{Span}((\varepsilon_i)_{i \in I}), H'_0 = \text{Span}((\varepsilon'_i)_{i \in I})$

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 :

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

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 :

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

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 :

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

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

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

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

4.2 Observables

As \widehat{U} is unitary, it cannot be self adjoint or trace class. So from this respect 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 : $\widehat{\Phi} = \Upsilon \circ \Phi \circ \Upsilon^{-1}$. The map Υ depends only on the observable, and not X .

For the new variable the observable is $\Phi' = \Phi \circ U$ and it is associated to the operator :

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

4.3 Change of units and projective representations

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.

However one can have another interpretation. We have seen that an infinite dimensional Hilbert space can always be endowed with a complex structure. V stays a real vector space, but then we can be more flexible in the choice of the linear map $\Upsilon : V \rightarrow H$. Its definition relies on the fundamental identities :

$$\forall X \in O : \Upsilon(X) = \sum_{i \in I} \langle \phi_i, \Upsilon(X) \rangle_H \varepsilon_i \in H_0$$

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

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

So $z\Upsilon$, where $|z| = 1$ still defines a chart which is compatible, both with the manifold and the complex Hilbert space structure. The positive kernel stays the same on $V : K_V(X, Y) = K(\Xi(X), \Xi(Y)) = K(z\Xi(X), z\Xi(Y)) = z\bar{z}K(\Xi(X), \Xi(Y))$ with a hermitian form on H .

Similarly the spectrum of the unitary operator \widehat{U} , in the theorem above, comprised of its eigen values (0 is excluded), is a subset of the complex numbers of module 1 (JCD Th.1062). So with a complex structure the set of its eigen vectors is larger than in the real case.

A change of variable $\widehat{U}\psi = z\psi$ gives a map $U : UX = X$: the same state X can be represented by any vector of H , up to multiplication by a complex number of module 1 (called a "ray") : this is a projective representation.

Notice that the previous result still holds : the quantities must be unitless, and the projective representation is not necessary. Anyway the Hilbert space H is defined up to an isometry.

4.4 Application to groups of transformations

Usually the changes of variables are a fundamental feature of the model, and have special properties related to a group G . It is then possible to precise the nature of the abstract Hilbert space H .

4.4.1 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 :

$$\forall \kappa \in T_1 G : \left(\widehat{U}'(1) \kappa \right)^* = - \left(\widehat{U}'(1) \kappa \right)$$

$\widehat{U}(\exp \kappa) = \exp \widehat{U}'(1) \kappa$ where the first exponential is taken on $T_1 G$ 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).

The observables are parametrized by g :

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

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

4.4.2 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 P 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).

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

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

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

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.

4.4.3 One parameter groups

An important case, related to the previous one, is if there is a map $\widehat{U} : \mathbb{R}_+ \rightarrow GL(H, H)$ such that :

$$\begin{aligned}\widehat{U}(t+t') &= \widehat{U}(t) \circ \widehat{U}(t') \\ \widehat{U}(0) &= Id\end{aligned}$$

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 :

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

S is normal and has a spectral resolution P :

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

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

5 THE EVOLUTION OF THE SYSTEM

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

i) t is just a parameter used to identify a state of the system : t gives its temporal location. Then the origin does not matter : if we take 1/1/2000 of 1/1/2010 we identifies the same state of the system.

ii) t is a parameter used to measure the duration of a phenomenon, usually the time elapsed since some specific event. Then of course the origin matters, and it is usually imposed.

The first case appears notably in theoretical Physics. The nature of time and its relations with spatial coordinates is a deep topic in Physics, upon which we will not dwell, but when t is understood as a coordinate to identify an event in the Universe the origin does not matter (this is an axiom in Galilean Geometry and, with some precaution regarding the choice of maps, this is still true in General Relativity). More generally the evolution of the system is assumed to be smooth, in the meaning that its evolution over any period of time $[t_1, t_2]$ is similar (but not necessarily identical) to its evolution over $[t_1 + T, t_2 + T]$: it follows the same rules. There is no privileged period in time.

In the second case it is common that the problem is "well posed" : the state of the system at t depends only on the state at t = 0. It happens when the

evolution of the system answers to hyperbolic differential equations. Then the variables belong to some family of maps such that the map is defined uniquely in the family by its initial value.

We will address both cases, with precise conditions.

To keep it simple we will call "time" the privileged argument t , but of course t can represent any other argument, in the same conditions. The variables can depend on other arguments (or some may not depend at all on t) and, as above, we will denote $X(t) = (X_k(t))_{k=1}^N$ the vector of the values of X at the time t . If the value of some variables is imposed they are considered as any other variable, and subject to the same conditions.

5.1 Fundamental theorems

5.1.1 Conditions 3

They address the case 2 with a well posed problem.

Condition 3 :

The variables $(X_k)_{k=1}^N$ are maps :

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

ii) The maps X belong to an open subset O of an infinite dimensional Fréchet space V

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

iv) $\forall t \in R : X(t) = X'(t) \Rightarrow X = X'$

We will assume that $0 \in R$ (but it can be any fixed point of R).

Proposition 13 *If the conditions 3 are met, then there are a Hilbert space F , an open $\tilde{O} \subset F$, a map : $\Theta : R \rightarrow \mathcal{L}(F; F)$ such that :*

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

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

$\Theta(t)$ is unitary

Proof. Let us first assume that $O = V$.

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 :

the evaluation map : $\mathcal{E}(t) : V \rightarrow F(t) : \mathcal{E}(t) X = X(t)$ is linear, onto,

it is injective : iii) reads : $\forall t \in R : \mathcal{E}(t) X = \mathcal{E}(t) X' \Rightarrow X = X'$

so it is bijective and continuous at $X = 0$ so continuous (V is metric).

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

Define on $F(0)$ the positive kernel :

$$K_{F(0)}(f, f') = K_V(\mathcal{E}(0)^{-1}f, \mathcal{E}(0)^{-1}f') = K_V(X, Y) = \langle \Upsilon X, \Upsilon Y \rangle_H$$

Let $(f_j)_{j \in J}$ be a basis of $F(0)$: $u \in F(0) \Leftrightarrow u = \sum_{j \in J} x_j f_j, (x_j)_{j \in J} \in \mathbb{R}_0^J$ and defines the bilinear symmetric definite positive form with coefficients : $\langle f_j, f_k \rangle_{W(0)} = K_{F(0)}(f_j, f_k)$

By the Gram-Schmidt procedure we can build an orthonormal basis $(\tilde{f}_j)_{j \in J}$ of $F(0)$: $F(0) = \text{Span}(\tilde{f}_j)_{j \in J}$ and the Hilbert vector space : $F = \left\{ \sum_{j \in J} \tilde{x}_j \tilde{f}_j, (\tilde{x}_j)_{j \in J} \in \ell^2(J) \right\}$

Define the map :

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

This is a bijective, continuous, linear map. The vector spaces are isomorphic, so one can take $F(0)$ as the common structure and $\Theta(t)$ is a bijective, continuous endomorphism $\Theta(t) \in \mathcal{L}(F(0); F(0))$ which can be extended to F .

$\Theta(t)$ is unitary on $F(0)$: $\langle u, v \rangle_{F(0)} = K_V(\mathcal{E}(0)^{-1}u, \mathcal{E}(0)^{-1}v)$. The bilinear map $\langle \Theta(t)u, \Theta(t)v \rangle_{F(0)}$ has a unique extension on F so $\Theta(t)$ is unitary on F .

$F(0)$ is an open subset of F , which is dense in F . $\mathcal{E}(t)$ is linear, bijective and continuous, this is a homeomorphism, the image $\mathcal{E}(0)$ of an open subset O of V is an open subset \tilde{O} of $F(0)$, that is of F . ■

Proposition 14 *If the conditions 1 are met there is a map : $\hat{\mathcal{E}} : R \rightarrow \mathcal{L}(H; F)$ such that :*

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

$\hat{\mathcal{E}}(t)$ is an isometry

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

Define the map :

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

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

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

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

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

Remark :

There is no continuity condition on the maps X , and similarly $\Theta(t)$ can be discontinuous. However the continuity of $\mathcal{E}(t)$ is crucial. Practically it involves some relation between the semi-norms on V and the norm on E . $\mathcal{E}(t)$ 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) $d(X, 0)_V \rightarrow 0 \Rightarrow \forall t \in \mathbb{R} : \|X(t)\|_E \rightarrow 0$ and the condition iii) is met, but this is not a general result.

5.1.2 Conditions 4

They address the case 1, which, as we will see, is a special case of the previous one.

Conditions 4 :

The variables $(X_k)_{k=1}^N$ are maps :

- i) $X = (X_k)_{k=1}^N :: \mathbb{R} \rightarrow E$ where E is a normed vector space
- ii) The maps X belong to an open subset O of an infinite dimensional Fréchet space V
- iii) The evaluation map : $\mathcal{E} : \mathbb{R} \times V \rightarrow F :: \mathcal{E}(t) X = X(t)$ 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 $\theta \in \mathbb{R}$

Proposition 15 *If the conditions 4 above are met, then 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

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

Proof. 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.

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 :

$$\begin{aligned}\widehat{U}(\theta) &= \sum_{n=0}^{\infty} \frac{\theta^n}{n!} \widehat{S}^n = \exp \theta \widehat{S} \\ \frac{d}{ds} \widehat{U}(s) |_{\theta=s} &= \left(\exp \theta \widehat{S} \right) \circ \widehat{S} \\ \widehat{S} &= \int_{Sp(S)} sP(s) \\ \left\| \widehat{U}(\theta) \right\| &= 1 \leq \exp \left\| \theta \widehat{S} \right\|\end{aligned}$$

$S = \Upsilon^{-1} \circ \widehat{S} \circ \Upsilon$ 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 :

$$\begin{aligned}U(\theta) &= \Upsilon^{-1} \circ \widehat{U}(\theta) \circ \Upsilon = \sum_{n=0}^{\infty} \frac{\theta^n}{n!} S^n = \exp \theta S \\ \forall \theta, t \in \mathbb{R} : U(\theta) X(t) &= X(t + \theta) = (\exp \theta S) X(t) \\ \mathcal{E}(t) \exp \theta S &= \mathcal{E}(t + \theta)\end{aligned}$$

Exchange θ, t and take $\theta = 0$:

$$\begin{aligned}\mathcal{E}(\theta) \exp tS &= \mathcal{E}(t + \theta) \\ \mathcal{E}(0) \exp tS &= \mathcal{E}(t) \in \mathcal{L}(V; F)\end{aligned}$$

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

:

$$\begin{aligned}\frac{d}{ds} U(s) X(0) |_{s=t} &= \frac{d}{ds} X(s) |_{s=t} = SX(t) \\ \Leftrightarrow \frac{d}{ds} \mathcal{E}(s) |_{s=t} &= S\mathcal{E}(t)\end{aligned}$$

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

Proposition 16 *If the conditions 4 are met, then there are a Hilbert space F , an open $\tilde{O} \subset F$, a continuous anti-hermitian map $\tilde{S} \in \mathcal{L}(F; F)$ such that :*

$$\begin{aligned}\forall X \in O \subset V : X(0) &\in \tilde{O} \subset F \\ \forall t : X(t) &= \left(\exp t\tilde{S} \right) (X(0)) \in F\end{aligned}$$

The maps X are smooth and :

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

Proof. 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 :

$$\begin{aligned}\forall t \in \mathbb{R} : X(t) &= (\exp tS \circ X)(0) \\ X(t) = X'(t) &\Rightarrow \forall \theta : X(t + \theta) = X'(t + \theta) \Leftrightarrow X = X'\end{aligned}$$

So the conditions 3 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.

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

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 in the conditions 3 :

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

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

The probability issues that we have seen previously take a more obvious meaning. The genuine variable is a map depending on t , the error which is made in the specification of X translates in systematic errors in the value of $X(t)$, which give the feeling that $X(t)$ itself is a random variable, following some precise probability law.

If the conditions 3 above are met then the estimation can be focused on finding an estimator of $\Theta(t)$. This is the starting point of a large class of statistical methods. First one restricts E to a finite dimensional vector space N , then the problem is to find an orthonormal basis of functions belonging to $L^2(R, N, dt)$ for a definite positive kernel K_N , which is usually chosen among simple functions.

If the conditions 4 above are met we have a great advantage, because we have already a good part of the specification. The estimation can be focused on the linear operator \tilde{S} from the relation : $\frac{d}{ds}X(s)|_{s=t} = \tilde{S}X(t)$. Moreover we have not made any assumption about the dimension of E (and then F) which can be finite. The problem then sums up to a classic linear estimation.

5.3 Stationary solutions

A classic problem of statistical mechanics is to look for specific configurations of the set of solutions $X(t)$, in particular for periodic or stationary solutions. We have some significant results for systems which meet the conditions above.

A solution is stationary if $X(t) = Ct$.

i) In the conditions 3 : $X(t) = \Theta(t)X(0) = X(0)$.

$X(0)$ must be an eigen vector with eigen value 1 of $\Theta(t)$ for all t in R . This is the vector subspace L of F of invariant vectors by $\{\Theta(t), t \in R\}$.

$\Theta(t)$ is unitary, so its eigen values belong to the unit circle, it may have the eigen value 1. Because X is defined by its initial value, one can say that, if it exists, a stationary state is unique.

From the Alaoglu-Birkhoff theorem one can tell (JCD Th.1157) that :

$$\forall u \in F, \forall t \in R : \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{p=0}^n (\Theta(t))^p u = P_L(u)$$

where P_L is the orthogonal projection of F on L .

ii) In the conditions 4 : $X(t) = Ct \Leftrightarrow \tilde{S}X(0) = 0 \Leftrightarrow X(0) \in \ker \tilde{S}$ then as above X is fully defined and $L = \ker \tilde{S}$.

6 DISCRETE VARIABLES AND PHASES TRANSITIONS

6.1 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) which coexist simultaneously in the same global system, follow different rules of behaviour, but interact together. We will see in the next section 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. Conversely one can use a set of quantitative data to estimate a discrete variable featuring these distinct behaviours, as in factorial analysis.

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 assumptions :

Conditions 5 :

Conditions 1 + the set O of possible states of the system has two connected components O_1, O_2

Proposition 17 *If the condition 5 is met 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 the conditions 5, to replace a discrete variable by a quantitative variable with the same features.

6.2 Phases Transitions

There is a large class of problems which, in another way, involve discrete quantities : transitions in the evolution of a system. They do not involve the maps X but the values $X(t)$ which are taken over a period of time in some vector space E . When one can distinguish specific subsets of E , that we will call phases (to avoid any confusion with states which involves the map X). The conditions in which these transitions happen are of special interest. Common cases in Physics are change of phases for solid or liquid bodies, the desintegration of a particle,..., in Economics a crisis or a recession, in Finances a rupture in the markets,...

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

Staying in the general model as in proposition 1, the variables X are maps which belong to an open $O \subset V$ of maps defined over some interval $R \subset \mathbb{R}$.

The first issue is the definition of the phases. The general idea is that they are significantly different so it makes sense to keep the same definition as above : 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 the conditions 4 certainly are not met, but the conditions 3 can hold (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 have 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 $X(t)$ stays, 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 , and clearly the conditions 4 do not apply.

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 $\hat{p}(y) = \frac{n(y)}{G(y)}$.

7 INTERACTING SYSTEMS

7.1 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 in the theory. To account for the interactions between the two systems the models are of the form :

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

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.

One can consider the direct product $S_1 \times S_2$, but then usually the variables Z_1, Z_2 are explicitly related, at least one should be dropped and doing so we obviously miss the interactions Z_1, Z_2 . Moreover it is not always obvious to measure these variables.

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

We consider the models without interactions. 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 6 :

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 18 *Whenever two systems S_1, S_2 interact, there is a model S encompassing the two systems and meeting the conditions 6 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. It is an extension of the common usage, for discrete variables $X_k \in \{x_{k1}, \dots, x_{kN_l}\}$, to consider the variable $Y \in \{(x_{11}, \dots, x_{kN_l}), k = 1 \dots p\}$ which takes in count the simultaneous values of each variable and is their discrete tensorial product with $N_1 \times \dots \times N_l$ positions.

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. So it requires the simultaneous knowledge of all the variables in the system.

7.1.1 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, notably in Quantum Mechanics.

The knowledge of the states (X_1, X_2) of both systems requires two vectors of I components each, that is $2 \times I$ scalars, and the knowledge of the state S requires a vector of I^2 components. So the measure of S requires more data, and brings more information, because it encompasses all the interactions. Moreover *a tensor is not necessarily the tensorial product of vectors* (if it is so it is said to be 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, and 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.

7.2 Population of N Similar Interacting Systems

An interesting case is many (and usually a variable number) similar systems, represented by the same model, interacting together. In a first step we will make the additional assumption that the number of individual systems (that we will call a microsystem) is fixed.

Conditions 7 :

There are N (a fixed and usually large number) systems $s = 1 \dots N$ which are represented by the same model, with variables $(X_s)_{s=1}^N$ satisfying the conditions 1 : X_s belong to an open O of an infinite dimensional separable Fréchet space V .

Notice that we do not require any other property from each individual system : similar means only that they can be represented by the same model, using the same variables, but each system is identified by a label s , and in particular they can have different "sizes".

For each microsystem the Hilbert space H and the linear map Υ are the same, but the vectors representing the states are different quantities (usually related by some laws defined over the whole population). We can apply the previous result : then 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. It is assumed that it is possible to identify each microsystem, which are labeled by $s = 1 \dots N$. 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 $(X_s)_{s=1}^N$.

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).

As said above for two microsystems, Ψ is not necessarily the tensorial product of vectors of H , so usually the knowledge of the state of the system requires more than the knowledge of the state of each system.

7.2.1 Observables of interacting systems

Observables can be devised in two ways.

First, as previously, as estimators of the state S : observables are then linear map from \mathbf{V}_N to some finite dimensional vector subspaces of \mathbf{V}_N . They are built by tensorial product of observables on V , which are associated to self adjoint operators on H , which are themselves extended to \mathbf{H}_N : $\hat{\Phi}_N = \Upsilon_N \circ \Phi_N \circ \Upsilon_N^{-1}$

Such observables are the only ones, meeting our usual requirements of linearity and compactness, which give a measure of the state of the total system, encompassing the interactions.

For primary observables a fixed subset J of I is chosen, from where there is a hilbertian basis $\tilde{\varepsilon}_{i_1} \otimes \dots \otimes \tilde{\varepsilon}_{i_N}$ with $i_1, \dots, i_N \in J$ and the value which is measured is the projection of Ψ on this basis. The probability to measure $\Phi_N(S)$ if the system is in the state S is :

$$\Pr(\Phi_N(S) | S) = \frac{\|\hat{\Phi}_N(\Upsilon_N(S))\|_{\mathbf{H}_N}^2}{\|\Upsilon_N(S)\|_{\mathbf{H}_N}^2}$$

Quantities such as $\Phi_N(S)$ are difficult or impossible to measure directly for large systems. It is common to use an indirect formalism to help to visualize S . The method is the following.

As for primary observables a finite subset J of I defines a finite dimensional vector subspace H_J of H and the projections $\hat{Y}_J : H \rightarrow H_J$

This projection extends to

$$\hat{\pi}_J : H_N \rightarrow \otimes_{s=1}^N H_J :: \hat{\pi}_J \left(\sum_{i_1 \dots i_N \in I} \psi^{i_1 \dots i_N} \tilde{\varepsilon}_{i_1} \otimes \dots \otimes \tilde{\varepsilon}_{i_N} \right) = \sum_{i_1 \dots i_N \in J} \psi^{i_1 \dots i_N} \tilde{\varepsilon}_{i_1} \otimes \dots \otimes \tilde{\varepsilon}_{i_N}$$

There is no canonical map :

$$\mathbf{H}_N = \otimes_{s=1}^N H \rightarrow H^N = \prod_{s=1}^N H$$

However one can look for a set of vectors $(\varphi_s)_{s=1}^N \in H_J^N$ such that for $\psi \in \otimes_{s=1}^N H_J$ the tensor $\varphi_1 \otimes \varphi_2 \dots \otimes \varphi_N$ gives a good approximation of ψ . With $\varphi_s = \sum_{j \in J} A_s^j \tilde{\varepsilon}_j : \varphi_1 \otimes \varphi_2 \dots \otimes \varphi_N = \sum_{j_1 \dots j_N \in J} A_1^{j_1} \dots A_N^{j_N} \tilde{\varepsilon}_{j_1} \otimes \dots \otimes \tilde{\varepsilon}_{j_N}$

So this sums up to find NJ scalars such that $A_1^{j_1} \dots A_N^{j_N} = \psi^{j_1 \dots j_N}$. As the number of components $\psi^{j_1 \dots j_N}$ is J^N usually there is no exact solution : this is an extension of the classic statistical problem to estimate a cross distribution by marginal distributions, and there are classic estimators. The usual interpretation of the result is probabilist : the state of each microsystem is seen as a random variable with expected value φ_s which combine independantly to give ψ . Of course the method can be transposed from ψ, H to X,V. However we have to keep in mind that the probability which is introduced is conventional, and does not imply anything about the actual behaviour of real systems.

This method is of particular interest when one considers global variables : global quantities measured over the whole of the system. They can be defined as maps $G : \mathbf{V}_N \rightarrow F$ valued in some finite dimensional vector space F, which is not necessarily a subspace of \mathbf{V}_N (it can be a scalar). As usual G denotes all the possible partial maps with the same properties. As such they provide only a limited information about the state S. The map G is transposed as : $\hat{G} : \mathbf{H}_N \rightarrow F : \hat{G} = G \circ \Upsilon_N^{-1}$.

Using the method above then the estimator of ψ is submitted to the constraints :

$$\hat{G} \left(\sum_{i_1 \dots i_N \in I} \psi^{i_1 \dots i_N} \tilde{\varepsilon}_{i_1} \otimes \dots \otimes \tilde{\varepsilon}_{i_N} \right) = \hat{G} (\varphi_1 \otimes \varphi_2 \dots \otimes \varphi_N)$$

Because the value of $G(S)$ provides additional information it increases the quality of the previous estimator of S by separable tensors. This construct has the advantage to give a simple picture. It has obvious merits when the number N of microsystems is large : the measure of the tensor S is difficult or impossible, and the knowledge of the individual states and of the interactions is of limited interest. This is all the more interesting when the microsystems are identical : they follow the same laws. Then the probability can be seen as the proportion of microsystems which have a given state, compatible with the measured quantity G as we are going to see.

7.3 Homogeneous systems

7.3.1 Definition

So far each microsystem is identified by a label s, and the state $S = X_1 \otimes \dots \otimes X_N$ is deemed different from $S = X_{\sigma(1)} \otimes \dots \otimes X_{\sigma(N)}$. 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 :

Conditions 8 :

In addition to the conditions 7, the microsystems are indistinguishable : any permutation of the N microsystems gives the same state of the total system.

Proposition 19 *The states Ψ of homogeneous systems belong to an open subset of the Hilbert space*

$$\mathbf{h} = \odot_{n_1} H_J \odot_{n_2} H_J \dots \odot_{n_p} H_J$$

where

$$0 \leq n_p \leq \dots \leq n_1 \leq N, n_1 + \dots + n_p = N$$

\odot is the symmetric tensorial product

$H_J = \widehat{Y}_J H$ is a N dimensional vector subspace of H

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.

If the conditions 8 are met, then we can implement the results proven previously. To U is associated a unitary operator \widehat{U} on \mathbf{H}_N such that $(\mathbf{H}_N, \widehat{U})$ is a unitary representation of $\mathfrak{S}(N)$.

The action of \widehat{U} on H_N is defined uniquely by linear extension of $\widehat{U}(\sigma)(\psi_1 \otimes \dots \otimes \psi_N) = \psi_{\sigma(1)} \otimes \dots \otimes \psi_{\sigma(N)}$ on separable tensors. $\Psi \in H_N$ reads in a Hilbert basis $(\tilde{\varepsilon}_i)_{i \in I}$ of H : $\Psi = \sum_{i_1 \dots i_N \in I} \Psi^{i_1 \dots i_N} \tilde{\varepsilon}_{i_1} \otimes \dots \otimes \tilde{\varepsilon}_{i_N}$ and :

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

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 $\mathbf{h} \subset H_N$ 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 $\mathbf{h} \subset \widehat{Y}_{J_N} H_N$ 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 $\tilde{\varepsilon}_{j_1} \otimes \dots \otimes \tilde{\varepsilon}_{j_N} \in \mathbf{h}, \forall \sigma \in \mathfrak{S}(N) : \widehat{U}(\sigma) \tilde{\varepsilon}_{j_1} \otimes \dots \otimes \tilde{\varepsilon}_{j_N} = \tilde{\varepsilon}_{j_{\sigma(1)}} \otimes \dots \otimes \tilde{\varepsilon}_{j_{\sigma(N)}} \in \mathbf{h}$
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} his 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) 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) Any tensor of \mathbf{h} is a linear combination of the separable tensors $\otimes_{n_1} \tilde{\varepsilon}_{j_1} \otimes_{n_2} \tilde{\varepsilon}_{j_2} \dots \otimes_{n_p} \tilde{\varepsilon}_{j_p}$ which can be seen as representing a configuration where n_k mi-

crossystems are in the same state $\tilde{\varepsilon}_{j_k}$. 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) 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 : $\Psi \in \odot_n H_J \Leftrightarrow S_n(\Psi) = n! \Psi$

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

7.3.2 Global observables of homogeneous systems

As global observable is defined as a linear map $G : \mathbf{V}_N \rightarrow F$ valued in some finite dimensional vector space F,

Proposition 20 *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 $\hat{G} : \odot_n H_J \rightarrow E :: \hat{G} = G \circ \Upsilon_N^{-1}$*

$\hat{G}(\Psi) = k \hat{G}(\otimes_{n_1} \tilde{\varepsilon}_{j_1} \otimes_{n_2} \tilde{\varepsilon}_{j_2} \dots \otimes_{n_p} \tilde{\varepsilon}_{j_p})$ 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. \hat{G} must be symmetric for a homogeneous system : $\forall \Psi \in \mathbf{h}, \forall \sigma \in \mathfrak{S}(N) : \hat{G} \circ \hat{U}(\sigma) \Psi = \hat{G} \Psi$

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 :

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

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

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

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

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

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

The map \hat{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

$\left\{ \widehat{G} \left(\otimes_{n_1} \widetilde{\varepsilon}_{j_1} \otimes_{n_2} \widetilde{\varepsilon}_{j_2} \dots \otimes_{n_p} \widetilde{\varepsilon}_{j_p} \right), \lambda, (\widetilde{\varepsilon}_{j_k})_{k=1}^p \right\}$ 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 $k, \lambda, (\widetilde{\varepsilon}_{j_k})_{k=1}^p$. 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 $\widetilde{\varepsilon}_{j_k}$ 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 $\widehat{G} \left(\otimes_{n_1} \varepsilon_{j_1} \otimes_{n_2} \varepsilon_{j_2} \dots \otimes_{n_p} \varepsilon_{j_p} \right)$ 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 $\widetilde{\varepsilon}_{j_k}$. So, from a practical point of view, this is a vector $\gamma = \widehat{G} \left(\otimes_{n_1} \widetilde{\varepsilon}_{j_1} \otimes_{n_2} \widetilde{\varepsilon}_{j_2} \dots \otimes_{n_p} \widetilde{\varepsilon}_{j_p} \right)$ which is measured, and from it $\lambda, (\widetilde{\varepsilon}_{j_k})_{k=1}^p$ are estimated.

7.3.3 Entropy

Let G be a global variable which encompasses all the global information which is available on the system. For the representation of the states of the system we can pick an arbitrary hermitian basis $(\widetilde{\varepsilon}_i)_{i \in I}$ of H , any subset J of cardinality N of I , then the map $\widehat{G} = G \circ \Upsilon_N$ is defined as above. The information on the system is summed up in $\gamma = \widehat{G}(\Psi) = \widehat{G} \left(\otimes_{n_1} \widetilde{\varepsilon}_{j_1} \otimes_{n_2} \widetilde{\varepsilon}_{j_2} \dots \otimes_{n_p} \widetilde{\varepsilon}_{j_p} \right)$ for unknown class λ and family $(\widetilde{\varepsilon}_{j_k})_{k=1}^p$. As J is arbitrary we will denote $(\varepsilon_j)_{j=1}^N$ the basis $(\widetilde{\varepsilon}_i)_{i \in J}$ of H_J and the family $(\widetilde{\varepsilon}_{j_k})_{k=1}^p$ is then the choice of non ordered p integers in $1 \dots N$. So for a given class λ there are C_N^p possible different bases, and the dimension of \mathbf{h} is given by the number of permutations $\frac{N!}{n_1! \dots n_p!}$.

We have seen previously that it is possible to give a more illuminating representation by a probabilist model. In this picture it is assumed that it is possible to define and measure the state of each microsystem. Then the configuration $(\otimes_{n_1} \varepsilon_{j_1} \otimes_{n_2} \varepsilon_{j_2} \dots \otimes_{n_p} \varepsilon_{j_p})$ is interpreted as n_k microstates are in the state represented by ε_{j_k} . And one attributes a probability to each configuration. In this picture it is assumed that each microsystem behaves independantly, and has a probability π_j to be in the state represented by ε_j and $\sum_{j=1}^N \pi_j = 1$.

Then the probability that we have $(n_k)_{k=1}^p$ microstates in the states $(\varepsilon_k)_{k=1}^p$ is $\frac{N!}{n_1! \dots n_p!} (\pi_{j_1})^{n_1} \dots (\pi_{j_p})^{n_p}$ and the expected value of γ is :

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

with

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

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 :

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

With the Lagrange multipliers a for $\sum_{j=1}^N \pi_j = 1$ and $\theta \in F^*$ for $\langle \gamma \rangle$ the solution is given by the equations :

$$i=1\dots N : \frac{\partial}{\partial \pi_i} \left(- \sum_{j=1}^N \pi_j \ln \pi_j - a \left(\sum_{j=1}^N \pi_j \right) - \theta(z) \right) = 0$$

That is :

$$- \ln \pi_i - 1 - a - \theta \left(\frac{\partial z}{\partial \pi_i} \right) = 0$$

$$\pi_i = \exp \left(-1 - a - \theta \left(\frac{\partial z}{\partial \pi_i} \right) \right)$$

$$\sum_{j=1}^N \pi_j = 1 \Rightarrow \exp(1 + a) = \sum_j \exp \left(-\theta \left(\frac{\partial z}{\partial \pi_j} \right) \right)$$

The estimation of π_i is then a Gibbs measure :

$$\pi_i = \frac{1}{Z} \exp \left(-\theta \left(\frac{\partial z}{\partial \pi_i} \right) \right)$$

$$Z = \sum_j \exp \left(-\theta \left(\frac{\partial z}{\partial \pi_j} \right) \right) \text{ is the distribution function.}$$

And we have the additional equations :

$$\langle \gamma \rangle = \frac{1}{Z^N} \sum_{\lambda} \frac{N!}{n_1! \dots n_p!} \sum_{1 \leq j_1 \leq \dots \leq j_p \leq N}$$

$$\times \left(\exp \left(-n_1 \theta \left(\frac{\partial z}{\partial \pi_{j_1}} \right) \right) \right) \dots \exp \left(-n_p \theta \left(\frac{\partial z}{\partial \pi_{j_p}} \right) \right) \widehat{G} \left(\otimes_{n_1} \varepsilon_{j_1} \otimes_{n_2} \varepsilon_{j_2} \dots \otimes_{n_p} \varepsilon_{j_p} \right)$$

$$\langle \gamma \rangle = \frac{1}{Z^N} \sum_{\lambda} \frac{N!}{n_1! \dots n_p!} \sum_{1 \leq j_1 \leq \dots \leq j_p \leq N} \left(\exp \left(-\theta \left(\sum_{k=1}^p n_k \frac{\partial z}{\partial \pi_{j_k}} \right) \right) \right) \widehat{G} \left(\otimes_{n_1} \varepsilon_{j_1} \otimes_{n_2} \varepsilon_{j_2} \dots \otimes_{n_p} \varepsilon_{j_p} \right)$$

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 homogeneisation 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.

7.3.4 Evolution of homogeneous systems

The evolution of homogeneous systems raises many interesting issues. But we must first precise our assumptions, basically they are a combination of the conditions 3 and 8 :

Conditions 9 :

i) The system is comprised of a fixed number N of microsystems $s = 1 \dots N$ which are represented by the same model, with variables $(X_s)_{s=1}^N$

ii) For each microsystem :

- 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

- $\forall t \in R$ the evaluation map : $\mathcal{E}(t) : O \rightarrow E : \mathcal{E}(t) X_s = X_s(t)$ is continuous

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

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

iv) The microsystems are indistinguishable : any permutation of the N microsystems gives the same evolution of the total system.

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

Proposition 21 *If the conditions 9 are met then 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 $\mathbf{f}(t)$ such that $(\mathbf{f}(t), \widehat{U}_F)$, where \widehat{U}_F is the permutation on $\otimes_N F$, is an irreducible representation of $\mathfrak{S}(N)$*

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 19 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 14 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 : $\widehat{\mathcal{E}}_N(t) (\otimes_N \psi_s) = \otimes_N X_s(t)$

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

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

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

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

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

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

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

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

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

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

$\forall \theta \in \mathfrak{S}(\lambda^c) : \widehat{U}_F(\theta) \left(\otimes_{n_1} \varphi_{j_1}(t) \otimes_{n_2} \varphi_{j_2}(t) \dots \otimes_{n_p} \varphi_{j_p}(t) \right) \neq \left(\otimes_{n_1} \varphi_{j_1}(t) \otimes_{n_2} \varphi_{j_2}(t) \dots \otimes_{n_p} \varphi_{j_p}(t) \right)$ and the tensors are linearly independant

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

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

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

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

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

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

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

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

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

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

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

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

For each t the space $\mathbf{f}(t)$ is defined by a hilbertian basis $(f_i)_{i \in I}$ of F, a finite subset J of I, a conjugacy class $\lambda(t)$ and a family of p vectors $(f_{j_k}(t))_{k=1}^p$ belonging to $(f_i)_{i \in J}$.. The set J is arbitrary but defined by \mathbf{h} , so it does not depend on t. For a given class of conjugacy different family 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 $S(t) \in \mathbf{f}_\lambda :$

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

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 $\mathbf{f}_\lambda :$

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

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

The distance between $S(t)$ and a given \mathbf{f}_{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 phases as classes of conjugacy. So, in a probabilist picture, one can assume that the probability for the system to be in a phase λ : $\Pr(S(t) \in \mathbf{f}_{\lambda})$ is a function of $\frac{\|\pi_{\lambda} S(t)\|^2}{\|S(t)\|^2}$. It can be estimated as seen previously from data on a past period, with the knowledge of both λ and $\frac{\|\pi_{\lambda} S(t)\|^2}{\|S(t)\|^2}$.

A global observable can be defined as above. This a symmetric map $G \in \mathcal{L}(\otimes_N F; W)$ so :

$$\begin{aligned} \text{if } S(t) \in \mathbf{f}_{\lambda} : G(S(t)) &= \sum_{\sigma \in \mathfrak{S}(\lambda^c)} S_{\lambda}^{\sigma}(t) G\left(\widehat{U}_F(\sigma) \otimes_{n_1} f_1 \otimes_{n_2} f_2 \dots \otimes_{n_p} f_p\right) = \\ G_{\lambda} \sum_{\sigma \in \mathfrak{S}(\lambda^c)} S_{\lambda}^{\sigma}(t) \\ g(t) = G(S(t)) &= \sum_{\lambda} G_{\lambda} C_{\lambda}(t) \text{ with} \\ G_{\lambda} &= G\left(\otimes_{n_1} f_1 \otimes_{n_2} f_2 \dots \otimes_{n_p} f_p\right) \\ C_{\lambda}(t) &= 0 \text{ if } S(t) \notin \mathbf{f}_{\lambda}, C_{\lambda}(t) = \sum_{\sigma \in \mathfrak{S}(\lambda^c)} S_{\lambda}^{\sigma}(t) \text{ if } S(t) \in \mathbf{f}_{\lambda} \Rightarrow C_{\lambda}(t) C_{\lambda'}(t) = \\ &\delta_{\lambda\lambda'} C_{\lambda}(t)^2 \end{aligned}$$

So there are two scalar functions which have a special interest : $C_{\lambda}(t)$ which gives $g(t)$ in a given phase, and $\frac{\|\pi_{\lambda} S(t)\|^2}{\|S(t)\|^2}$ which gives the probability to be in a phase λ . As above the value of $S_{\lambda}^{\sigma}(t)$ cannot be measured, $\sum_{\sigma \in \mathfrak{S}(\lambda^c)} S_{\lambda}^{\sigma}(t)$ only is accessible.

The definition of the entropy must be changed. The concept of entropy is linked to the distribution of the states of microsystems, in a probabilist point of view, and should vary with t . So one cannot keep fixed probabilities π_i , and it is difficult to ignore the scalars $C_{\lambda}(t)$. The distribution of the states can be measured by the distribution of $\rho_{\lambda}(t) = \frac{\|\pi_{\lambda} S(t)\|^2}{\|S(t)\|^2}$ which can be the measure of the probability that a system is in the phase λ . So a proxy for the entropy can be : $E(t) = - \sum_{\lambda} \rho_{\lambda}(t) \ln \rho_{\lambda}(t)$

7.4 Variable number of interacting systems

If the number N of interacting systems can vary it is still possible to implement most of the previous result. It makes sense only for homogeneous systems, so the set J is fixed from the microstates.

The Fock space is : $\oplus_{N=0}^{\infty} (\otimes_N H_J)$. Its vectors are defined by an infinite ordered set $(\Psi_0, \Psi_1, \Psi_2, \dots)$ where $\Psi_N \in \otimes_N H_J$. $N = 0$ corresponds to real scalars (JCD 13.5). If the model represents maps $(X_s)_{s=1}^N$ for each microsystem then we have a unique vector $(\Psi_0, \Psi_1, \Psi_2, \dots)$ which encompasses all the possible

values of N . Each Ψ_N belongs to a vector space \mathbf{h}_N corresponding to some class of conjugation λ of $\mathfrak{S}(N)$.

There is an extensive literature on the Fock spaces. Let us just remind the main results.

This is a Hilbert space, with hilbertian bases defined as infinite extension of hilbertian bases of $\otimes_N H_J$.

Any operator on the Hilbert spaces can be extended to a linear continuous operator on the Fock space.

For each Fock space $\oplus_{N=1}^{\infty} (\otimes_N H_J)$ there is a number operator \widehat{N} , self adjoint, whose, dense, domain is :

$$D(\widehat{N}) = \left\{ \Psi_N \in \otimes_N H_J, \sum_{N \geq 0} N^2 \|\Psi_N\|^2 < \infty \right\}$$

$$\widehat{N}(\Psi) = (0, \Psi_1, 2\Psi_2, \dots, N\Psi_N, \dots)$$

The annihilation operator cuts a tensor at its beginning :

$$a_N : H_J \rightarrow \mathcal{L}(\otimes_N H; \otimes_{N-1} H) ::$$

$$a_N(v)(\Psi_1 \otimes \Psi_2 \dots \otimes \Psi_N) = \frac{1}{\sqrt{N}} \langle v, \Psi_1 \rangle \Psi_2 \otimes \Psi_2 \dots \otimes \Psi_N$$

The creation operator adds a vector to a tensor at its beginning :

$$a_N^* : H \rightarrow \mathcal{L}(\otimes_N H; \otimes_{N+1} H) ::$$

$$a_N^*(v)(\Psi_1 \otimes \Psi_2 \dots \otimes \Psi_N) = \sqrt{N+1} v \otimes \Psi_1 \otimes \Psi_2 \dots \otimes \Psi_N$$

a_N^* is the adjoint of a_N and a_N, a_N^* can be extended to the Fock space as a, a^* .

If the model represents the evolution of interacting microsystems then we can assume that at a given time the number $N(t)$ of microsystems is fixed. The state $S(t)$ belongs to one of the vector spaces $\otimes_N F_J$.

8 CONCLUSION

The results presented here provide a framework which can be usefully implemented in various domains.

In Physics it gives an interpretation of the "axioms" of Quantum Mechanics, which I have further studied in a previous paper (2014). Basically the computational methods are the same, but they are simpler and, from my point of view, safer in that one can clearly exposes the conditions in which they may be used.

The more fructuous applications can be found in Economics. Economists are familiar with models, statistics and advanced mathematics. However it does not seem that Hilbert spaces have been largely studied, except for some statistical methods. The tools presented here can be useful, notably to study the phases and in theories of general equilibrium.

By extension they can find some applications in other social sciences, whenever quantitative models are considered.

jc.dutailly@free.fr

9 BIBLIOGRAPHY

- R.D.Anderson *Some open questions in infinite dimensional topology* Proceeding of the 3d Prague symposium Praha (1972)
- O.Bratelli, D.W.Robinson *Operators algebras and quantum statistical mechanics* Springer (2002)
- J.C.Dutailly *Mathematics for theoretical physics* arXiv:1209-5665v2 [math-ph] 4 feb 2014
- J.C.Dutailly *Particles and fields* hal-0933043, version 1, 19 janvier 2014
- Tepper L.Gill, G.R.Pantsulaia, W.W.Zachary *Constructive analysis in infinitely many variables* arXiv 1206-1764v2 [math-FA] 26 june 2012
- D.W.Henderson *Infinite dimensional manifolds are open subsets of Hilbert spaces* (1969) Internet paper
- E.T.Jaynes *Where do we stand on maximum entropy ?* Paper for a MIT conference (1978)
- H.Torunczyk *Characterizing Hilbert spaces topology* Fundamental mathematica (1981)