Lecture 7
QUANTUM PROBABILITY

Stéphane ATTAL

Abstract Bell Inequalities and the failure of hidden variable approaches show that random phenomena of Quantum Mechanics cannot be modeled by classical Probability Theory. The aim of Quantum Probability Theory is to provide an extension of the classical theory of probability which allows to describe those quantum phenomena. Quantum Probability Theory does not add anything to the axioms of Quantum Mechanics, it just emphasizes the probabilistic nature of them. This lecture is devoted to introducing this quantum extension of Probability Theory, its connection with the quantum mechanical axioms. We also go a step further by introducing the Toy Fock spaces and by showing how they hide very rich quantum and classical probabilistic structures.

For reading this chapter the reader should be familiar with the basic notions of Quantum Mechanics, but also of Probability Theory and discrete time stochastic processes. If necessary read Lectures 4 and 5.

Stéphane ATTAL
Institut Camille Jordan, Université Lyon 1, France
e-mail: attal@math.univ-lyon1.fr

7.1 Observables and Laws

There exists plenty of different definitions of a quantum probability space in the literature, with different levels of generality. Most commonly Quantum Probability Theory is defined at the level of von Neumann algebras and normal states. In this lecture we consider a lower level of generality in order to stick to the most usual axioms of Quantum Mechanics, that is, the Hilbert space level, where states are density matrices and observables are self-adjoint operators.

7.1.1 States and Observables

Definition 7.1. A *quantum probability space* is a pair $(\mathcal{H},\rho)$ where $\mathcal{H}$ is a separable Hilbert space and ρ is a density matrix on $\mathcal{H}$. Such a ρ on $\mathcal{H}$ is called a *state* on $\mathcal{H}$. The set of states on $\mathcal{H}$ is denoted by $\mathcal{S}(\mathcal{H})$.

Proposition 7.2. *The set of states $\mathcal{S}(\mathcal{H})$ is convex. The extremal points of $\mathcal{S}(\mathcal{H})$ are the pure states $\rho=|u\rangle\langle u|$.*

Proof. The convexity of the set $\mathcal{S}(\mathcal{H})$ is obvious, let us characterize the extremal points of this set. Recall that any state ρ on $\mathcal{H}$ admits a canonical decomposition

$$\rho=\sum_{n\in\mathbb{N}}\lambda_n\,|\phi_n\rangle\langle\phi_n|\,,$$

where the λ_n are positive with $\sum_n \lambda_n = 1$ and the ϕ_n's form an orthonormal basis of $\mathcal{H}$. From this representation it is clear that if ρ is extremal then it is of the form $|u\rangle\langle u|$. Conversely, take $\rho = |u\rangle\langle u|$ and assume that $\rho = \lambda\rho_1 + (1-\lambda)\rho_2$ is a convex combination of two states. Put $\mathsf{P} = \mathsf{I} - |u\rangle\langle u|$, so that $\mathrm{Tr}\,(\rho\,\mathsf{P}) = 0$ and thus $\mathrm{Tr}\,(\rho_1\,\mathsf{P}) = \mathrm{Tr}\,(\rho_2\,\mathsf{P}) = 0$. If for every $i = 1, 2$ we put $\rho_i = \sum_j \lambda^i_j\,|u^i_j\rangle\langle u^i_j|$, we then get

$$\mathrm{Tr}\,(\rho_i\,\mathsf{P}) = \sum_j \lambda^i_j\,\langle u^i_j\,,\,\mathsf{P}u^i_j\rangle = 0\,,$$

that is, $\sum_j \lambda^i_j\,|\langle u^i_j, u\rangle|^2 = 1$. The only possibility is that only one of λ^i_j is not null (and thus equal to 1) and the corresponding u^i_j is equal to u. This gives $\rho_1 = \rho_2$. $\square$

Definition 7.3. An *observable* on $\mathcal{H}$ is a self-adjoint operator on $\mathcal{H}$. The space of observables on $\mathcal{H}$ is denoted by $\mathcal{O}(\mathcal{H})$.

Let A be an observable on $\mathcal{H}$. By the von Neumann Spectral Theorem there exists a spectral measure ξ associated to A such that

$$\mathsf{A} = \int_{\mathbb{R}} \lambda\,\mathrm{d}\xi(\lambda)\,.$$

If ρ is a state on $\mathcal{H}$ then the mapping

$$\begin{aligned} \mu : \mathrm{Bor}(\mathbb{R}) &\longrightarrow [0,1] \\ E &\longmapsto \mathrm{Tr}\,(\rho\,\xi(E)) \end{aligned}$$

is a probability measure on $\mathbb{R}$. This probability measure μ is called the *law*, or *distribution* of A under the state ρ.

The following proposition gives another characterization of the distribution μ of A.

Proposition 7.4. *Let $(\mathcal{H}, \rho)$ be a quantum probability space and A be an observable on $\mathcal{H}$. Then the law of A under the state ρ is the only probability measure μ on $\mathbb{R}$ such that, for all bounded Borel function f on $\mathbb{R}$ we have*

$$\int_{\mathbb{R}} f(x)\,\mathrm{d}\mu(x) = \mathrm{Tr}\,\big(\rho\,f(\mathsf{A})\big)\,. \tag{7.1}$$

Proof. Let μ be the law of A. We write

$$\rho = \sum_{n\in\mathbb{N}} \lambda_n\,|u_n\rangle\langle u_n|$$

and as a consequence

$$\mathrm{Tr}\,\big(\rho\,f(\mathsf{A})\big) = \sum_{n\in\mathbb{N}} \lambda_n\,\langle u_n\,,\,f(\mathsf{A})\,u_n\rangle$$

where the series is absolutely convergent. Let ξ denote the spectral measure of A and let μ_n be the measure given by $\mu_n(E) = \langle u_n\,,\,\xi(E)\,u_n\rangle = \|\xi(E)\,u_n\|^2$, for all n. We then have

$$\mathrm{Tr}\,\big(\rho\,f(\mathsf{A})\big) = \sum_{n\in\mathbb{N}} \lambda_n \int_{\mathbb{R}} f(x)\,\mathrm{d}\mu_n(x)\,.$$

On the other hand, if μ is the law of A we have

$$\mu(E) = \sum_{n\in\mathbb{N}} \lambda_n\,\langle u_n, \xi(E)\,u_n\rangle = \sum_{n\in\mathbb{N}} \lambda_n\,\mu_n(E)\,,$$

that is, μ is the measure $\sum_{n\in\mathbb{N}} \lambda_n\,\mu_n$. This proves that the relation (7.1) holds when μ is the law of A.

Conversely, if Equation (7.1) holds for a measure μ, then automatically

$$\mu(E) = \mathrm{Tr}\,(\rho\,\mathbb{1}_E(A)) = \mathrm{Tr}\,(\rho\,\xi(E))\,,$$

for all $E \in \mathrm{Bor}(\mathbb{R})$. This says exactly that μ is the law of A. $\square$

Another nice characterization is obtained in terms of the Fourier transform.

Theorem 7.5. *Let $(\mathcal{H}, \rho)$ be a quantum probability space and let A be an observable on $\mathcal{H}$. Then the function*

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \quad \mathbb{R} \\ t &\longmapsto \mathrm{Tr}\,(\rho\,\mathrm{e}^{\mathrm{i}t\mathsf{A}}) \end{aligned}$$

is the Fourier transform of some probability measure μ on $\mathbb{R}$. This measure is the law of A under the state ρ.

Proof. For any $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ and $t_1, \dots, t_n \in \mathbb{R}$ we have

$$\begin{aligned} \sum_{k,j=1}^n \lambda_k\,\overline{\lambda}_j\,f(t_k - t_j) &= \sum_{k,j=1}^n \lambda_k\,\overline{\lambda}_j\,\mathrm{Tr}\,\left(\rho\,\mathrm{e}^{\mathrm{i}(t_k - t_j)\mathsf{A}}\right) \\ &= \sum_{k,j=1}^n \lambda_k\,\overline{\lambda}_j\,\mathrm{Tr}\,\left(\rho\,\mathrm{e}^{\mathrm{i}t_k\mathsf{A}}\big(\mathrm{e}^{\mathrm{i}t_j\,\mathsf{A}}\big)^*\right) \\ &= \mathrm{Tr}\,\left(\rho\left(\sum_{k=1}^n \lambda_k\,\mathrm{e}^{\mathrm{i}t_k\,\mathsf{A}}\right)\left(\sum_{j=1}^n \lambda_j\,\mathrm{e}^{\mathrm{i}t_j w\,\mathsf{A}}\right)^*\right)\,. \end{aligned}$$

As ρ is positive, the quantity above is positive. By Bochner's criterion this proves that f is the Fourier transform of some probability measure μ on $\mathbb{R}$. By Proposition 7.4 and by uniqueness of the Fourier transform, this measure is the law of A. $\square$

We now make a remark concerning the action of unitary conjugation in this setup.

Proposition 7.6. *Let ρ be a state and A be an observable on $\mathcal{H}$. If U is a unitary operator on $\mathcal{H}$, then $\mathsf{U}^*\mathsf{A}\,\mathsf{U}$ is still an observable and $\mathsf{U}\,\rho\,\mathsf{U}^*$ is still a state. Furthermore, the law of $\mathsf{U}^*\mathsf{A}\,\mathsf{U}$ under the state ρ is the same as the law of A under the state $\mathsf{U}\,\rho\,\mathsf{U}^*$.*

Proof. The facts that $\mathsf{U}^*\mathsf{A}\,\mathsf{U}$ is still an observable and $\mathsf{U}\,\rho\,\mathsf{U}^*$ is still a state are obvious and left to the reader. We have

$$\mathrm{Tr}\,\left(\rho\,\mathsf{U}^*\mathsf{A}\,\mathsf{U}\right) = \mathrm{Tr}\,\left(\mathsf{U}\,\rho\,\mathsf{U}^*\,\mathsf{A}\right).$$

Furthermore, for any bounded function f on $\mathbb{R}$ we have, by the functional calculus

$$f(\mathsf{U}^*\mathsf{A}\,\mathsf{U}) = \mathsf{U}^* f(\mathsf{A})\,\mathsf{U}\,,$$

which gives finally

$$\mathrm{Tr}\,(\rho\, f(\mathsf{U}^*\mathsf{A}\,\mathsf{U})) = \mathrm{Tr}\,(\mathsf{U}\,\rho\,\mathsf{U}^*\, f(\mathsf{A}))\,.$$

By Proposition 7.4, this says exactly that the law of $\mathsf{U}^*\mathsf{A}\,\mathsf{U}$ under the state ρ is the same as the law of A under the state $\mathsf{U}\,\rho\,\mathsf{U}^*$. $\square$

7.1.2 Observables vs Random Variables

We stop for a moment our list of definitions and properties and we focus on the relations between observables and usual random variables in Probability Theory.

For the simplicity of the discussion we consider only the case of pure states, the general case can be easily deduced by a procedure called the G.N.S. representation, which is not worth developing for this discussion.

A Hilbert space $\mathcal{H}$ and a wave function φ being given, let us see that any observable on $\mathcal{H}$ can be viewed as a classical random variable. Let A be an observable on $\mathcal{H}$, i.e. a self-adjoint operator on $\mathcal{H}$. Consider the probability measure

$$E \mapsto \mu(E) = \left\|\mathbb{1}_E(\mathsf{A})\,\varphi\right\|^2$$

on $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$. Consider the operator U from $L^2(\mathbb{R}, \mathrm{Bor}(\mathbb{R}), \mu)$ to $\mathcal{H}$ given by $\mathsf{U}f = f(\mathsf{A})\,\varphi$. This is a well-defined operator and an isometry, for the von Neumann Spectral Theorem gives

$$\left\|f(\mathsf{A})\,\varphi\right\|^2 = \int_{\mathbb{R}} |f(x)|^2\;\mathrm{d}\mu(x)\,.$$

We deduce easily from this definition

$$\mathsf{U}^* f(\mathsf{A})\, \varphi = f$$

and

$$\mathsf{U}^* \varphi = \mathbb{1}\,.$$

The operator $\mathsf{U}^* \mathsf{A}\, \mathsf{U}$ is then clearly equal to the operator $\mathcal{M}_X$ of multiplication by the function $X(x) = x$ on $\mathbb{R}$. Furthermore, we have

$$\begin{aligned}\langle\, \varphi\,,\, f(\mathsf{A})\, \varphi\,\rangle &= \langle\, \mathsf{U}^*\, \varphi\,,\, \mathsf{U}^* f(\mathsf{A}) \mathsf{U}\, \mathsf{U}^*\, \varphi\,\rangle\\ &= \langle\, \mathbb{1}\,,\, \mathcal{M}_{f(X)}\, \mathbb{1}\,\rangle\\ &= \int_{\mathbb{R}} f(X(x))\, \mathrm{d}\mu(x)\,.\end{aligned}$$

This means that the probability distribution associated to the observable A, in the state φ, as described in the setup of Quantum Probability Theory, is the same as the usual probability distribution of X as a random variable on $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}), \mu)$.

Being given a single observable A on $\mathcal{H}$ together with a state φ is thus the same as being given a classical random variable X on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The observable A specifies the random variable X as a function on Ω, the state φ specifies the underlying probability measure $\mathbb{P}$ and therefore specifies the probability distribution of X.

In the converse direction, it is easy to see a classical probabilistic setup as a particular case of the quantum setup. Let X be a real random variable defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Put $\mathcal{H} = L^2(\Omega, \mathcal{F}, \mathbb{P})$ and $\varphi = \mathbb{1}$. The operator $\mathsf{A} = \mathcal{M}_X$ is then a self-adjoint operator on $\mathcal{H}$. Let us compute the probability distribution of the observable A in the state φ, in the sense of Quantum Probability. We get

$$\begin{aligned}\langle\, \varphi\,,\, f(\mathsf{A})\, \varphi\,\rangle &= \langle\, \mathbb{1}\,,\, \mathcal{M}_{f(X)}\, \mathbb{1}\,\rangle\\ &= \int_{\mathbb{R}} f(X(\omega))\, \mathrm{d}\mathbb{P}(\omega)\\ &= \mathbb{E}[f(X)]\,.\end{aligned}$$

Thus, the distribution of the observable A, in the quantum sense is the same as the probability distribution of the random variable X in the usual sense.

One could be tempted to conclude that there is no difference between the probabilistic formalism of Quantum Mechanics and the usual Probability Theory. But this is only true for **one** observable, or actually for a family of *commuting* observables. Indeed, once considering two observables A and B on $\mathcal{H}$ which do not commute, the picture is completely different. Each of the observables A and B can be seen as concrete random variables X and Y respectively, but **on some different probability spaces**. Indeed, if A and B do not commute they cannot be diagonalized simultaneously and they are represented as multiplication operators on some **different** probability spaces (otherwise they would commute!).

Furthermore, their associated probability spaces have nothing to do together. They cannot be put together, as one usually does in the case of independent random variables, by taking the tensor product of the two spaces, because the two random variables X and Y are not independent. As operators on $\mathcal{H}$ the observables A and B may have some strong relations, such as $[\mathsf{A}, \mathsf{B}] = \lambda \mathrm{I}$, for example (this is the case for momentum and position observables in Quantum Mechanics).

Is there a way, with the random variables X, Y and their associated probability spaces, to give an account of such a relation? Aspect's experiment and Bell's inequalities prove that this is actually impossible! As an example, the spin of an electron in two different directions gives rise to two Bernoulli random variables but each one on its own probability space. These two Bernoulli random variables are not at all independent, they depend of each other in a way which is impossible to express in classical terms. The only way to express their dependency is to represent them as 2×2 spin matrices on $\mathbb{C}^2$ and work with the quantum mechanical axioms!

As a conclusion, each single observable in Quantum Mechanics is like a concrete random variable, but on its **own** probability space, "with its own dice". Two non-commuting observables have different associated probability spaces, "different dices"; they cannot be represented by independent random variables or anything else (even more complicated) in the classical language of probability. The only way to express their mutual dependency is to consider them as two self-adjoint operators on the same Hilbert space and to compute their probability distributions as the quantum mechanical axioms dictate us.

In the quantum formalism it is possible to make operations on non commuting observables A, B, for example to add them and to consider the result as a new observable. The distribution of $\mathsf{A} + \mathsf{B}$ has then nothing to do with the convolution of the distribution of A by the one of B, it depends in a complicated way on the relations between A and B as operators on $\mathcal{H}$. We shall see examples later on in this course.

The case of commuting observables is equivalent to usual case of a pair of random variables. Let us here give some details about that.

Let A and B be two observables on $\mathcal{H}$ which commute (in the case of unbounded operators this means that their spectral measures commute). Then there exists a spectral measure $d\xi(x, y)$ on $\mathbb{R} \times \mathbb{R}$ such that

$$\mathsf{A} = \int_{\mathbb{R}^2} x \, \mathrm{d}\xi(x, y) \quad \text{and} \quad \mathsf{B} = \int_{\mathbb{R}^2} y \, \mathrm{d}\xi(x, y)\,.$$

In other words, A and B can be diagonalized simultaneously, they can be represented as multiplication operators on the same probability space.

The pair (A, B) admits a law in the quantum sense. Indeed, if ρ is a state on $\mathcal{H}$, then the mapping

$$\begin{aligned} \mu : \mathrm{Bor}(\mathbb{R}\times\mathbb{R}) &\longrightarrow \qquad [0,1] \\ E\times F \quad &\longmapsto \mathrm{Tr}\,\big(\rho\,\xi(E,F)\big) \end{aligned}$$

extends to a probability measure on $\mathbb{R}\times\mathbb{R}$. This probability measure is called the *law of the pair* (A,B) under the state ρ. Another way to understand this law is to say that if A and B commute they admit a two variable functional calculus and, for every bounded measurable function f on $\mathbb{R}\times\mathbb{R}$, we have

$$\int_{\mathbb{R}^2} f(x,y)\,\mathrm{d}\mu(x,y) = \mathrm{Tr}\,\left(\rho\, f(\mathsf{A},\mathsf{B})\right)\,.$$

From the Fourier transform point of view, the law μ is obtained by noticing that if A and B commute then the function

$$(t_1,t_2)\longmapsto \mathrm{Tr}\,\left(\rho\,\mathrm{e}^{\mathrm{i}(t_1\mathsf{A}+t_2\mathsf{B})}\right)$$

satisfies Bochner's criterion for two variable functions and hence is the Fourier transform of some probability measure μ on $\mathbb{R}\times\mathbb{R}$. Clearly, this measure is the law of (A,B) under ρ as defined above.

At this stage, it is interesting to note a way to produce independent random variables in the quantum probability context.

Proposition 7.7. *Let* A *and* B *be two observables on* $\mathcal{H}$ *and* $\mathcal{K}$ *respectively. Let* μ *and* ν *denote their laws under the states* ρ *and* τ*, respectively. On the Hilbert space* $\mathcal{H}\otimes\mathcal{K}$ *consider the commuting observables*

$$\widehat{\mathsf{A}} = \mathsf{A}\otimes \mathrm{I} \qquad \textit{and} \qquad \widehat{\mathsf{B}} = \mathrm{I}\otimes\mathsf{B}\,.$$

Then, under the state $\rho\otimes\tau$*, the pair of observables* $(\widehat{\mathsf{A}},\widehat{\mathsf{B}})$ *follows the law* $\mu\otimes\nu$*, that is, they are independent observables with the same individual law as* A *and* B *respectively.*

Proof. Clearly $\widehat{\mathsf{A}}$ and $\widehat{\mathsf{B}}$ commute, hence they admit a law of pair, let us compute this law. For every bounded real functions f and g we have

$$f(\widehat{\mathsf{A}}) = f(\mathsf{A})\otimes\mathrm{I} \qquad \text{and} \qquad g(\widehat{\mathsf{B}}) = \mathrm{I}\otimes g(\mathsf{B})\,.$$

Thus

$$\begin{aligned} \mathrm{Tr}\,\left((\rho\otimes\tau)\, f(\widehat{\mathsf{A}})g(\widehat{\mathsf{B}})\right) &= \mathrm{Tr}\,\left(\big(\rho\, f(\mathsf{A})\big)\otimes\big(\tau\, g(\mathsf{B})\big)\right) \\ &= \mathrm{Tr}\,\big(\rho\, f(\mathsf{A})\big)\;\mathrm{Tr}\,\big(\tau\, g(\mathsf{B})\big) \\ &= \int_{\mathbb{R}} f(x)\,\mathrm{d}\mu(x)\,\int_{\mathbb{R}} g(y)\,\mathrm{d}\nu(y) \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} f(x)g(y)\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y)\,. \end{aligned}$$

Taking $g = \mathbb{1}$, the identity above shows that $\widehat{\mathsf{A}}$ has the same law under $\rho \otimes \tau$ as the one of A under the state ρ. The corresponding property of course holds for B. Taking general f and g, the identity above shows the independence of $\widehat{\mathsf{A}}$ and $\widehat{\mathsf{B}}$. □

7.1.3 The Role of Multiplication Operators

Together with the discussion we had above on the connection between observables and classical random variables, I want to put the emphasis here on the particular role of multiplication operators in the quantum setup.

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a family of real random variables on $(X_i)_{i\in I}$ which are of interest for some reasons. Imagine we have found a natural unitary isomorphism $\mathsf{U} : L^2(\Omega, \mathcal{F}, \mathbb{P}) \to \mathcal{H}$ with some Hilbert space $\mathcal{H}$. The question we want to discuss here is the following: Where can we read on $\mathcal{H}$ the random variables X_i, with all their probabilistic properties (individual distributions, relations with each other such as independence or dependence, joint distributions, functional calculus, etc.)?

This is certainly not by looking at the images $h_i = \mathsf{U}X_i$ of the X_i's in $\mathcal{H}$. Indeed, these elements h_i of $\mathcal{H}$ could be almost any element of $\mathcal{H}$ by choosing correctly the unitary operator U. They carry no probabilistic information whatsoever on the random variables X_i.

The pertinent objects to look at are the operators

$$\mathsf{A}_i = \mathsf{U}\, \mathcal{M}_{X_i}\, \mathsf{U}^* \,,$$

that is, the push-forward of the operators of multiplications by the X_i's. Indeed, these operators are a commuting family of self-adjoint operators on $\mathcal{H}$. The functional calculus for such families of operators gives easily

$$f(\mathsf{A}_{i_1}, \dots, \mathsf{A}_{i_n}) = \mathsf{U}\, \mathcal{M}_{f(X_{i_1},\dots,X_{i_n})}\, \mathsf{U}^* \,.$$

If we put $\phi = \mathsf{U}\mathbb{1}$, then ϕ is a pure state of $\mathcal{H}$ and we have

$$\begin{aligned}
\mathrm{Tr}\,(|\phi\rangle\langle\phi|\, f(\mathsf{A}_{i_1}, \dots, \mathsf{A}_{i_n})) &= \langle \phi\,,\, f(\mathsf{A}_{i_1}, \dots, \mathsf{A}_{i_n})\phi\rangle \\
&= \langle \mathbb{1}\,,\, \mathsf{U}^*\, f(\mathsf{A}_{i_1}, \dots, \mathsf{A}_{i_n})\, \mathsf{U}\, \mathbb{1}\rangle \\
&= \left\langle \mathbb{1}\,,\, \mathcal{M}_{f(X_{i_1},\dots,X_{i_n})}\, \mathbb{1}\right\rangle \\
&= \mathbb{E}\left[f(X_{i_1}, \dots, X_{i_n})\right] \,.
\end{aligned}$$

The law of the n-uplet $(\mathsf{A}_{i_1}, \dots, \mathsf{A}_{i_n})$ as observables in the state $|\phi\rangle\langle\phi|$ is thus the same as the usual law of the n-uplet $(X_{i_1}, \dots, X_{i_n})$.

The operators A_i contain all the probabilistic informations on the X_i's: individual and joint laws, functional calculus, etc. The operators A_i play exactly the same role in $\mathcal{H}$ as the random variables X_i play in $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

The only informations concerning the X_i's that one cannot recover from the A_i's are the samples, that is, the individual values $X_i(\omega)$ for $\omega \in \Omega$.

We end this discussion with an extension to a situation which is very common in Quantum Noise Theory. Together with the situation described above, we have another Hilbert space $\mathcal{K}$ and an operator on $\mathcal{K} \otimes \mathsf{H}$ of the form

$$\mathsf{T} = \sum_{i=1}^{n} \mathsf{B}_i \otimes \mathsf{A}_i \,,$$

for some bounded operators B_i on $\mathcal{K}$. On the other hand, consider the random operator

$$\begin{array}{rl} \mathsf{S} : \Omega \longrightarrow & \mathcal{B}(\mathcal{K}) \\ \omega \longmapsto & \mathsf{S}(\omega) = \sum_{i=1}^{n} X_i(\omega) \mathsf{B}_i \,. \end{array}$$

We want to make clear the connections between T and S and show that they are one and the same thing.

Let us define the "multiplication" operator associated to S:

$$\begin{array}{rcl} \widehat{\mathsf{S}} : L^2(\Omega, \mathcal{F}, \mathbb{P}) \otimes \mathcal{H} & \longrightarrow & L^2(\Omega, \mathcal{F}, \mathbb{P}) \otimes \mathcal{H} \\ f \otimes h & \longmapsto & \sum_{i=1}^{n} X_i\, f \otimes \mathsf{B}_i\, h \,. \end{array}$$

In other words

$$\widehat{\mathsf{S}} = \sum_{i=1}^{n} \mathcal{M}_{X_i} \otimes \mathsf{B}_i \,.$$

Then the connection becomes clear:

$$\mathsf{T} = (\mathsf{U} \otimes \mathrm{I})\, \widehat{\mathsf{S}}\, (\mathsf{U}^* \otimes \mathrm{I}) \,.$$

Even more than just this unitary equivalence, the operator T contains all the probabilistic informations of the random operator S; any kind of computation that one wants to perform on S can be made in the same way on T. The functional calculus is the same, the expectations and the laws are computed is the same way as described above.

7.1.4 Events

In this subsection we discuss the notion of *event* in Quantum Probability. In classical probability theory an event E can be identified to the random variable $X = \mathbb{1}_E$. That is, a random variable X such that $X^2 = \overline{X} = X$. Conversely, every random variable X which satisfies the above relation is the indicator function of some event. On the space $L^2(\Omega, \mathcal{F}, \mathbb{P})$, the indicator function $X = \mathbb{1}_E$ can be identified with its multiplication operator $f \mapsto \mathbb{1}_E f$, it is thus an orthogonal projection of $L^2(\Omega, \mathcal{F}, \mathbb{P})$, whose range is the subspace of functions with support included in E.

Definition 7.8. Following the analogy with the classical case, we define an *event* on a quantum probability space $(\mathcal{H}, \rho)$ to be any orthogonal projector on $\mathcal{H}$. We denote by $\mathcal{P}(\mathcal{H})$ the set of events of $\mathcal{H}$. In particular, if A is an observable on $\mathcal{H}$, with spectral measure ξ_{A} and if E is a Borel set of $\mathbb{R}$, the projector $\xi_{\mathsf{A}}(E)$ can be interpreted as the event "A belongs to E". Up to a normalization constant it is the state one would obtain after a measurement of the observable A, if one had observed "the value of the measurement belongs to the set E". In other words, the subspace $\mathrm{Ran}\,\xi_{\mathsf{A}}(E)$ is the subspace of wave functions for which a measure of A would give a result in E with probability 1; this subspace is indeed the subspace corresponding to the knowledge "A belongs to E" on the system.

This terminology coincides also with the definitions and results of previous subsections, for by Proposition 7.4 we have

$$\mathrm{Tr}\,(\rho\, \xi_{\mathsf{A}}(E)) = \mathrm{Tr}\,(\rho\, \mathbb{1}_E(\mathsf{A})) = \int_{\mathbb{R}} \mathbb{1}_E(x)\, \mathrm{d}\mu(x) = \mu(E)$$

which says that $\mathrm{Tr}\,(\rho\, \xi_{\mathsf{A}}(E))$ is the probability that A belongs to E under the state ρ.

Definition 7.9. Together with this quantum notion of event we make use the following terminology:

– If P_1, P_2 are two events and if $\mathsf{P}_1 \leq \mathsf{P}_2$, we say that P_1 *implies* P_2.

– The operators 0 and I are the *null* and *certain* events.

– The *complementary* of the event P is the event $\mathrm{I} - \mathsf{P}$.

– If $\mathsf{P}_1, \ldots, \mathsf{P}_n$ are events then $\vee_{i=1}^n \mathsf{P}_i$ is the orthogonal projector onto the subspace generated by $\cup_{i=1}^n \mathrm{Ran}\,\mathsf{P}_i$; it is the event "*occurrence of at least one of the* P_i*'s*". In the same way $\wedge_{i=1}^n \mathsf{P}_i$ is the orthogonal projector onto $\cap_{i=1}^n \mathrm{Ran}\,\mathsf{P}_i$; it is the event "*simultaneous occurrence of all the* P_i*'s*".

Definition 7.10. When a state ρ is given, the mapping $\alpha : \mathsf{P} \mapsto \mathrm{Tr}\,(\rho\, \mathsf{P})$ behaves like a probability measure on $\mathcal{P}(\mathcal{H})$ in the sense that

i) $\alpha(\mathrm{I}) = 1$

ii) $\alpha\left(\sum_{i\in\mathbb{N}} \mathsf{P}_i\right) = \sum_{i\in\mathbb{N}} \alpha(\mathsf{P}_i)$, for any family $(\mathsf{P}_i)_{i\in\mathbb{N}}$ of pairwise orthogonal projections.

A mapping $\alpha : \mathcal{P}(\mathcal{H}) \to \mathbb{R}$ which satisfies i) and ii) is called a *probability measure* on $\mathcal{P}(\mathcal{H})$. Note that every probability measure α on $\mathcal{P}(\mathcal{H})$ satisfies:

iii) $\alpha(\mathsf{P}) \in [0, 1]$ for all $\mathsf{P} \in \mathcal{P}(\mathcal{H})$.

iv) $\alpha(\mathrm{I} - \mathsf{P}) = 1 - \alpha(\mathsf{P})$ for all $\mathsf{P} \in \mathcal{P}(\mathcal{H})$.

v) $\alpha(0) = 0$.

One may naturally wonder what is the general form of all probability measures on $\mathcal{P}(\mathcal{H})$. The answer is given by the celebrated Gleason's theorem. This theorem is very deep and has many applications, developments and consequences in the literature. Its proof is very long and heavy. As we do not really need it in the following (we just present it as a remark), we state it without proof. One may consult the book [Par92], pp. 31–40, for a complete proof.

Theorem 7.11 (Gleason's theorem). *Let $\mathcal{H}$ be a separable Hilbert space with dimension greater than or equal to 3. A mapping $\alpha : \mathcal{P}(\mathcal{H}) \to \mathbb{R}$ is a probability measure on $\mathcal{P}(\mathcal{H})$ if and only if there exists a density matrix ρ on $\mathcal{H}$ such that*

$$\alpha(\mathsf{P}) = \mathrm{Tr}\,(\rho\,\mathsf{P})$$

for all $\mathsf{P} \in \mathcal{P}(\mathcal{H})$.

The correspondence $\rho \mapsto \alpha$ between density matrices on $\mathcal{H}$ and probability measures on $\mathcal{P}(\mathcal{H})$ is a bijection.

7.2 Quantum Bernoulli

We shall now enter into concrete examples of Quantum Probability spaces and specific properties of this quantum setup. In this section we aim to explore the simplest situation possible, the quantum analogue of a Bernoulli random walk. We shall see that this simple situation in classical probability becomes incredibly richer in the quantum context.

7.2.1 Quantum Bernoulli Random Variables

Before going to the Bernoulli random walks, we study the quantum analogue of Bernoulli random variables. This is to say that we study now the simplest non-trivial example of a quantum probability space: the space $\mathcal{H} = \mathbb{C}^2$.

Definition 7.12. Let (e_1, e_2) be the canonical basis of $\mathbb{C}^2$. The matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are the so-called *Pauli matrices*. There are also usually denoted by $\sigma_x, \sigma_y, \sigma_z$, respectively.

With the help of these matrices we have very useful parametrizations of observables and states.

Proposition 7.13.
1) Together with the identity matrix I*, the Pauli matrices form a real basis of the space* $\mathcal{O}(\mathcal{H})$*. That is, any observable* A *on* $\mathcal{H}$ *can be written*

$$\mathsf{A} = t\,\mathrm{I} + \sum_{i=1}^{3} x_i\,\sigma_i$$

for some $t, x_1, x_2, x_3 \in \mathbb{R}$*. In particular we get* $\mathrm{Tr}\,(\mathsf{A}) = 2t$*.*
2) Putting

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2}\,,$$

the eigenvalues of A *are* $\left\{t - \|\mathbf{x}\|\,,\ t + \|\mathbf{x}\|\right\}$*.*
3) Any state ρ *is of the form*

$$\rho = \frac{1}{2}\begin{pmatrix} 1+x & y-iz \\ y+iz & 1-x \end{pmatrix}$$

with $x, y, z \in \mathbb{R}$ *satisfying* $x^2 + y^2 + z^2 \leq 1$*. In particular,* ρ *is a pure state if and only if* $x^2 + y^2 + z^2 = 1$*.*

Proof. 1) and 2) are immediate and left to the reader. Let us prove 3). A state ρ on $\mathcal{H}$ is an observable which has trace 1 and positive eingenvalues. Hence by 1) and 2), a state ρ on $\mathcal{H}$ is of the form

$$\rho = \frac{1}{2}\left(\mathrm{I} + \sum_{i=1}^{3} x_i\,\sigma_i\right)$$

with $\|\mathbf{x}\| \leq 1$. As a consequence the space $\mathcal{S}(\mathcal{H})$ identifies to $\overline{B(0,1)}$, the closed unit ball of $\mathbb{R}^3$, with the same convex structure. As a consequence, the extreme points of $\mathcal{S}(\mathcal{H})$, the pure states, correspond to the extreme points of $\overline{B(0,1)}$, that is, the unit sphere S^2 of $\mathbb{R}^3$. $\square$

The result 3) above means that one can always write a pure state as:

$$\begin{aligned}\rho &= \frac{1}{2}\begin{pmatrix} 1+\cos(\varphi) & \mathrm{e}^{-\mathrm{i}\theta}\sin(\varphi) \\ \mathrm{e}^{\mathrm{i}\theta}\sin(\varphi) & 1-\cos(\varphi) \end{pmatrix} \\ &= \begin{pmatrix} \cos^2\left(\frac{\varphi}{2}\right) & \mathrm{e}^{-\mathrm{i}\theta}\sin\left(\frac{\varphi}{2}\right)\cos\left(\frac{\varphi}{2}\right) \\ \mathrm{e}^{\mathrm{i}\theta}\sin\left(\frac{\varphi}{2}\right)\cos\left(\frac{\varphi}{2}\right) & \sin^2\left(\frac{\varphi}{2}\right) \end{pmatrix},\end{aligned}$$

for some $\theta \in [0, 2\pi]$ and $\varphi \in [0, \pi]$. We recognize the operator $|u\rangle\langle u|$ where

$$u = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\frac{\theta}{2}} \cos\left(\frac{\varphi}{2}\right) \\ \mathrm{e}^{\mathrm{i}\frac{\theta}{2}} \sin\left(\frac{\varphi}{2}\right) \end{pmatrix} .$$

This way, one cleary describes all unitary vectors of $\mathbb{C}^2$ up to a phase factor, that is, one describes all rank one projectors $|u\rangle\langle u|$.

We shall now compute the law of quantum observables in this context.

Proposition 7.14. *Consider an orthonormal basis of $\mathbb{C}^2$ for which the state ρ is diagonal:*

$$\rho = \begin{pmatrix} p & 0 \\ 0 & q \end{pmatrix} ,$$

with $0 \leq q \leq p \leq 1$ and $p + q = 1$. If A is any observable

$$\mathsf{A} = t\mathrm{I} + \sum_{i=1}^{3} x_i \, \sigma_i \, ,$$

then the law of A under the state ρ is the law of a random variable which takes the values $t - \|\mathbf{x}\|$ and $t + \|\mathbf{x}\|$ with probability

$$\frac{1}{2}\Big(1 + \frac{x_3}{\|\mathbf{x}\|}(q-p)\Big) \quad \text{and} \quad \frac{1}{2}\Big(1 - \frac{x_3}{\|\mathbf{x}\|}(q-p)\Big)$$

respectively.

Proof. The eigenvalues of A are $t - \|\mathbf{x}\|$ and $t + \|\mathbf{x}\|$, as we have already seen. The eigenvectors are easily obtained too, they are of the form

$$v_1 = \lambda \begin{pmatrix} -x_1 + ix_2 \\ \|\mathbf{x}\| + x_3 \end{pmatrix} \qquad \mathrm{v}_2 = \mu \begin{pmatrix} x_1 - ix_2 \\ \|\mathbf{x}\| - x_3 \end{pmatrix} ,$$

$\lambda, \mu \in \mathbb{C}$, respectively.

Taking the normalized version u_1, u_2 of these vectors, the probability of measuring A to be equal to $t - \|\mathbf{x}\|$ is then

$$\begin{aligned} \langle u_1 \, , \, \rho \, u_1 \rangle &= \frac{1}{2\,\|\mathbf{x}\|\,(\|\mathbf{x}\| + x_3)} \, \left(p(x_1^2 + x_2^2) + q(\|\mathbf{x}\| + x_3)^2\right) \\ &= \frac{p(\|\mathbf{x}\| - x_3) + q(\|\mathbf{x}\| + x_3)}{2\,\|\mathbf{x}\|} \\ &= \frac{1}{2}\Big(1 + \frac{x_3}{\|\mathbf{x}\|}(q-p)\Big) . \end{aligned}$$

This also gives immediately the other probability. □

In the particular case $p = 1$ and $q = 0$, i.e. if $\rho = |e_1\rangle\langle e_1|$, we get all probability distribution on $\mathbb{R}$ with (at most) 2 point support. Note that this situation would be impossible in a classical probability context. Indeed, let $\Omega = \{0, 1\}$ and $\mathbb{P}$ be a fixed probability measure on Ω with $\mathbb{P}(1) = p$ and $\mathbb{P}(0) = q$. Then a random variable X on Ω can take any couple of values (x, y) in $\mathbb{R}$ but only with $\mathbb{P}(X = x) = p$ and $\mathbb{P}(X = y) = q$. In quantum probability, we observe that, under a *fixed* state ρ, the set of observables gives rise to the whole range of two-point support laws. The space of observables is much richer than in the classical context.

We end up this subsection with an illustration of non-commutativity in this context. Let $\rho = |e_1\rangle\langle e_1|$. The observables σ_1, σ_2 both have the distribution $\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$ under that state. But the observable

$$\sigma_1 + \sigma_2 = \begin{pmatrix} 0 & 1-i \\ 1+i & 0 \end{pmatrix}$$

has spectrum $\{-\sqrt{2}, \sqrt{2}\}$ and distribution

$$\frac{1}{2}\delta_{-\sqrt{2}} + \frac{1}{2}\delta_{\sqrt{2}}\,.$$

Adding two Bernoulli random variables with values $\{-1, +1\}$, we have ended up with a Bernoulli random variable with values in $\{-\sqrt{2}, \sqrt{2}\}$! This fact would be of course impossible in the classical context, it comes from the fact that σ_1 and σ_2 do not commute, they do not admit a pair law. Let us check this fact explicitely. When computing

$$\varphi(t_1, t_2) = \langle e_1\,,\, \mathrm{e}^{\mathrm{i}(t_1\sigma_1 + t_2\sigma_2)} e_1\rangle$$

one easily finds $\varphi(t_1, t_2) = \cos\left(\sqrt{t_1^2 + t_2^2}\right)$ which is not the Fourier transform of a positive measure on $\mathbb{R}^2$. Indeed, Bochner's positivity criterion

$$\sum_{i,j=1}^{n} \lambda_i \bar{\lambda}_j\, \varphi(t_i - t_j, s_i - s_j) \geq 0$$

fails when taking $n = 3$, $\lambda_1 = \lambda_2 = \lambda_3 = 1$, $t_1 = s_2 = t_3 = s_3 = 2\pi/3$ and $s_1 = t_2 = 0$.

7.2.2 The Toy Fock Space

In this subsection we go beyond the simple example studied above and we focus on the very rich and fundamental structure of a chain of quantum Bernoulli random variables. The constructions and results of this subsection are fundamental in discrete time Quantum Noise Theory.

Consider the space $\mathbb{C}^2$ with a fixed orthonormal basis which we shall denote by $\{\mathcal{X}^0, \mathcal{X}^1\}$. Physically, one can think of this basis as representing a two-level system with its ground state $\mathcal{X}^0$ and excited state $\mathcal{X}^1$, or a spin state space with spin down $\mathcal{X}^0$ and spin up $\mathcal{X}^1$ states etc.

Definition 7.15. We define the *Toy Fock space*[1] to be the countable tensor product

$$\mathrm{T}\Phi = \bigotimes_{\mathbb{N}^*} \mathbb{C}^2,$$

associated to the stabilizing sequence $(\mathcal{X}^0)_{n\in\mathbb{N}^*}$. In other words, it is the Hilbert space whose Hilbertian orthonormal basis is given by the elements

$$e_1 \otimes e_2 \otimes \cdots \otimes e_n \otimes \cdots$$

where for each i we have $e_i = \mathcal{X}^0$ or $\mathcal{X}^1$, but only a finite number of e_i are equal to $\mathcal{X}^1$.

Definition 7.16. Let us describe the above orthonormal basis in another useful way. Let

$$\mathcal{P} = \mathcal{P}_f(\mathbb{N}^*)$$

be the set of finite subsets of $\mathbb{N}^*$. An element σ of $\mathcal{P}$ is thus a finite sequence of integers $\sigma = \{i_1 < \ldots < i_n\}$. For every $\sigma = \{i_1 < \ldots < i_n\} \in \mathcal{P}$ we put $\mathcal{X}_\sigma$ to be the basis element $e_1 \otimes \cdots \otimes e_n \otimes \cdots$ of $\mathrm{T}\Phi$ given by

$$e_j = \begin{cases} \mathcal{X}^1 & \text{if } j = i_k \text{ for some } k, \\ \mathcal{X}^0 & \text{otherwise.} \end{cases}$$

Note the particular case

$$\mathcal{X}_\emptyset = \mathcal{X}^0 \otimes \ldots \otimes \mathcal{X}^0 \otimes \ldots$$

Every element f of $\mathrm{T}\Phi$ writes in unique way as

$$f = \sum_{\sigma\in\mathcal{P}} f(\sigma)\, \mathcal{X}_\sigma\,,$$

with

$$\|f\|^2_{\mathrm{T}\Phi} = \sum_{\sigma\in\mathcal{P}} |f(\sigma)|^2 < \infty\,.$$

In other words, with these notations, the space $\mathrm{T}\Phi$ naturally identifies to the space $\ell^2(\mathcal{P})$.

The fundamental structure of $\mathrm{T}\Phi$ is the one of a countable tensor product of copies of $\mathbb{C}^2$ with respect to a given basis vector $\mathcal{X}^0$. Actually a more

[1] The toy Fock space was called "bébé Fock" in french, by P.-A. Meyer. Exercise: find the joke.

abstract definition for $\mathrm{T}\Phi$ is possible as follows: $\mathrm{T}\Phi$ is the separable Hilbert space whose orthonormal basis is fixed and chosen to be indexed by $\mathcal{P}$, the set of finite subsets of $\mathbb{N}^*$. This is enough to carry all the properties we need for the space $\mathrm{T}\Phi$.

Definition 7.17. Let $n \leq m$ be any fixed elements of $\mathbb{N}^*$. We denote by $\mathrm{T}\Phi_{[n,m]}$ the subspace of $\mathrm{T}\Phi$ generated by the χ_σ such that $\sigma \subset [n,m]$. The space $\mathrm{T}\Phi_{[n,m]}$ is thus the subspace of $\mathrm{T}\Phi$ made of those $f = \sum_{\sigma\in\mathcal{P}} f(\sigma)\chi_\sigma$ such that $f(\sigma) = 0$ once $\sigma \not\subset [n,m]$.

Among those spaces $\mathrm{T}\Phi_{[n,m]}$, we denote by $\mathrm{T}\Phi_{n]}$ the space $\mathrm{T}\Phi_{[1,n]}$. We also write $\mathrm{T}\Phi_{[n}$ for the space $\mathrm{T}\Phi_{[n,+\infty[}$, with obvious definition.

Following the same idea, we denote by $\mathcal{P}_{[n,m]}$ the set of finite subsets of $\mathbb{N}^* \cap [n,m]$ and we have the corresponding notations $\mathcal{P}_{n]}$ and $\mathcal{P}_{[n}$. We then get the natural identification

$$\mathrm{T}\Phi_{[n,m]} = \ell^2(\mathcal{P}_{[n,m]})\,.$$

For a $\sigma \in \mathcal{P}$, we put $\sigma_{[n,m]} = \sigma \cap [n,m]$. In the same way we define $\sigma_{n]} = \sigma \cap [1,n]$ and $\sigma_{[n} = \sigma \cap [n,+\infty[$.

With these notations, the fundamental structure of $\mathrm{T}\Phi$ is reflected by the following result.

Proposition 7.18. *Let $n < m \in \mathbb{N}^*$ be fixed. The mapping*

$$\begin{array}{rcl} \mathsf{U} : \mathrm{T}\Phi_{n]} \otimes \mathrm{T}\Phi_{[n+1,m-1]} \otimes \mathrm{T}\Phi_{[m} & \longrightarrow & \mathrm{T}\Phi \\ f \otimes g \otimes h & \longmapsto & k \end{array}$$

where

$$k(\sigma) = f(\sigma_{n]})\, g(\sigma_{[n+1,m-1]})\, h(\sigma_{[m})\,,$$

for all $\sigma \in \mathcal{P}$, extends to a unitary operator.

Proof. The operator U is isometric for

$$\begin{aligned}
&\langle \mathsf{U}(f\otimes g\otimes h)\,,\, \mathsf{U}(f'\otimes g'\otimes h')\rangle_{\mathrm{T}\Phi} = \\
&\quad = \sum_{\sigma\in\mathcal{P}} \overline{f}(\sigma_{n]})\, \overline{g}(\sigma_{[n+1,m-1]})\, \overline{h}(\sigma_{[m})\, f'(\sigma_{n]})\, g'(\sigma_{[n+1,m-1]})\, h'(\sigma_{[m}) \\
&\quad = \sum_{\alpha\in\mathcal{P}_{n]}} \overline{f}(\alpha)\, f'(\alpha) \sum_{\beta\in\mathcal{P}_{[n+1,m-1]}} \overline{g}(\beta)\, g'(\beta) \sum_{\gamma\in\mathcal{P}_{[m}} \overline{h}(\gamma)\, h'(\gamma) \\
&\quad = \langle f\,,\, f'\rangle_{\mathrm{T}\Phi_{n]}}\, \langle g\,,\, g'\rangle_{\mathrm{T}\Phi_{[n+1,m-1]}}\, \langle h\,,\, h'\rangle_{\mathrm{T}\Phi_{[m}} \\
&\quad = \langle f\otimes g\otimes h\,,\, f'\otimes g'\otimes h'\rangle_{\mathrm{T}\Phi_{n]}\otimes\mathrm{T}\Phi_{[n+1,m-1]}\otimes\mathrm{T}\Phi_{[m}}\,.
\end{aligned}$$

But U is defined on a dense subspace and thus extends to an isometry on the whole of $\mathrm{T}\Phi_{n]} \otimes \mathrm{T}\Phi_{[n-1,m+1]} \otimes \mathrm{T}\Phi_{[m}$. Its range is dense in $\mathrm{T}\Phi$ for it contains all the χ_σ's. Thus U is unitary. $\square$

We have completely described the structure of the *space* $\mathrm{T}\Phi$. We now turn to the *operators* on $\mathrm{T}\Phi$.

Definition 7.19. We shall make use of a particular basis, which is not the usual Pauli matrix basis, as in the previous subsection. In the orthonormal basis $\{\chi^0, \chi^1\}$ of $\mathbb{C}^2$ we consider the basis of matrices

$$\boxed{a_0^0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad a_1^0 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad a_0^1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad a_1^1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.}$$

In other words, we have

$$\boxed{a_j^i \, \chi^k = \delta_{k,i} \, \chi^j}$$

for all $i, j, k = 0, 1$.

From these basic operators on $\mathbb{C}^2$, we built basic operators on $\mathrm{T}\Phi$ by considering the operators $a_j^i(n)$ which act as a_j^i on the n-th copy of $\mathbb{C}^2$ and as the identity on the other copies. This means, when acting on the natural orthonormal basis of $\mathrm{T}\Phi$

$$\begin{aligned}
a_0^0(n)\chi_\sigma &= \mathbb{1}_{n \notin \sigma} \, \chi_\sigma \\
a_1^0(n)\chi_\sigma &= \mathbb{1}_{n \notin \sigma} \, \chi_{\sigma \cup \{n\}} \\
a_0^1(n)\chi_\sigma &= \mathbb{1}_{n \in \sigma} \, \chi_{\sigma \setminus \{n\}} \\
a_1^1(n)\chi_\sigma &= \mathbb{1}_{n \in \sigma} \, \chi_\sigma \, .
\end{aligned}$$

In some sense the operators $a_j^i(n)$ form a basis of $\mathcal{B}(\mathcal{H})$. Of course, one needs to be precise with the meaning of such a sentence, for now we deal with an infinite dimensional space. This could be done easily in terms of von Neumann algebras: "*The von Neumann algebra generated by the operators $a_j^i(n)$, $i, j = 0, 1$, $n \in \mathbb{N}^*$, is the whole of $\mathcal{B}(\mathrm{T}\Phi)$.*" But this would bring us too far and would need some knowledge on operator algebras which is not required for this lecture.

Let us describe how random walks can be produced with the help of the operators $a_j^i(n)$.

Theorem 7.20. *Let A be an observable on $\mathbb{C}^2$, with coefficients*

$$\mathsf{A} = \sum_{i,j=0,1} \alpha_j^i \, a_j^i .$$

The observable A admits a certain law μ under the state χ^0. For all $n \in \mathbb{N}^*$ define the observable

$$\mathsf{A}(n) = \sum_{i,j=0,1} \alpha_j^i \, a_j^i(n)$$

on $\mathrm{T}\Phi$. Then the sequence $(\mathsf{A}(n))_{n\in\mathbb{N}}$ in $\mathcal{B}(\mathrm{T}\Phi)$ is commutative and, under the state $\chi_\emptyset$, it has the law of a sequence of independent random variables, each of which has the same law μ.

Proof. This is an easy application of Proposition 7.7 and Proposition 7.18 when noticing that:
i) for every $n \in \mathbb{N}^*$, in the following splitting of $\mathrm{T}\Phi$:

$$\mathrm{T}\Phi \simeq \mathrm{T}\Phi_{n-1]} \otimes \mathrm{T}\Phi_{[n,n]} \otimes \mathrm{T}\Phi_{[n+1}$$

the operator $\mathsf{A}(n)$ is of the form $\mathrm{I} \otimes \mathsf{A}(n) \otimes \mathrm{I}$.
ii) the state $\chi_\emptyset$ on $\mathrm{T}\Phi$ is the tensor product $\otimes_{n\in\mathbb{N}^*}\Omega$ of the ground states of each copy of $\mathbb{C}^2$.
Details are left to the reader. $\square$

Of course in the result above one can easily make the α^i_j's depend on n also; this allows to produce any sequence of independent random variables. Summing up these random variables, this provides an easy way of realizing any classical random walk, in law, on $\mathrm{T}\Phi$. We shall see in next subsection that the Toy Fock space provides much more than that: it can realize the multiplication operators of any classical Bernoulli random walk on its canonical space, by means of linear combinations of the operators $a^i_j(n)$.

7.2.3 Probabilistic Interpretations

In this section we explore the so-called *probabilistic interpretations* of $\mathrm{T}\Phi$, that is, we shall realize on $\mathrm{T}\Phi$ the multiplication operators associated to any Bernoulli random walk, by means of only linear combinations of the operators $a^i_j(n)$.

Definition 7.21. Let $p \in]0,1[$ be fixed and $q = 1-p$. Consider a classical Bernoulli sequence $(\nu_n)_{n\in\mathbb{N}^*}$, that is, a sequence of independent identically distributed Bernoulli random variables ν_n with law $p\delta_1 + q\delta_0$. We realize this random walk on its canonical space $(\Omega, \mathcal{F}, \mu_p)$ where $\Omega = \{0,1\}^{\mathbb{N}^*}$, $\mathcal{F}$ is the σ-field generated by finite-based cylinders and μ_p is the only probability measure on $(\Omega, \mathcal{F})$ which makes the coordinate mappings

$$\begin{aligned} \nu_n : \Omega &\longrightarrow \{0,1\} \\ \omega &\longmapsto \omega_n \end{aligned}$$

being independent, identically distributed with law $p\delta_1 + q\delta_0$.

Definition 7.22. We center and normalize ν_n by putting

$$X_n = \frac{\nu_n - p}{\sqrt{pq}}$$

so that $\mathbb{E}[X_n]=0$ and $\mathbb{E}[X_n^2]=1$. The random variables X_n are then independent, they take the values

$$\begin{cases} \sqrt{\frac{q}{p}}\,, & \text{with probability } p\,,\\ -\sqrt{\frac{p}{q}}\,, & \text{with probability } q\,. \end{cases}$$

For every $\sigma\in\mathcal{P}$ we put

$$X_\sigma = \begin{cases} X_{i_1}\cdots X_{i_n} & \text{if } \sigma=\{i_1,\ldots,i_n\}\,,\\ \mathbb{1} & \text{if } \sigma=\emptyset\,, \end{cases}$$

where $\mathbb{1}$ is the deterministic random variable always equal to 1.

Proposition 7.23. *The set $\{X_\sigma\,;\ \sigma\in\mathcal{P}\}$ forms an orthonormal basis of the Hilbert space $L^2(\Omega,\mathcal{F},\mu_p)$.*

Proof. Let us first check that the set $\{X_\sigma\,;\ \sigma\in\mathcal{P}\}$ forms an orthonormal set. Let $\alpha,\beta\in\mathcal{P}$ be fixed, we have

$$\begin{aligned} \langle X_\alpha\,,\,X_\beta\rangle_{L^2(\Omega,\mathcal{F},\mu_p)} &= \mathbb{E}_{\mu_p}[X_\alpha\,X_\beta]\\ &= \mathbb{E}_{\mu_p}[X_{\alpha\setminus\beta}\,(X_{\alpha\cap\beta})^2\,X_{\beta\setminus\alpha}]\\ &= \mathbb{E}_{\mu_p}[X_{\alpha\setminus\beta}]\,\mathbb{E}_{\mu_p}[X^2_{\alpha\cap\beta}]\,\mathbb{E}_{\mu_p}[X_{\beta\setminus\alpha}]\,, \end{aligned}$$

for the random variables X_i and X_j are independent once $i\neq j$. As we have $\mathbb{E}[X_\sigma]=0$ for all $\sigma\neq\emptyset$, we get that $\langle X_\alpha\,,\,X_\beta\rangle_{L^2(\Omega,\mathcal{F},\mu_p)}$ vanishes unless $\alpha\setminus\beta=\beta\setminus\alpha=\emptyset$, that is, unless $\alpha=\beta$. In the case where $\alpha=\beta$ we get

$$\langle X_\alpha\,,\,X_\beta\rangle_{L^2(\Omega,\mathcal{F},\mu_p)} = \mathbb{E}_{\mu_p}[X_\sigma^2]=1\,.$$

This proves the orthonormality.

Let us now prove the totality of the X_σ's. Had we replaced $\mathbb{N}^*$ by $\{1,\ldots,N\}$ in the above definition of Ω, we would directly conclude that the X_σ form a basis for they are orthonormal and have cardinal 2^N, the dimension of the space $L^2(\Omega,\mathcal{F},\mu_p)$ in that case. We conclude in the general case $\Omega=\{0,1\}^{\mathbb{N}^*}$ by noticing that any $f\in L^2(\Omega,\mathcal{F},\mu_p)$ can be approached by finitely supported functions. □

From this proposition we see that there is a natural isomorphism between $L^2(\Omega,\mathcal{F},\mu_p)$ and the toy Fock space $\mathrm{T}\Phi$: it consists in identifying the orthonormal basis $\{X_\sigma\,,\ \sigma\in\mathcal{P}\}$ of $L^2(\Omega,\mathcal{F},\mu_p)$ with the orthonormal basis $\{\chi_\sigma\,,\ \sigma\in\mathcal{P}\}$ of $\mathrm{T}\Phi$. The space $L^2(\Omega,\mathcal{F},\mu_p)$ is now denoted by $\mathrm{T}\Phi_p$ and is called the *p-probabilistic interpretation* of $\mathrm{T}\Phi$. The space $\mathrm{T}\Phi_p$ is an exact reproduction of $\mathrm{T}\Phi$, but it is concretely represented as the space of L^2 functionals of some random walk. The element X_σ of $\mathrm{T}\Phi$ interprets in $\mathrm{T}\Phi_p$ as a concrete random variable X_σ.

All the spaces $\mathrm{T}\Phi_p$ identify pairwise this way and they all identify to $\mathrm{T}\Phi$. But this identification only belongs to the level of Hilbert spaces. As we have already discussed in Subsection 7.1.3, the fact that the random variable X_n in the p-probabilistic interpretation $\mathrm{T}\Phi_p$ is represented by the basis vector $\chi_{\{n\}}$ in $\mathrm{T}\Phi$ does not mean much. All the probabilistic properties of X_n such as its law, its independence with respect to the other X_m's, ... all these informations are lost in this identification. The actual representation of the random variable X_n of $\mathrm{T}\Phi_p$ in $\mathrm{T}\Phi$ should be the push forward of the *multiplication operator* by X_n.

Let us be more precise.

Definition 7.24. Let $\mathsf{U}_p : \mathrm{T}\Phi_p \longmapsto \mathrm{T}\Phi$ be the unitary operator which realizes the basis identification between $\mathrm{T}\Phi_p$ and $\mathrm{T}\Phi$. Let $\mathcal{M}_{X_n}$ be the operator of multiplication by X_n in $\mathrm{T}\Phi_p$. We shall consider the operator

$$\mathcal{M}^p_{X_n} = \mathsf{U}_p\, \mathcal{M}_{X_n}\, \mathsf{U}_p^*$$

on $\mathrm{T}\Phi$, which is the representation of the random variable X_n but in the space $\mathrm{T}\Phi$.

We shall now prove a striking result which shows that these different random variables X_n, for different p's, are represented on $\mathrm{T}\Phi$ by means of a very simple linear combination of the $a^i_j(n)$'s.

Theorem 7.25. *For all $p \in]0,1[$ and every $n \in \mathbb{N}^*$, the operator $\mathcal{M}^p_{X_n}$ on $\mathrm{T}\Phi$ is given by*

$$\boxed{\mathcal{M}^p_{X_n} = a^0_1(n) + a^1_0(n) + c_p\, a^1_1(n)\,,} \tag{7.2}$$

where

$$\boxed{c_p = \frac{q-p}{\sqrt{pq}}\,.}$$

The mapping $p \mapsto c_p$ is a bijection from $]0,1[$ to $\mathbb{R}$.

Proof. Consider a $f \in \mathrm{T}\Phi_p$, the product $X_n\, f$ is clearly determined by the products $X_n X_\sigma$, $\sigma \in \mathcal{P}$. If n does not belong to σ then by definition of our basis we get

$$X_n X_\sigma = X_{\sigma \cup \{n\}}\,,$$

that is, there is nothing to compute. If n belongs to σ then

$$X_n X_\sigma = X_n^2 X_{\sigma \setminus \{n\}}\,.$$

In other words, the product $X_n f$ on $\mathrm{T}\Phi_p$ is determined by the X_n^2, $n \in \mathbb{N}^*$.

Lemma 7.26. *On $\mathrm{T}\Phi_p$ we have, for all $n \in \mathbb{N}^*$*

$$\boxed{X_n^2 = \mathbb{1} + c_p\, X_n} \tag{7.3}$$

where

$$c_p = \frac{q-p}{\sqrt{pq}}\,.$$

Proof (of the lemma). We have on $\mathrm{T}\Phi_p$

$$\begin{aligned} X_n^2 &= \left(\frac{\nu_n - p}{\sqrt{pq}}\right)^2 = \frac{\nu_n^2 - 2p\nu_n + p^2}{pq} \\ &= \frac{(1-2p)\nu_n}{pq} + \frac{p}{q} \\ &= \frac{q-p}{pq}\big(\sqrt{pq}X_n + p\mathbb{1}\big) + \frac{p}{q}\mathbb{1} \\ &= \mathbb{1} + \frac{q-p}{\sqrt{pq}}X_n\,. \end{aligned}$$

This proves (7.3) and the lemma.

Coming back to the proof of the theorem and applying Lemma 7.26, we get for all $\sigma \in \mathcal{P}$

$$\begin{aligned} X_n\, X_\sigma &= \mathbb{1}_{n\notin\sigma}\, X_{\sigma\cup\{n\}} + \mathbb{1}_{n\in\sigma}(\mathbb{1} + c_p X_n)\, X_{\sigma\setminus\{n\}} \\ &= \mathbb{1}_{n\notin\sigma}\, X_{\sigma\cup\{n\}} + \mathbb{1}_{n\in\sigma}\, X_{\sigma\setminus\{n\}} + c_p\mathbb{1}_{n\in\sigma}\, X_\sigma \end{aligned}$$

This means that on $\mathrm{T}\Phi$ we have

$$\mathcal{M}^p_{X_n} X_\sigma = \mathbb{1}_{n\notin\sigma} X_{\sigma\cup\{n\}} + \mathbb{1}_{n\in\sigma} X_{\sigma\setminus\{n\}} + c_p\,\mathbb{1}_{n\in\sigma} X_\sigma\,.$$

We recognize the action of the operator

$$a^0_1(n) + a^1_0(n) + c_p\, a^1_1(n)$$

on the basis of $\mathrm{T}\Phi$. This gives (7.2).

Finally, we have

$$c_p = \frac{q-p}{\sqrt{pq}} = \frac{1-2p}{\sqrt{p(1-p)}}$$

and as a function of p it is easy to check that it is a bijection from $]0,1[$ to $\mathbb{R}$. □

What we have seen throughout this section is fundamental. On a very simple space $\mathrm{T}\Phi$ we have been able to reproduce a continuum of different probabilistic situation: all the Bernoulli random walk with any parameter $p \in]0,1[$. We are able to completely reproduce inside $\mathrm{T}\Phi$ all these different probability spaces, all these different laws. Even more surprising is the fact that this representation is achieved only with the help of linear combinations of three basic processes $(a_0^1(n))_{n\in\mathbb{N}^*}$, $(a_1^0(n))_{n\in\mathbb{N}^*}$, $(a_1^1(n))_{n\in\mathbb{N}_*}$. These three quantum processes play the role of three fundamental quantum Bernoulli random walk from which every classical Bernoulli random walk can be constructed.

7.2.4 Brownian Motion and Poisson Process

The fact that one gets all classical Bernoulli random variables from the Toy Fock space has many important consequences, some of them are fundamental in the theory of quantum noises. We shall give here a taste of it by showing that the toy Fock space gives rise to two fundamental stochastic processes: the Brownian motion and the Poisson process.

First of all, as we wish to give some heuristic of a continuous-time limit, we shall now have our Toy Fock indexed by $h\mathbb{N}^*$, for some small real parameter $h > 0$, instead of $\mathbb{N}^*$.

In Theorem 7.25, consider the case $p = 1/2$. The random variables X_n then take the values ± 1 with probability $1/2$. In that case we get $c_{\frac{1}{2}} = 0$, which means that the operators

$$\mathsf{S}_{nh} = \sum_{i=1}^{n} a_1^0(ih) + a_0^1(ih)\,,\ n \in \mathbb{N}^*\,,$$

are a commutative sequence of observables on $\mathrm{T}\Phi$, their distribution in the state $X_\emptyset$ is the one of a symmetric Bernoulli random walk indexed by $h\mathbb{N}^*$. Even more, they are the multiplication operators by these random walks on their canonical space. This means in particular that all the functional calculus that can be usually derived from these random walks, can be obtained in the same way with the sequence of operators (S_{nh}).

Renormalizing (S_{nh}), we consider the sequence

$$\mathsf{Q}_{nh} = \sum_{i=1}^{n} \sqrt{h}\left(a_1^0(ih) + a_0^1(ih)\right)\,,\ n \in \mathbb{N}^*\,,$$

which is a random walk converging to a Brownian motion when h tends to 0. This is a convergence in law for the associated random walk, but when regarding the operators Q_{nh} this is far more. It is a convergence at the level of multiplication operators, it is a convergence which respects the functional calculus of the Brownian motion in the limit. For example, computing

$$\begin{aligned}\sum_{i=2}^{n}(\mathsf{Q}_{ih}-\mathsf{Q}_{(i-1)h})^2 &= \sum_{i=1}^{n-1} h\left(a_1^0(ih)+a_0^1(ih)\right)^2\\ &= \sum_{i=2}^{n} h\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}_{ih}^2\\ &= \sum_{i=2}^{n} h\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}_{ih}\\ &= h\,\mathrm{I}\,,\end{aligned}$$

gives directly the quadratic variation of the Brownian motion:

$$(\mathrm{d}W_t)^2 = \mathrm{d}t$$

or more precisely

$$[W\,,\,W]_t = t\,.$$

Even the Ito formula is encoded in this matrix representation. For example, the famous formula

$$W_t^2 = 2\int_0^t W_s\,\mathrm{d}W_s + t$$

appears as a simple continuous time limit of the following computation

$$\begin{aligned}\mathsf{Q}_{nh}^2 &= \sum_{i,j=1}^{n} h\left(a_1^0(ih)+a_0^1(ih)\right)\left(a_1^0(jh)+a_0^1(jh)\right)\\ &= 2\sum_{i=2}^{n}\left(\sum_{j=1}^{i-1} h\left(a_1^0(jh)+a_0^1(jh)\right)\right)\left(a_1^0(ih)+a_0^1(ih)\right)\\ &\quad+\sum_{i=1}^{n} h\left(a_1^0(ih)+a_0^1(ih)\right)^2\\ &= 2\sum_{i=2}^{n}\mathsf{Q}_{(i-1)h}\big(\mathsf{Q}(ih)-\mathsf{Q}((i-1)h)\big)+h\,\mathrm{I}\,.\end{aligned}$$

Where the situation becomes even more striking is the way the Poisson process enters into the game also, in the Toy Fock space.

Coming back to the random walks of Theorem 7.25, but with a particular choice of the probability p now. Indeed, let us take $p = h$, where h is the same parameter as the time step of the random walk. In order to make the computation not too huggly we prefer to take

$$p = \frac{h}{1+h}\,,\qquad q = \frac{1}{1+h}\,,$$

so that

$$c_p = \frac{1-h}{\sqrt{h}}\,.$$

The associated random variable X_{nh} takes the values $1/\sqrt{h}$ and $-\sqrt{h}$ with probability p and q respectively. The random walk

$$S_{nh} = \sum_{i=1}^{n} \sqrt{h}\, X_{ih}$$

is a Bernoulli random walk, with time step h, which takes the value $-h$ very often (with probability $1-h$, more or less) and the value 1 very rarely (with probability h, more or less). In the continuous time limit it is a compensated Poisson process.

From the point of view of the Toy Fock space representation we obtain the operators

$$\begin{aligned}
\mathsf{X}_{nh} &= \sum_{i=1}^{n} \sqrt{h}\left(a_1^0(ih) + a_0^1(ih) + c_p\, a_1^1(ih)\right) \\
&= \sum_{i=1}^{n} \sqrt{h}\left(a_1^0(ih) + a_0^1(ih)\right) + (1-h)a_1^1(ih)\,.
\end{aligned}$$

This operator family is supposed to behave like a compensated Poisson process, in the limit $h \to 0$. As a consequence, the family

$$\mathsf{N}_{nh} = \mathsf{X}_{nh} + nh\mathrm{I} = \sum_{i=1}^{n} \sqrt{h}\left(a_1^0(ih) + a_0^1(ih)\right) + a_1^1(ih) + ha_0^0(ih)$$

should represent a standard Poisson process in the limit. In the same way as for the Brownian motion, one can recover from these operators all the properties of the Poisson process, of its functional calculus. For example, the matrix

$$\mathsf{A} = \sqrt{h}\left(a_1^0(ih) + a_0^1(ih)\right) + a_1^1(ih) + ha_0^0(ih) = \begin{pmatrix} h & \sqrt{h} \\ \sqrt{h} & 1 \end{pmatrix}_{(ih)}$$

satisfies

$$\mathsf{A}^2 = (1+h)\mathsf{A}\,.$$

This relation gives in the continuous time limit the fundamental characterization of the Poisson process

$$(\mathrm{d}N_t)^2 = \mathrm{d}N_t\,,$$

or else

$$[N\,,\,N]_t = N_t\,.$$

7.3 Higher Level Chains

The aim of this section is to present the natural extension of the Toy Fock space structure, but for higher number of levels on each sites. From a quantum mechanical point of view this is clearly a richer physical structure. From the probabilistic interpretation point of view it gives access to all random walks in $\mathbb{R}^N$. But the interesting point is that the probabilistic interpretations of the $N+1$-level chain gives rise to a very particular probabilistic structure: the *obtuse random variables*.

7.3.1 Structure of Higher Level Chains

Let $\mathcal{H}$ be any separable Hilbert space. Let us fix a particular Hilbertian basis $(\mathcal{X}^i)_{i\in\mathcal{N}\cup\{0\}}$ for the Hilbert space $\mathcal{H}$, where we assume (for notational purposes) that $0 \notin \mathcal{N}$. This particular choice of notations is motivated both by physical interpretations (we see the $\mathcal{X}^i$, $i \in \mathcal{N}$, as representing, for example, different possible excited states of a quantum system, the vector $\mathcal{X}^0$ represents the "ground state" of the quantum system) and by a mathematical necessity (for defining a countable tensor product of Hilbert space one needs to specify one vector in each Hilbert space).

Definition 7.27. Let $\mathrm{T}\Phi$ be the tensor product $\bigotimes_{\mathbb{N}^*} \mathcal{H}$ with respect to the stabilizing sequence $\mathcal{X}^0$. In other words, an orthonormal basis of $\mathrm{T}\Phi$ is given by the family $\{\mathcal{X}_\sigma\,;\ \sigma \in \mathcal{P}\}$ where

– the set $\mathcal{P} = \mathcal{P}(\mathbb{N}^*, \mathcal{N})$ is the set of finite subsets $\{(n_1, i_1), \ldots, (n_k, i_k)\}$ of $\mathbb{N}^* \times \mathcal{N}$ such that the n_i's are mutually different. Another way to describe the set $\mathcal{P}$ is to identify it to the set of sequences $(\sigma_n)_{n\in\mathbb{N}^*}$ with values in $\mathcal{N} \cup \{0\}$ which take a value different from 0 only finitely many times;

– $\mathcal{X}_\sigma$ denotes the vector

$$\mathcal{X}^0 \otimes \ldots \otimes \mathcal{X}^0 \otimes \mathcal{X}^{i_1} \otimes \mathcal{X}^0 \otimes \ldots \otimes \mathcal{X}^0 \otimes \mathcal{X}^{i_2} \otimes \ldots$$

where $\mathcal{X}^{i_1}$ appears in the n_1-th copy of $\mathcal{H}$, where $\mathcal{X}^{i_2}$ appears in the n_2-th copy of $\mathcal{H}$ etc.

The physical signification of this basis is easy to understand: we have a chain of quantum systems, indexed by $\mathbb{N}^*$. The space $\mathrm{T}\Phi$ is the state space of this chain, the vector $\mathcal{X}_\sigma$ with $\sigma = \{(n_1, i_1), \ldots, (n_k, i_k)\}$ represents the state in which exactly k sites are excited: the site n_1 in the state $\mathcal{X}^{i_1}$, the site n_2 in the state $\mathcal{X}^{i_2}$ etc, all the other sites being in their ground state $\mathcal{X}^0$.

Definition 7.28. This particular choice of a basis gives $\mathrm{T}\Phi$ the same particular structure as the spin chain. If we denote by $\mathrm{T}\Phi_{n]}$ the space generated by the $\mathcal{X}_\sigma$ such that $\sigma \subset \{1, \ldots, n\} \times \mathcal{N}$ and by $\mathrm{T}\Phi_{[m}$ the one generated by the

χ_σ such that $\sigma \subset \{m, m+1, \ldots\} \times \mathcal{N}$, we get an obvious natural isomorphism between $\mathrm{T}\Phi$ and $\mathrm{T}\Phi_{n-1]} \otimes \mathrm{T}\Phi_{[n}$ given by

$$[f \otimes g](\sigma) = f\left(\sigma \cap \{1, \ldots, n-1\} \times \mathcal{N}\right) g\left(\sigma \cap \{n, \ldots\} \times \mathcal{N}\right).$$

Definition 7.29. Put $\{a^i_j; i, j \in \mathcal{N} \cup \{0\}\}$ to be the natural basis of $\mathcal{B}(\mathcal{H})$, that is,

$$a^i_j(\chi^k) = \delta_{i,k}\, \chi^j$$

for all $i, j, k \in \mathcal{N} \cup \{0\}$. We denote by $a^i_j(n)$ the natural ampliation of the operator a^i_j to $\mathrm{T}\Phi$ which acts on the copy number n as a^i_j and which acts as the identity on the other copies. That is, in terms of the basis χ_σ,

$$a^i_j(n)\chi_\sigma = 1\!\!1_{(n,i)\in\sigma}\, \chi_{\sigma\setminus(n,i)\cup(n,j)}$$

if neither i nor j is zero, and

$$\begin{aligned} a^i_0(n)\chi_\sigma &= 1\!\!1_{(n,i)\in\sigma}\, \chi_{\sigma\setminus(n,i)},\\ a^0_j(n)\chi_\sigma &= 1\!\!1_{n\notin\sigma}\, \chi_{\sigma\cup(n,j)},\\ a^0_0(n)\chi_\sigma &= 1\!\!1_{n\notin\sigma}\, \chi_\sigma, \end{aligned}$$

where $n \notin \sigma$ actually means "for all i in $\mathcal{N}$, $(n,i) \notin \mathcal{N}$"; we could equivalently use the notation $(n,0) \in \sigma$ instead of $n \notin \sigma$, having in mind the interpretation of σ as a sequence in $\mathcal{N}$ with finitely many non-zero terms.

In the case $\mathcal{N} = \{1, \ldots, N\}$ we say that $\mathrm{T}\Phi$ is *Toy Fock space with multiplicity* $N+1$. In the case $\mathcal{N} = \mathbb{N}^*$ we say that $\mathrm{T}\Phi$ is the *Toy Fock space with infinite multiplicity.*

7.3.2 Obtuse Systems

The probabilistic interpretations of higher level chains is carried, in a natural way, by particular random variables, the *obtuse random variables.* We shall see in the next subsections that these random variables have very interesting algebraic properties which characterize their behavior (and which encodes the behavior of their continuous-time limit). Before entering into all those properties, we start with *obtuse systems.*

Definition 7.30. Let $N \in \mathbb{N}^*$ be fixed. An *obtuse system in* $\mathbb{R}^N$ is a family of $N+1$ vectors $v_1, \ldots, v_{N+1}$ such that

$$\langle v_i\,,\, v_j\rangle = -1$$

for all $i \neq j$. In that case we put

$$\widehat{v}_i = \begin{pmatrix} 1 \\ v_i \end{pmatrix} \in \mathbb{R}^{N+1}\,,$$

so that

$$\langle \widehat{v}_i\,,\,\widehat{v}_j\rangle = 0$$

for all $i \neq j$. They then form an orthogonal basis of $\mathbb{R}^{N+1}$. We put

$$p_i = \frac{1}{\|\widehat{v}_i\|^2} = \frac{1}{1+\|v_i\|^2}\,,$$

for $i = 1, \ldots N+1$.

Lemma 7.31. *With the notations above, we have*

$$\sum_{i=1}^{N+1} p_i = 1 \tag{7.4}$$

and

$$\sum_{i=1}^{N+1} p_i\, v_i = 0\,. \tag{7.5}$$

Proof. We have, for all j,

$$\left\langle \sum_{i=1}^{N+1} p_i\,\widehat{v}_i\,,\,\widehat{v}_j \right\rangle = p_j \|\widehat{v}_j\|^2 = 1 = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}\,,\,\widehat{v}_j \right\rangle\,.$$

As the $\widehat{v}_j$'s form a basis, this means that

$$\sum_{i=1}^{N+1} p_i\,\widehat{v}_i = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\,.$$

This gives the two announced equalities. □

Lemma 7.32. *With the notations above, we also have*

$$\sum_{i=1}^{N+1} p_i\,|v_i\rangle\langle v_i| = \mathrm{I}_{\mathbb{R}^N}\,. \tag{7.6}$$

Proof. As the vectors $(\sqrt{p_i}\,\widehat{v}_i)_{i\in\{1,\ldots,N+1\}}$ form an orthonormal basis of $\mathbb{R}^{N+1}$ we have

$$\mathrm{I}_{\mathbb{R}^{N+1}} = \sum_{i=1}^{N+1} p_i\,|\widehat{v}_i\rangle\langle\widehat{v}_i|\,.$$

Now, for all $i = 1, \ldots, N+1$, put

$$u = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \widetilde{v}_i = \begin{pmatrix} 0 \\ v_i \end{pmatrix} ,$$

so that $\widehat{v}_i = u + \widetilde{v}_i$. We get

$$\begin{aligned} \mathrm{I}_{\mathbb{R}^{N+1}} &= \sum_{i=1}^{N+1} p_i \, |u + \widetilde{v}_i\rangle\langle u + \widetilde{v}_i| \\ &= \sum_{i=1}^{N+1} p_i \, |u\rangle\langle u| + \sum_{i=1}^{N+1} p_i \, |u\rangle\langle \widetilde{v}_i| + \sum_{i=1}^{N+1} p_i \, |\widetilde{v}_i\rangle\langle u| + \sum_{i=1}^{N+1} p_i \, |\widetilde{v}_i\rangle\langle \widetilde{v}_i| \, . \end{aligned}$$

Using (7.4) and (7.5), this simplifies into

$$\mathrm{I}_{\mathbb{R}^{N+1}} = |u\rangle\langle u| + \sum_{i=1}^{N+1} p_i \, |\widetilde{v}_i\rangle\langle \widetilde{v}_i| \, .$$

In particular we have

$$\sum_{i=1}^{N+1} p_i \, |v_i\rangle\langle v_i| = \mathrm{I}_{\mathbb{R}^N} \, ,$$

that is, the announced equality. $\square$

Let us consider two examples (that we shall meet again later in this section). On $\mathbb{R}^2$, the 3 vectors

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \qquad v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} , \qquad v_3 = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$$

form an obtuse system of $\mathbb{R}^2$. The associated p_i's are then respectively

$$p_1 = \frac{1}{2} \, , \qquad p_2 = \frac{1}{3} \, , \qquad p_3 = \frac{1}{6} \, .$$

We shall be interested also in another example. Let $h > 0$ be a parameter, which shall be though of as small. On $\mathbb{R}^2$, associated to the probabilities

$$p_1 = \frac{1}{2} \, , \qquad p_2 = h \, , \qquad p_3 = \frac{1}{2} - h \, ,$$

the 3 vectors

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \qquad v_2 = \begin{pmatrix} -1 \\ \left(\frac{1-2h}{h}\right)^{1/2} \end{pmatrix} \qquad v_3 = \begin{pmatrix} -1 \\ -2\left(\frac{h}{1-2h}\right)^{1/2} \end{pmatrix}$$

form an obtuse system of $\mathbb{R}^2$.

We end this subsection with some useful properties of obtuse systems. First, we prove an independence property for obtuse systems.

Proposition 7.33. *Every strict sub-family of an obtuse family is linearly free.*

Proof. Let $\{v_1, \ldots, v_{N+1}\}$ be an obtuse family of $\mathbb{R}^N$. Let us show that $\{v_1, \ldots, v_N\}$ is free, which would be enough for our claim.

If we had $v_N = \sum_{i=1}^{N-1} \lambda_i \, v_i$ then, taking the scalar product with v_N we get $\|v_N\|^2 = \sum_{i=1}^{N-1} -\lambda_i$, whereas taking the scalar product with v_{N+1} gives $-1 = \sum_{i=1}^{N-1} -\lambda_i$, whence a contradiction. $\square$

Now we prove a kind of uniqueness result for obtuse systems.

Proposition 7.34. *Let $\{v_1, \ldots, v_{N+1}\}$ be an obtuse system of $\mathbb{R}^N$ having $\{p_1, \ldots, p_{N+1}\}$ as associated probabilities. Then the following assertions are equivalent.*

i) The family $\{w_1, \ldots, w_{N+1}\}$ is an obtuse system on $\mathbb{R}^N$ with same respective probabilities $\{p_1, \ldots, p_{N+1}\}$.

ii) There exists an orthogonal operator U on $\mathbb{R}^N$ such that $w_i = \mathsf{U}\, v_i$, for all $i = 1, \ldots, N+1$.

Proof. One direction is obvious. If $w_i = \mathsf{U}\, v_i$, for all $i = 1, \ldots, N+1$ and for some orthogonal operator U, then the scalars products $\langle v_i \, , \, v_j \rangle$ and $\langle w_i \, , \, w_j \rangle$ are equal, for each pair (i, j). This shows that $\{w_1, \ldots, w_{N+1}\}$ is an obtuse system with the same probabilities.

In the converse direction, if $v_1, \ldots, v_{N+1}$ and $w_1, \ldots, w_{N+1}$ are obtuse systems associated to the same probabilities $p_1, \ldots, p_{N+1}$, then

$$\langle v_i \, , \, v_j \rangle = \langle w_i \, , \, w_j \rangle$$

for all i, j. The sub-families $\{v_1, \ldots, v_N\}$ and $\{w_1, \ldots, w_N\}$ are bases of $\mathbb{R}^N$, by Proposition 7.33, and their elements have same respective norms. The map $\mathsf{U} : \mathbb{R}^N \to \mathbb{R}^N$, such that $\mathsf{U}\, v_i = w_i$ for all $i = 1, \ldots, N$, is thus orthogonal, by the conservation of scalar products. Finally, the vector v_{N+1} is a certain linear combination of the v_i's, $i = 1, \ldots, N$, but w_{N+1} is the same linear combination of the w_i's, $i = 1, \ldots, N$, by the conservation of scalar products. Hence $\mathsf{U}\, v_{N+1}$ is equal to w_{N+1} and the proposition is proved. $\square$

7.3.3 Obtuse Random Variables

Obtuse systems are strongly related to a certain class of random variables.

Definition 7.35. Consider a random variable X, with values in $\mathbb{R}^N$. We shall denote by $X^1, \ldots, X^N$ the coordinates of X in $\mathbb{R}^N$. We say that X is *centered* if its expectation is 0, that is, if $\mathbb{E}[X^i] = 0$ for all i. We say that X is *normalized* if its covariance matrix is I, that is, if

$$\operatorname{cov}(X^i, X^j) = \mathbb{E}[X^i\, X^j] - \mathbb{E}[X^i]\,\mathbb{E}[X^j] = \delta_{i,j}\,,$$

for all $i, j = 1, \ldots N$.

Definition 7.36. Consider a random variable X, with values in $\mathbb{R}^N$, which can take only $N+1$ different non-null values $v_1, \ldots, v_{N+1}$, with strictly positive probability $p_1, \ldots, p_{N+1}$, respectively. We consider the canonical version of X, that is, we consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where $\Omega = \{1, \ldots, N+1\}$, where $\mathcal{F}$ is the full σ-algebra of Ω, where the probability measure $\mathbb{P}$ is given by $\mathbb{P}(\{i\}) = p_i$ and the random variable X is given by $X(i) = v_i$, for all $i \in \Omega$. The coordinates of v_i are denoted by v_i^k, for $k = 1, \ldots, N$, so that $X^k(i) = v_i^k$.

In the same way as previously, we put

$$\widehat{v}_i = \begin{pmatrix} 1 \\ v_i \end{pmatrix} \in \mathbb{R}^{N+1}\,,$$

for all $i = 1, \ldots, N+1$.

We shall also consider the deterministic random variable X^0 on $(\Omega, \mathcal{F}, \mathbb{P})$, which is always equal to 1. For $i = 0, \ldots, N$ let $\widetilde{X}^i$ be the random variable defined by

$$\widetilde{X}^i(j) = \sqrt{p_j}\, X^i(j)$$

for all $i = 0, \ldots, N$ and all $j = 1, \ldots, N+1$.

Proposition 7.37. *With the notations above, the following assertions are equivalent.*

1) X is centered and normalized.

2) The $(N+1)\times(N+1)$-matrix $\left(\widetilde{X}^i(j)\right)_{i,j}$ is an orthogonal matrix.

3) The $(N+1)\times(N+1)$-matrix $\left(\sqrt{p_i}\,\widehat{v}_i^j\right)_{i,j}$ is an orthogonal matrix.

4) The family $\{v_1, \ldots, v_{N+1}\}$ is an obtuse system with

$$p_i = \frac{1}{1 + \|v_i\|^2}\,,$$

for all $i = 1, \ldots, N+1$.

Proof.

1) $\Rightarrow$ 2): Since the random variable X is centered and normalized, each component X^i has a zero mean and the scalar product in L^2 between two components X^i and X^j is equal to $\delta_{i,j}$. Hence, for all i in $\{1, \ldots, N\}$, we get

$$\mathbb{E}[X^i] = 0 \quad \Longleftrightarrow \quad \sum_{k=1}^{N+1} p_k\, v_k^i = 0\,, \tag{7.7}$$

and for all $i, j = 1, \ldots N$,

$$\mathbb{E}[X^i\, X^j] = \delta_{i,j} \quad \Longleftrightarrow \quad \sum_{k=1}^{N+1} p_k\, v_k^i\, v_k^j = \delta_{i,j}\,. \tag{7.8}$$

Now, using Eqs. (7.7) and (7.8), we get, for all $i, j = 1, \ldots, N$

$$\begin{aligned}
\left\langle \widetilde{X}^0\,,\, \widetilde{X}^0 \right\rangle &= \sum_{k=1}^{N+1} p_k = 1\,,\\
\left\langle \widetilde{X}^0\,,\, \widetilde{X}^i \right\rangle &= \sum_{k=1}^{N+1} \sqrt{p_k}\,\sqrt{p_k}\, v_k^i = 0\,,\\
\left\langle \widetilde{X}^i\,,\, \widetilde{X}^j \right\rangle &= \sum_{k=1}^{N+1} \sqrt{p_k}\, v_k^j\, \sqrt{p_k} v_k^i = \delta_{i,j}\,.
\end{aligned}$$

The orthogonal character follows immediately.

2) $\Rightarrow$ 1): Conversely, if the matrix $\left(\widetilde{X}^i(j)\right)_{i,j}$ is orthogonal, the scalar products of column vectors give the mean 0 and the covariance I for the random variable X.

2) $\Leftrightarrow$ 3): The matrix $\left(\sqrt{p_j}\,\widehat{v}_i^j\right)_{i,j}$ is the transpose matrix of $\left(\widetilde{X}^i(j)\right)_{i,j}$. Therefore, if one of these two matrices is orthogonal, its transpose matrix is orthogonal too.

3) $\Leftrightarrow$ 4): The matrix $\left(\sqrt{p_j}\,\widehat{v}_j^i\right)_{i,j}$ is orthogonal if and only if

$$\left\langle \sqrt{p_i}\,\widehat{v}_i\,,\, \sqrt{p_j}\,\widehat{v}_j \right\rangle = \delta_{i,j}\,,$$

for all $i, j = 1, \ldots, N+1$. On the other hand, the condition $\left\langle \sqrt{p_i}\,\widehat{v}_i\,,\, \sqrt{p_i}\,\widehat{v}_i \right\rangle = 1$ is equivalent to $p_i\,(1+\|v_i\|^2) = 1$, whereas the condition $\left\langle \sqrt{p_i}\,\widehat{v}_i\,,\, \sqrt{p_j}\,\widehat{v}_j \right\rangle = 0$ is equivalent to $\sqrt{p_i}\,\sqrt{p_j}\,(1+\langle v_i\,,\, v_j\rangle) = 0$, that is, $\langle v_i\,,\, v_j\rangle = -1$. This gives the result. □

Definition 7.38. Because of the equivalence between 1) and 4) above, the random variables in $\mathbb{R}^N$ which take only $N+1$ different values with strictly positive probability, which are centered and normalized, are called *obtuse random variables in* $\mathbb{R}^N$.

7.3.4 Generic Character of Obtuse Random Variables

We shall here apply the properties of obtuse systems to obtuse random variable, in order to show that these random variables somehow generate all the finitely supported probability distributions on $\mathbb{R}^N$.

First of all, an immediate consequence of Proposition 7.34 is that obtuse random variables on $\mathbb{R}^N$ with a prescribed probability distribution $\{p_1, \ldots, p_{N+1}\}$ are essentially unique.

Proposition 7.39. *Let X be an obtuse random variable of $\mathbb{R}^N$ with associated probabilities $\{p_1, \ldots, p_{N+1}\}$. Then the following assertions are equivalent.*

i) The random variable Y is an obtuse random variable on $\mathbb{R}^N$ with same probabilities $\{p_1, \ldots, p_{N+1}\}$.

ii) There exists an orthogonal operator U on $\mathbb{R}^N$ such that $Y = \mathsf{U}X$ in distribution.

Having proved that uniqueness, we shall now prove that obtuse random variables generate all the random variables (at least with finite support). First of all, a rather simple remark which shows that the choice of taking $N+1$ different values is the minimal one for centered and normalized random variables in $\mathbb{R}^N$.

Proposition 7.40. *Let X be a centered and normalized random variable in $\mathbb{R}^d$, taking n different values. Then we must have*

$$n \geq d + 1\,.$$

Proof. Let X be centered and normalized in $\mathbb{R}^d$, taking the values $v_1, \ldots, v_n$ with probabilities $p_1, \ldots, p_n$ and with $n \leq d$, that is, $n < d + 1$. The relation

$$0 = \mathbb{E}[X] = \sum_{i=1}^{n} p_i v_i$$

shows that $\mathrm{Rank}\{v_1, \ldots, v_n\} < n$. On the other hand, the relation

$$\mathrm{I}_{\mathbb{C}^d} = \mathbb{E}\left[|X\rangle\langle X|\right] = \sum_{i=1}^{n} p_i |v_i\rangle\langle v_i|$$

shows that $\mathrm{Rank}\{v_1, \ldots, v_n\} \geq d$. This gives the result. □

We can now state the theorem which shows how general, finitely supported, random variables on $\mathbb{R}^d$ are generated by the obtuse ones. We concentrate only on centered and normalized random variables, for they obviously generate all the others, up to an affine transform of $\mathbb{R}^d$.

Theorem 7.41. *Let $n \geq d+1$ and let X be a centered and normalized random variable in $\mathbb{R}^d$, taking n different values $v_1, \ldots, v_n$, with probabilities $p_1, \ldots, p_n$ respectively.*

If Y is any obtuse random variable on $\mathbb{R}^{n-1}$ associated to the probabilities $p_1, \ldots, p_n$, then there exists a partial isometry A from $\mathbb{R}^{n-1}$ to $\mathbb{R}^d$, with $\operatorname{Ran} A = \mathbb{R}^d$*, such that*

$$X = \mathsf{A}\, Y$$

in distribution.

Proof. Assume that the obtuse random variable Y takes the values $w_1, \ldots, w_n$ in $\mathbb{R}^{n-1}$. The family $\{w_1, \ldots, w_{n-1}\}$ is linearly independent by Proposition 7.33, hence there exists a linear map $\mathsf{A} : \mathbb{R}^{n-1} \to \mathbb{R}^d$ such that $\mathsf{A}\, w_i = v_i$ for all $i < n$. Now we have

$$p_n\, v_n = -\sum_{i<n} p_i\, v_i = -\sum_{i<n} p_i\, \mathsf{A}\, w_i = \mathsf{A}\left(-\sum_{i<n} p_i\, w_i\right) = p_n\, \mathsf{A}\, w_n\,.$$

Hence the relation $\mathsf{A}\, w_i = v_i$ holds for all $i \leq n$.

We have proved the relation $X = \mathsf{A}\, Y$ in distribution, with A being a linear map from $\mathbb{R}^{n-1}$ to $\mathbb{R}^d$. The fact that X is normalized can be written as $\mathbb{E}[XX^*] = \mathrm{I}_d$. But

$$\mathbb{E}[XX^*] = \mathbb{E}[\mathsf{A}\, YY^* \mathsf{A}^*] = \mathsf{A}\, \mathbb{E}[YY^*]\, \mathsf{A}^* = \mathsf{A}\, \mathrm{I}_n\, \mathsf{A}^* = \mathsf{A}\, \mathsf{A}^*\,.$$

Hence A must satisfy $\mathsf{A}\, \mathsf{A}^* = \mathrm{I}_d$, which is exactly saying that A is a partial isometry with range $\mathbb{R}^d$. $\square$

7.3.5 Doubly-symmetric 3-tensors

Obtuse random variables are naturally associated to some 3-tensors with particular symmetries. This is what we shall prove here.

Definition 7.42. A *3-tensor on* $\mathbb{R}^n$ is an element of $(\mathbb{R}^N)^* \otimes \mathbb{R}^N \otimes \mathbb{R}^N$, that is, a linear map from $\mathbb{R}^N$ to $\mathbb{R}^N \otimes \mathbb{R}^N$. Coordinate-wise, it is represented by a collection of coefficients $(\mathsf{S}^{ij}_k)_{i,j,k=1,\ldots,n}$. It acts on $\mathbb{R}^N$ as

$$(\mathsf{S}(x))^{ij} = \sum_{k=1}^{n} \mathsf{S}^{ij}_k x^k\,.$$

We shall see below that obtuse random variables on $\mathbb{R}^N$ have a naturally associated 3-tensor on $\mathbb{R}^{N+1}$. Note that, because of our notation choice $X^0, X^1, \ldots, X^N$, the 3-tensor is indexed by $\{0, 1, \ldots, N\}$ instead of $\{1, \ldots, N+1\}$.

Proposition 7.43. *Let X be an obtuse random variable in $\mathbb{R}^N$. Then there exists a unique 3-tensor S on $\mathbb{R}^{N+1}$ such that*

$$X^i\, X^j = \sum_{k=0}^N \mathsf{S}_k^{ij}\, X^k\,, \tag{7.9}$$

for all $i, j = 0, \ldots, N$. This 3-tensor S is given by

$$\mathsf{S}_k^{ij} = \mathbb{E}[X^i\, X^j\, X^k]\,, \tag{7.10}$$

for all $i, j, k = 0, \ldots N$.

Proof. As X is an obtuse random variable, that is, a centered and normalized random variable in $\mathbb{R}^N$ taking exactly $N+1$ different values, the random variables $\{X^0, X^1, \ldots, X^N\}$ are orthonormal in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, hence they form an orthonormal basis of $L^2(\Omega, \mathcal{F}, \mathbb{P})$, for the latter space is $N+1$-dimensional. These random variables being bounded, the products $X^i\, X^j$ are still elements of $L^2(\Omega, \mathcal{F}, \mathbb{P})$, hence they can be written, in a unique way, as linear combinations of the X^k's. As a consequence, there exists a unique 3-tensor S on $\mathbb{R}^{N+1}$ such that

$$X^i\, X^j = \sum_{k=0}^N \mathsf{S}_k^{ij}\, X^k$$

for all $i, j = 0, \ldots, N$. In particular we have

$$\mathbb{E}[X^i\, X^j\, X^k] = \sum_{l=0}^N \mathsf{S}_l^{ij}\, \mathbb{E}[X_l\, X_k] = \mathsf{S}_k^{ij}\,.$$

This shows Identity (7.10). □

This 3-tensor S has quite some symmetries, let us detail them.

Proposition 7.44. *Let S be the 3-tensor associated to an obtuse random variable X on $\mathbb{R}^N$. Then the 3-tensor S satisfies the following relations, for all $i, j, k, l = 0, \ldots, N$*

$$\mathsf{S}_0^{ij} = \delta_{ij}\,, \tag{7.11}$$

$$\mathsf{S}_k^{ij} \quad \textit{is symmetric in} \quad (i, j, k)\,, \tag{7.12}$$

$$\sum_{m=0}^N \mathsf{S}_m^{ij}\, \mathsf{S}_m^{kl} \quad \textit{is symmetric in} \quad (i, j, k, l)\,, \tag{7.13}$$

Proof.

Relation (7.11) is immediate for

$$\mathsf{S}_0^{ij} = \mathbb{E}[X^i\, X^j] = \delta_{ij}\,.$$

Equation (7.12) comes directly from Formula (7.10) which shows a clear symmetry in (i, j, k).
By (7.9) we have

$$X^i\, X^j = \sum_{m=0}^{N} \mathsf{S}^{ij}_m\, X^m\,,$$

whereas

$$X^k\, X^l = \sum_{n=0}^{N} \mathsf{S}^{kl}_n\, X^n\,.$$

Altogether this gives

$$\mathbb{E}\left[X^i\, X^j\, X^k\, X^l\right] = \sum_{m=0}^{N} \mathsf{S}^{ij}_m\, \mathsf{S}^{kl}_m\,.$$

But the left hand side is clearly symmetric in (i, j, k, l) and (7.13) follows.
□

7.3.6 The Main Diagonalization Theorem

We are going to leave for a moment the obtuse random variables and concentrate only on the symmetries we have obtained above. The relation (7.11) is really specific to obtuse random variables, we shall leave it for a moment. We concentrate on the relations (7.12) and (7.13) which have important consequences for the 3-tensor.

Definition 7.45. A 3-tensor S on $\mathbb{R}^{N+1}$ which satisfies (7.12) and (7.13) is called a *doubly-symmetric 3-tensor on* $\mathbb{R}^{N+1}$.

The main result concerning doubly-symmetric 3-tensors in $\mathbb{R}^{N+1}$ is that they are the exact generalization for 3-tensors of normal matrices for 2-tensors: they are exactly those 3-tensors which can be diagonalized in some orthonormal basis of $\mathbb{R}^{N+1}$.

Definition 7.46. A 3-tensor S on $\mathbb{R}^{N+1}$ is said to be *diagonalizable in some orthonormal basis* $(a_m)_{m=0}^N$ of $\mathbb{R}^{N+1}$ if there exist real numbers $(\lambda_m)_{m=0}^N$ such that

$$\mathsf{S} = \sum_{m=0}^{N} \lambda_m\, a_m^* \otimes a_m \otimes a_m\,. \tag{7.14}$$

In other words

$$\mathsf{S}(x) = \sum_{m=0}^{N} \lambda_m\, \langle a_m\,,\, x\rangle\, a_m \otimes a_m \tag{7.15}$$

for all $x \in \mathbb{R}^{N+1}$.

Note that, as opposed to the case of 2-tensors (that is, matrices), the "eigenvalues" λ_m are not completely determined by the representation (7.15). Indeed, if we put $\widetilde{a}_m = s_m\, a_m$ for all m, where the s_m's are any modulus one real numbers, then the $\widetilde{a}_m$'s still form an orthonormal basis of $\mathbb{R}^{N+1}$ and we have

$$\mathsf{S}(x) = \sum_{m=0}^{N} s_m\, \lambda_m\, \langle \widetilde{a}_m\,,\, x\rangle\; \widetilde{a}_m \otimes \widetilde{a}_m\,.$$

Hence the λ_m's are only determined up to a sign, only their modulus is determined by the representation (7.15).

Definition 7.47. Actually, there are more natural objects that can be associated to diagonalizable 3-tensors; they are the *orthogonal families in* $\mathbb{R}^N$. Indeed, if S is diagonalizable as above, for all m such that $\lambda_m \neq 0$ put $v_m = \lambda_m\, a_m$. The family $\{v_m\,;\ m = 1, \ldots, K\}$ is then an orthogonal family in $\mathbb{R}^{N+1}$ and we have

$$\mathsf{S}(v_m) = v_m \otimes v_m$$

for all m. In terms of the v_m's, the decomposition (7.15) of S becomes

$$\mathsf{S}(x) = \sum_{m=1}^{K} \frac{1}{\|v_m\|^2}\ \langle v_m\,,\, x\rangle\; v_m \otimes v_m\,. \tag{7.16}$$

This is the form of diagonalization we shall retain for 3-tensors. Be aware that in the above representation the vectors are orthogonal, but not normalized anymore. Also note that they only represent the eigenvectors of S associated to non-vanishing eigenvalues.

We can now state the main theorem.

Theorem 7.48. *A 3-tensor S on $\mathbb{R}^{N+1}$ is diagonalizable in some orthonormal basis if and only if it is doubly-symmetric.*

More precisely, the formulas

$$\mathcal{V} = \left\{v \in \mathbb{R}^{N+1} \setminus \{0\}\,;\ \mathsf{S}(v) = v \otimes v\right\}\,,$$

and

$$\mathsf{S}(x) = \sum_{v \in \mathcal{V}} \frac{1}{\|v\|^2}\ \langle v\,,\, x\rangle\; v \otimes v\,,$$

establish a bijection between the set of orthogonal systems $\mathcal{V}$ in $\mathbb{R}^{N+1} \setminus \{0\}$ and the set of complex doubly-symmetric 3-tensors S.

Proof. First step: let $\mathcal{V} = \{v_m\,;\, m = 1, \ldots, K\}$ be an orthogonal family in $\mathbb{R}^{N+1} \setminus \{0\}$. Put

$$\mathsf{S}_k^{ij} = \sum_{m=1}^{K} \frac{1}{\|v_m\|^2}\, v_m^i\, v_m^j\, v_m^k\,,$$

for all $i, j, k = 0, \ldots, N$. We shall check that S is a complex doubly-symmetric 3-tensor in $\mathbb{R}^N$. The symmetry of S^{ij}_k in i, j, k is obvious from the definition. This gives (7.12).

We have

$$\begin{aligned}
\sum_{m=0}^N \mathsf{S}^{ij}_m \, \mathsf{S}^{kl}_m &= \sum_{m=0}^N \sum_{n,p=1}^K \frac{1}{\|v_n\|^2} \, \frac{1}{\|v_p\|^2} \, v_n^i \, v_n^j \, v_n^m \, v_p^m \, v_p^k \, v_p^l \\
&= \sum_{n,p=1}^K \frac{1}{\|v_n\|^2} \, \frac{1}{\|v_p\|^2} \, v_n^i \, v_n^j \, \langle v_p \, , \, v_n \rangle \, v_p^k \, v_p^l \\
&= \sum_{n=1}^K \frac{1}{\|v_n\|^2} \, v_n^i \, v_n^j \, v_n^k \, v_n^l
\end{aligned}$$

and the symmetry in i, j, k, l is obvious. This gives (7.13).

We have proved that the formula

$$\mathsf{S}(x) = \sum_{v \in \mathcal{V}} \frac{1}{\|v\|^2} \, \langle v \, , \, x \rangle \, v \otimes v \tag{7.17}$$

defines a complex doubly-symmetric 3-tensor if $\mathcal{V}$ is any family of (non-vanishing) orthogonal vectors.

Second step: now given a complex doubly-symmetric 3-tensor S of the form (7.17), we shall prove that the set $\mathcal{V}$ coincides with the set

$$\widehat{\mathcal{V}} = \{ v \in \mathbb{C}^N \setminus \{0\} \, ; \; \mathsf{S}(v) = v \otimes v \} \, .$$

Clearly, if $y \in \mathcal{V}$ we have by (7.17)

$$\mathsf{S}(y) = y \otimes y \, .$$

This proves that $\mathcal{V} \subset \widehat{\mathcal{V}}$. Now, let $v \in \widehat{\mathcal{V}}$. On one side we have

$$\mathsf{S}(v) = v \otimes v \, ,$$

on the other side we have

$$\mathsf{S}(v) = \sum_{z \in \mathcal{V}} \frac{1}{\|z\|^2} \, \langle z \, , \, v \rangle \, z \otimes z \, .$$

In particular, applying $\langle y | \in \mathcal{V}^*$ to both sides, we get

$$\langle y \, , \, v \rangle \, v = \langle y \, , \, v \rangle \, y$$

and thus either v is orthogonal to y or $v = y$. This proves that v is one of the elements y of $\mathcal{V}$, for it were orthogonal to all the $y \in \mathcal{V}$ we would get $v \otimes v = \mathsf{S}(v) = 0$ and v would be the null vector.

We have proved that $\mathcal{V}$ coincides with the set

$$\{v \in \mathbb{R}^N \setminus \{0\}\,;\ \mathsf{S}(v) = v \otimes v\}\,.$$

Third step: now we shall prove that all complex doubly-symmetric 3-tensors S on $\mathbb{R}^{N+1}$ are diagonalizable in some orthonormal basis.

The property (7.12) indicates that the matrices

$$\mathsf{S}_k = (\mathsf{S}_k^{ij})_{i,j=0,\ldots,N}$$

are real symmetric hence they are all diagonalizable in some orthonormal basis. The properties (7.12) and (7.13) imply

$$\sum_{m=0}^{N} \mathsf{S}_j^{im}\,\mathsf{S}_l^{mk} = \sum_{m=0}^{N} \mathsf{S}_l^{im}\,\mathsf{S}_j^{mk}\,.$$

In other words

$$(\mathsf{S}_j\,\mathsf{S}_l)_{ik} = (\mathsf{S}_l\,\mathsf{S}_j)_{ik}\,,$$

for all i, k, all j, l. The matrices S_k commute pairwise. Thus the matrices S_k can be simultaneously diagonalized: there exists an orthogonal matrix $\mathsf{U} = (u^{ij})_{i,j=0,\cdots,N}$ such that, for all k in $\{0, \cdots, N\}$,

$$\mathsf{S}_k = \mathsf{U}\,\mathsf{D}_k\,\mathsf{U}^*\,, \tag{7.18}$$

where the matrices D_k are diagonal: $\mathsf{D}_k = diag(\lambda_k^0, \cdots, \lambda_k^N)$. As a consequence, the coefficient S_k^{ij} can be written as

$$\mathsf{S}_k^{ij} = \sum_{m=0}^{N} \lambda_k^m\, u^{im}\, u^{jm}\,.$$

Let us denote by a_m the mth column vector of U, that is, $a_m = (u^{lm})_{l=0,\cdots,N}$. Moreover, we denote by λ^m the vector of λ_k^m, for $k = 0, \cdots, N$. Since the matrix U is orthogonal, the vectors a_m form an orthonormal basis of $\mathbb{R}^{N+1}$. We have

$$\mathsf{S}_k^{ij} = \sum_{m=0}^{N} \lambda_k^m\, a_m^i\, a_m^j\,.$$

Our aim now is to prove that λ^m is proportional to a_m. To this end, we shall use the symmetry properties of S. From the simultaneous reduction (7.18), we get

$$\mathsf{S}_j\,\mathsf{S}_q = \mathsf{U}\,\mathsf{D}_j\,\mathsf{D}_q\,\mathsf{U}^*\,.$$

Therefore, we have

$$(\mathsf{S}_j\,\mathsf{S}_q)_{i,r} = \sum_{m=0}^{N} \mathsf{S}_j^{im}\,\mathsf{S}_q^{mr} = \sum_{m=0}^{N} a_m^i\,\lambda_j^m\,\lambda_q^m\,a_m^r\,.$$

In particular we have, for all $p \in \{0,\dots,N\}$

$$\begin{aligned}\sum_{i,j,q,r=0}^{N} (\mathsf{S}_j\,\mathsf{S}_q)_{i,r}\,a_p^i\,\lambda_j^p\,\lambda_q^p\,a_p^r &= \sum_{m=0}^{N} \langle a_m\,,\,a_p\rangle\;\langle\lambda^m\,,\,\lambda^p\rangle\;\langle\lambda^p\,,\,\lambda^m\rangle\;\langle a_p\,,\,a_m\rangle\\ &= \|\lambda^p\|^4\,. \qquad (7.19)\end{aligned}$$

Note that

$$\sum_{m=0}^{N} \mathsf{S}_j^{im}\mathsf{S}_q^{mr}$$

is also symmetric in (j,r). Applying this, the expression (7.19) is also equal to

$$\begin{aligned}&\sum_{i,j,q,r=0}^{N}\sum_{m=0}^{N} a_m^i\,\lambda_r^m\,\lambda_q^m\,a_m^j\,a_p^i\,\lambda_j^p\,\lambda_q^p\,a_p^r =\\ &\qquad= \sum_{m=0}^{N} \langle a_p\,,\,\lambda^m\rangle\;\langle\lambda^m\,,\,\lambda^p\rangle\;\langle a_m\,,\,\lambda^p\rangle\;\langle a_p\,,\,a_m\rangle\\ &\qquad= |\langle a_p\,,\,\lambda^p\rangle|^2\;\|\lambda^p\|^2\,.\end{aligned}$$

This gives

$$|\langle a_p\,,\,\lambda^p\rangle| = \|\lambda^p\| = \|a_p\|\;\|\lambda^p\|\,.$$

This is a case of equality in Cauchy-Schwarz inequality, hence there exists $\mu_p \in \mathbb{R}$ such that $\lambda^p = \mu_p\,a_p$, for all $p = 0,\dots,N$. This way, the 3-tensor S can be written as

$$\mathsf{S}_k^{ij} = \sum_{m=0}^{N} \mu_m\,a_m^i\,a_m^j\,a_m^k\,. \qquad (7.20)$$

In other words

$$\mathsf{S}(x) = \sum_{m=0}^{N} \mu_m\;\langle a_m\,,\,x\rangle\;a_m\otimes a_m\,.$$

We have obtained the orthonormal diagonalization of S. The proof is complete. $\square$

7.3.7 Back to Obtuse Random Variables

The theorem above is a general diagonalization theorem for 3-tensors. For the moment it does not take into account the relation (7.11). When we make it enter into the game, we see obtuse systems appearing.

Theorem 7.49. *Let* S *be a doubly-symmetric 3-tensor on* $\mathbb{R}^{N+1}$ *satisfying also the relation*

$$\mathsf{S}_0^{ij} = \delta_{ij}$$

for all $i, j = 0, \ldots, N$*. Then the orthogonal system* $\mathcal{V}$ *such that*

$$\mathsf{S}(x) = \sum_{v \in \mathcal{V}} \frac{1}{\|v\|^2} \langle v\,,\, x \rangle \; v \otimes v \tag{7.21}$$

is made of exactly $N+1$ *vectors* $v_1, \ldots, v_{N+1}$*, all of them satisfying* $v_i^0 = 1$*. In particular the family of* $N+1$ *vectors of* $\mathbb{R}^N$*, obtained by restricting the* v_i*'s to their* N *last coordinates, forms an obtuse system in* $\mathbb{R}^N$*.*

Proof. First assume that $\mathcal{V} = \{v_1, \ldots, v_K\}$. By hypothesis, we have

$$\mathsf{S}_k^{ij} = \sum_{m=1}^{K} \frac{1}{\|v_m\|^2} \, v_m^i \, v_m^j \, v_m^k \,,$$

for all $i, j, k = 0, \ldots, N$. With the supplementary property (7.11) we have in particular

$$\mathsf{S}_0^{ij} = \sum_{m=1}^{K} \frac{1}{\|v_m\|^2} \, v_m^i \, v_m^j \, v_m^0 = \delta_{ij}$$

for all $i, j = 0, \ldots, N$.

Consider the orthonormal family of $\mathbb{R}^{N+1}$ made of the vectors $e_m = v_m / \|v_m\|$. We have obtained above the relation

$$\sum_{m=0}^{K} v_m^0 \, |e_m\rangle\langle e_m| = \mathrm{I}\,,$$

as matrices acting on $\mathbb{R}^{N+1}$. The above is thus a spectral decomposition of the identity matrix, this implies that the e_m's are exactly $N+1$ vectors and that all the v_m^0 are equal to 1.

This proves the first part of the theorem. The last part concerning obtuse systems is now obvious and was already noticed when we introduced obtuse systems. □

In particular we have proved the following theorem.

Theorem 7.50. *The set of doubly-symmetric 3-tensors* S *on* $\mathbb{R}^{N+1}$ *which satisfy also the relation*

$$\mathsf{S}_0^{ij} = \delta_{ij}$$

for all $i, j = 0, \ldots, N$, *is in bijection with the set of distributions of obtuse random variables* X *on* $\mathbb{R}^N$. *The bijection is described by the following, with the convention* $X^0 = \mathbb{1}$.

– *The random variable* X *is the only (in distribution) random variable satisfying*

$$X^i X^j = \sum_{k=0}^{N} \mathsf{S}_k^{ij} X^k \,,$$

for all $i, j = 1, \ldots, N$.

– *The 3-tensor* S *is obtained by*

$$\mathsf{S}_k^{ij} = \mathbb{E}[X^i \, X^j \, X^k] \,,$$

for all $i, j, k = 0, \ldots, N$.

In particular the different possible values taken by X *in* $\mathbb{R}^N$ *coincide with the vectors* $w_n \in \mathbb{R}^N$, *made of the last* N *coordinates of the eigenvectors* v_n *associated to* S *in the representation (7.21). The associated probabilities are then* $p_n = 1/(1 + \|w_n\|^2) = 1/\|v_n\|^2$.

7.3.8 Probabilistic Interpretations

Definition 7.51. Let X be an obtuse random variable in $\mathbb{R}^N$, with associated 3-tensor S and let $(\Omega, \mathcal{F}, \mathbb{P}_\mathsf{S})$ be the canonical space of X. Note that we have added the dependency on S for the probability measure $\mathbb{P}_\mathsf{S}$. The reason is that, when changing the obtuse random variable X on $\mathbb{R}^N$, the canonical space Ω and the canonical σ-field $\mathcal{F}$ do not change, only the canonical measure $\mathbb{P}$ does change.

We have seen that the space $L^2(\Omega, \mathcal{F}, \mathbb{P}_\mathsf{S})$ is a $N+1$-dimensional Hilbert space and that the family $\{X^0, X^1, \ldots, X^N\}$ is an orthonormal basis of that space. Hence for every obtuse random variable X, with associated 3-tensor S, we have a natural unitary operator

$$\begin{array}{rcl} \mathsf{U}_\mathsf{S} : L^2(\Omega, \mathcal{F}, \mathbb{P}_\mathsf{S}) & \longrightarrow & \mathbb{C}^{N+1} \\ X^i & \longmapsto & e_i \,, \end{array}$$

where $\{e_0, \ldots, e_N\}$ is the canonical orthonormal basis of $\mathbb{C}^{N+1}$. The operator U_S is called the *canonical isomorphism* associated to X.

Definition 7.52. On the space $L^2(\Omega, \mathcal{F}, \mathbb{P}_\mathsf{S})$, for each $i = 0, \ldots, N$, we consider the multiplication operator

$$\begin{array}{rcl}\mathcal{M}_{X^i} : L^2(\Omega, \mathcal{F}, \mathbb{P}_\mathsf{S}) & \longrightarrow & L^2(\Omega, \mathcal{F}, \mathbb{P}_\mathsf{S}) \\ Y & \longmapsto & X^i\, Y\,, \end{array}$$

Definition 7.53. On the space $\mathbb{C}^{N+1}$, with canonical basis $\{e_0, \dots, e_N\}$ we consider the basic matrices a^i_j, for $i, j = 0, \dots, N$ defined by

$$a^i_j\, e_k = \delta_{i,k}\, e_j\,.$$

We shall see now that, when carried out on the same canonical space by U_S, the obtuse random variables of $\mathbb{R}^N$ admit a simple and compact matrix representation in terms of their 3-tensor.

Theorem 7.54. *Let X be an obtuse random variable on $\mathbb{R}^N$, with associated 3-tensor S and canonical isomorphism U_S. Then we have, for all $i, j = 0, \dots, N$*

$$\mathsf{U}_\mathsf{S}\, \mathcal{M}_{X^i}\, \mathsf{U}_\mathsf{S}^* = a^0_i + a^i_0 + \sum_{j,k=1}^N S^{jk}_i\, a^j_k\,. \tag{7.22}$$

for all $i = 1, \dots, N$.

Proof. We have, for any fixed $i \in \{1, \dots, N\}$, for all $j = 0, \dots, N$

$$\begin{aligned}\mathsf{U}_\mathsf{S}\, \mathcal{M}_{X^i}\, \mathsf{U}_\mathsf{S}^*\, e_j &= \mathsf{U}_\mathsf{S}\, \mathcal{M}_{X^i}\, X^j \\ &= \mathsf{U}_\mathsf{S}\, X^i\, X^j \\ &= \mathsf{U}_\mathsf{S} \sum_{k=0}^N \mathsf{S}^{ij}_k X^k \\ &= \sum_{k=0}^N \mathsf{S}^{ij}_k\, e_k\,.\end{aligned}$$

Hence the operator $\mathsf{U}_\mathsf{S}\, \mathcal{M}_{X^i}\, \mathsf{U}_\mathsf{S}^*$ has the same action on the orthonormal basis $\{e_0, \dots, e_N\}$ as the operator

$$\sum_{j,k=0}^N \mathsf{S}^{ij}_k\, a^j_k\,.$$

Using the symmetries of S (Proposition 7.44) and the identity (7.11), we get

$$\begin{aligned}\mathsf{U}_\mathsf{S}\, \mathcal{M}_{X^i}\, \mathsf{U}_\mathsf{S}^* &= \sum_{k=0}^N \mathsf{S}^{i0}_k\, a^0_k + \sum_{j=1}^N \mathsf{S}^{ij}_0\, a^j_0 + \sum_{j,k=1}^N \mathsf{S}^{ij}_k\, a^j_k \\ &= a^i_0 + a^0_i + \sum_{j,k=1}^N \mathsf{S}^{jk}_i\, a^j_k\,.\end{aligned}$$

This proves the representation (7.22). $\square$

Once again this is a remarkable point that we get here. All these different obtuse random variables X of $\mathbb{R}^N$ (and by the way, all the random variables in $\mathbb{R}^N$ with finite support) can be represented as very simple linear combinations of the basic operators a^i_j. As we have discussed above, all the probabilistic properties of X (law, independence, functional calculus etc.) are carried by these linear combinations of basic operators.

Let us check how this works with our examples. We first start with the obtuse random variable X in $\mathbb{R}^2$ taking the values

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\ , \qquad v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}\ , \qquad v_3 = \begin{pmatrix} -1 \\ -2 \end{pmatrix},$$

with respective probabilities

$$p_1 = \frac{1}{2}\ , \qquad p_2 = \frac{1}{3}\ , \qquad p_3 = \frac{1}{6}\,.$$

Adding a third coordinate 1 to each vector, we get the following orthogonal system of $\mathbb{R}^3$:

$$\widehat{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\ , \qquad \widehat{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}\ , \qquad \widehat{v}_3 = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}\ .$$

We compute the matrices $|\widehat{v}_i\rangle\langle\widehat{v}_i|$ and get respectively

$$|\widehat{v}_1\rangle\langle\widehat{v}_1| = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$|\widehat{v}_2\rangle\langle\widehat{v}_2| = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}$$

$$|\widehat{v}_3\rangle\langle\widehat{v}_3| = \begin{pmatrix} 1 & -1 & -2 \\ -1 & 1 & 2 \\ -2 & 2 & 4 \end{pmatrix}\ .$$

The matrices S_k, of multiplication by X^k, are given by

$$\mathsf{S}_k^{ij} = \sum_{m=0^2} p_m v^i_m v^i_j v^i_k\,.$$

This gives

$$\mathsf{S}_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$\mathsf{S}_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$
$$\mathsf{S}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & -1 \\ 1 & -1 & -1 \end{pmatrix}.$$

These are the matrices associated to the multiplication operators by X^0, X^1 and X^2, respectively. We recover facts that can be easily checked :

$$\mathsf{S}_0 = \mathrm{I}, \qquad (X^1)^2 = X^0, \qquad X^1X^2 = -X^2, \qquad (X^2)^2 = X^0 - X^1 - X^2 .$$

As a consequence, the two operators

$$\mathsf{X}^1 = a_0^1 + a_1^0 - a_2^2, \qquad \mathsf{X}^2 = a_0^2 + a_2^0 - a_2^1 - a_1^2 - a_2^2$$

have exactly the same probabilistic properties, in the sense of Quantum Probability on $\mathbb{C}^3$ in the state $|e_0\rangle\langle e_0|$, as the pair of random variables (X^1, X^2).

Let us now compute the 3-tensor associated to our second example, that is the obtuse random X on $\mathbb{R}^2$ which takes the values

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} -1 \\ \left(\frac{1-2h}{h}\right)^{1/2} \end{pmatrix} \qquad v_3 = \begin{pmatrix} -1 \\ -2\left(\frac{h}{1-2h}\right)^{1/2} \end{pmatrix}$$

with respective probabilities

$$p_1 = \frac{1}{2}, \qquad p_2 = h, \qquad p_3 = \frac{1}{2} - h .$$

The same kind of computation as above gives

$$\mathsf{S}_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$\mathsf{S}_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$
$$\mathsf{S}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & -1 \\ 1 & -1 & \frac{1-4h}{\sqrt{h(1-2h)}} \end{pmatrix}.$$

At the level of this lecture we do not have the tools for computing rigorously the continuous-time limits of the random walks associated to the two examples above, but we shall end this lecture by just describing these continuous-time limits. In both cases consider the stochastic process

$$Y_{nh} = \sum_{k=0}^{n} \sqrt{h}\, X_{kh} \,,$$

where the X_{kh} are independent copies of the random variable X.

In the case of the first example, the process Y converges, when h tends to 0, to a two dimensional Brownian motion.

In the case of our second example, the limit process is a Brownian motion in the first coordinate and a standard compensated poisson process in the second coordinate.

Notes

Many parts of this chapter are inspired from the two reference books of Parthasarathy [Par92] and Meyer [Mey93], and also from the course by Biane [Bia95]. The two first references contain long developments concerning the theory of quantum probability. Our approach in this chapter is closer to the one of Parthasarathy.

The notion of Toy Fock space and its probabilistic interpretation as natural space for all sequences of Bernoulli random variables, is due to P.-A. Meyer. It first appeared in [Mey86]. The way we present it in this chapter is an extended version, developed by Attal and Pautrat in [AP06].

Real obtuse random variables and their diagonalization theorem were introduced by Attal and Emery in [AÉ94].

References

[AÉ94] S. Attal and M. Émery. Équations de structure pour des martingales vectorielles. In *Séminaire de Probabilités, XXVIII*, volume 1583 of *Lecture Notes in Math.*, pages 256–278. Springer, Berlin, 1994.

[AP06] Stéphane Attal and Yan Pautrat. From repeated to continuous quantum interactions. *Ann. Henri Poincaré*, 7(1):59–104, 2006.

[Bia95] Philippe Biane. Calcul stochastique non-commutatif. In *Lectures on probability theory (Saint-Flour, 1993)*, volume 1608 of *Lecture Notes in Math.*, pages 1–96. Springer, Berlin, 1995.

[Mey86] P.-A. Meyer. Éléments de probabilités quantiques. I–V. In *Séminaire de Probabilités, XX, 1984/85*, volume 1204 of *Lecture Notes in Math.*, pages 186–312. Springer, Berlin, 1986.

[Mey93] Paul-André Meyer. *Quantum probability for probabilists*, volume 1538 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1993.

[Par92] K. R. Parthasarathy. *An introduction to quantum stochastic calculus*, volume 85 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1992.