Sheet 1

Basic bibliography

- Meyer, Paul A. *Quantum Probability for Probabilists*. 2nd edition. Berlin; New York: Springer, 1995.
- Parthasarathy, K. R. *An Introduction to Quantum Stochastic Calculus*. Springer Science & Business Media, 1992.
- Biane, Philippe. "Calcul Stochastique Non-Commutatif." In *Lectures on Probability Theory*, edited by Pierre Bernard, 1–96. Lecture Notes in Mathematics 1608. Springer Berlin Heidelberg, 1995. https://doi.org/10.1007/BFb0095746.

1 What is quantum probability?

QP is a non-commutative generalisation of probability.

classical probability	quantum probability
$L^2(\Omega,\mathcal{F},\mathbb{P})$	Hilbert space $\mathcal{H}$
$1\in L^2(\Omega,\mathcal{F},\mathbb{P})$	distinguished vector $u\in\mathcal{H}$, or a state ρ
random variable X	self-adjoint operator A
law of X	$f\mapsto \mathrm{Tr}[\rho f(A)]$,
expectation of X	$\mathrm{Tr}[\rho A]$

We start with a finite probability space $|\Omega|<+\infty$.

A real random variable X can be seen as a self-adjoint operator O_X on $\mathcal{H}=L^2(\Omega,\mathcal{F},\mathbb{P})$, diagonal on the o.n.b. $(e_\omega=\mathbb{I}_\omega/\mathbb{P}(\omega))_{\omega\in\Omega}$ and such that $O_X e_\omega = X(\omega)e_\omega$. The average of X is given by $\mathbb{E}[X]=\langle 1, O_X 1\rangle$ where $1\in\mathcal{H}$ is the function $1(\omega)=1$ for all $\omega\in\Omega$.

On the other hand given a n–dimensional Hilbert space $\mathcal{H}$ with a distinguished vector $u\in\mathcal{H}$ and a self-ajoint operator A there exists an ONB family of eigenvectors $(\varphi_i)_{i=1,\dots,n}$ for A corresponding to eigenvalues $(\lambda_i)_i$. Then letting $\Omega=\{1,\dots,n\}$, $\mu(\{i\})=\langle\varphi_i,u\rangle$, $X(i)=\lambda_i$ we have

$$A=\sum_i \lambda_i |\varphi_i\rangle\langle\varphi_i|, \qquad \mathbb{E}[X]=\sum_i \lambda_i\mu(\{i\})=\sum_i \lambda_i\langle u|\varphi_i\rangle\langle\varphi_i|\,u\rangle=\langle u, Au\rangle$$

Given two vectors $v,w\in\mathcal{H}$ the notation $|v\rangle\langle w|$ stands for the operator such that $|v\rangle\langle w|\,\varphi=\langle w,\varphi\rangle v$ for any $\varphi\in\mathcal{H}$. More generally, for any bounded function f

$$\mathbb{E}[f(X)]=\sum_i f(\lambda_i)\mu(\{i\})=\langle u, f(A)\,u\rangle$$

where the operator $f(A)$ is defined by $f(A)\varphi_i=f(\lambda_i)\varphi_i$ for all $i=1,\dots,n$.

General states corresponds to convex combination of the previous construction:

$$\sum_i p_i\langle u_i, f(A)\,u_i\rangle=\mathrm{Tr}[\rho f(A)] \qquad \text{with} \qquad \rho=\sum_i p_i|u_i\rangle\langle u_i|$$

where $\|u_i\| = 1$ and $\sum_i p_i = 1$. The operator ρ is positive definite $\langle u, \rho u\rangle \geqslant 0$ therefore self–adjoint and also of trace 1. We call all such operators *states* or *density matrices* and denote their set as $\mathcal{S}(\mathcal{H}) = \{\rho \in \mathcal{L}(\mathcal{H})\colon \rho \geqslant 0, \mathrm{Tr}(\rho) = 1\}$. Pure states $\rho_u = |u\rangle\langle u|$ are extremals in the convext set of all states. Note that pure states are defined modulo a conjugation with a phase factor $\rho_{e^{i\varphi}u} = |e^{i\varphi}u\rangle\langle e^{i\varphi}u| = |u\rangle\langle u| = \rho_u$. For pure states we have $\mathrm{Tr}(\rho_u A) = \langle u, Au\rangle$.

The expectation of a self-adjoint element on the state ρ is defined as the map $A \mapsto \mathrm{Tr}(\rho A)$. The law μ of the random variable A is given by the measure μ on $\sigma(A)$ (the spectrum of A) such that

$$\mu(f) = \mathrm{Tr}[\rho f(A)],$$

for all bounded measurable functions on $\sigma(A)$. In the discrete setting $\sigma(A) = \{\lambda_i\}_i$.

If A, B commute then their joint law is the measure μ on $\sigma(A) \times \sigma(B)$ such that $\mu(f) = \mathrm{Tr}[\rho f(A, B)]$ for all (bounded, measurable) f: $\sigma(A) \times \sigma(B) \to \mathbb{R}$. In this case $f(A, B)$ is defined by $f(A, B)\varphi_i = f(\lambda_i(A), \lambda_i(B))\varphi_i$ where $(\varphi_i)_i$ is a o.n.b. of $\mathcal{H}$ made of joint eigenvectors of A and B and where $\lambda_i(A), \lambda_i(B)$ are the corresponding eigenvalues.

The notion of law for multiple of random variables is well defined only if the random variables are commuting. The basic problem is that there is no canonical way to define $f(A, B)$ for non-commuting operators. For example, if $f(x, y) = x^2y^2$ then we have also $f(x, y) = x\,y\,x\,y$ but in general

$$A^2B^2 \neq ABAB.$$

A quantum probability space is an Hilbert space $\mathcal{H}$ together with a state $\rho \in \mathcal{S}(\mathcal{H})$. Random variables are the self–adjoint elements of $\mathcal{L}(\mathcal{H})$.

The aim of QP is to compute and study the laws of given collections of self–adjoint operators.

Later on more details on axiomatization. This is the von Neumann model of QP.

2 Non-commutative Bernoulli space

The simplest example is given by an Hilbert space of dimension 2:

$$\mathcal{H} = \mathrm{span}_{\mathbb{C}}(|0\rangle, |1\rangle)$$

or $|+1\rangle, |-1\rangle$, or $|\mathrm{fundamental}\rangle, |\mathrm{excited}\rangle$, or $|\uparrow\rangle, |\downarrow\rangle$, etc..

Any state has the form

$$\rho = \begin{pmatrix} \alpha & r \\ \bar{r} & 1-\alpha \end{pmatrix}, \qquad \alpha \in [0,1], r \in \mathbb{C}, \alpha(1-\alpha) - |r|^2 \geqslant 0.$$

When $\alpha(1-\alpha) = |r|^2$ this is an orthogonal projector on a complex line giving a pure state.

Consider the basis on the space of 2×2 matrices (Pauli matrices)

$$\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Another common notation for $\sigma_x, \sigma_y, \sigma_z$ is $\sigma_1, \sigma_2, \sigma_3$. This is an orthonormal basis of $M_2(\mathbb{C})$ with the scalar product $\langle A, B\rangle = \frac{1}{2}\mathrm{Tr}(A^*B)$. They verify the commutation relations

$$[\sigma_x, \sigma_y] = 2i\sigma_z, \qquad [\sigma_z, \sigma_x] = 2i\,\sigma_y, \qquad [\sigma_y, \sigma_z] = 2i\,\sigma_x,$$

anticommutation relations

$$\{\sigma_a, \sigma_b\} = 2\delta_{a,b}, \qquad a, b \in \{x, y, z\},$$

and they are involutions

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = -i\sigma_x\sigma_y\sigma_z = \mathbb{I}.$$

A generic self–adjoint element is given by

$$A = x\sigma_x + y\sigma_y + z\sigma_z + t\mathbb{I}$$

for which

$$\mathrm{Tr}[A] = 2t, \qquad \mathrm{Tr}[A^2] = 2(x^2 + y^2 + z^2 + t^2),$$

this implies that the spectrum of A is $\sigma(A) = \{t \pm \sqrt{x^2 + y^2 + z^2}\}$. Therefore the matrix

$$\sigma_{\theta,\varphi} = \begin{pmatrix} \cos\theta & e^{-i\varphi}\sin\theta \\ e^{i\varphi}\sin\theta & -\cos\theta \end{pmatrix}$$

represents a generic Bernoulli random variable with spectrum $\{-1, +1\}$. On the other hand all the pure states can be written as

$$\rho_{\theta,\varphi} = \frac{1}{2}(\mathbb{I} + \sigma_{\theta,\varphi}) = |\theta, \varphi\rangle\langle\theta, \varphi|$$

where $|\theta, \varphi\rangle$ is the eigenvector of $\sigma_{\theta,\varphi}$ with eigenvalue $+1$, namely $\sigma_{\theta,\varphi}|\theta, \varphi\rangle = |\theta, \varphi\rangle$.

3 The Stern–Gerlach experiment

Otto Stern and Walther Gerlach conducted in Frankfurt in 1922 the experience described in Figure 1 (left). A beam of atoms experience an intense magnetic field and as a consequence is deflected. Upon detection by means of a screen the arrival positions of the atoms reveals a quantized patterns, in contrast with classical theory of the magnetic moment of atoms which would require a continuous distribution of arrival positions due to the uniform distributions of the magnetic moment within the atom's population escaping from the oven. Figure 1 (right) shows the actual images obtained in the original experiment.

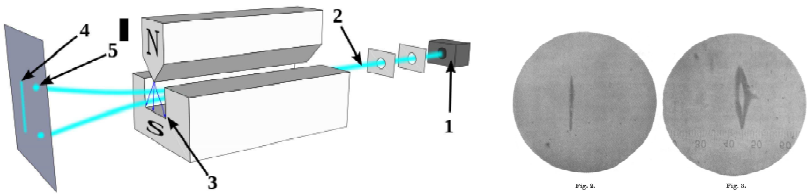
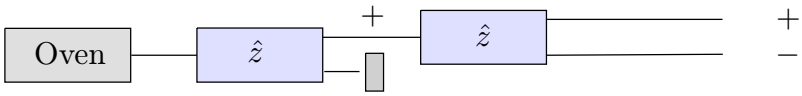


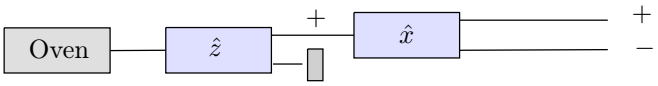
Figure 1. Left: Stern–Gerlach experiment: silver atoms travel through an inhomogeneous magnetic field and are deflected up or down depending on their spin. 1: furnace. 2: beam of silver atoms. 3: inhomogeneous magnetic field. 4: expected result. 5: what was actually observed. [from Wikipedia https://en.wikipedia.org/wiki/Stern–Gerlach_experiment]. Right: the experimental result of the Stern-Gerlach experiment. The beam has split into two components. From [Gerlach, Walther, and Otto Stern. "Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld." *Zeitschrift für Physik* 9, no. 1 (December 1, 1922): 349–52. https://doi.org/10.1007/BF01326983.]

The Stern–Gerlach experiment is a proof of the existence of a quantized magnetic moment for the electron. Indeed the silver atoms have atomic number 47. In its fundamental state, 46 of these electrons do not contributed to the magnetic moment since they come in pairs of opposite intrinstic magnetic moment (*spin*) and in a spatially symmetric state which do not generate any angular momentum. Only the last electron, whose spatial distributions is also symmetric, has an uncompensated intrinsic magnetic moment which consitute the only relevant contribution to the total magnetic moment of the atom. This magnetic moment interacts with the non–uniform magnetic field deflecting the trajectory of the atom. The presence of two well separated tracks means that this spin comes only in two varieties, oriented in the direction of the magnetic field or in the opposite direction.

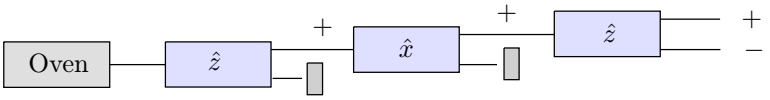
So the spin of the electron is a Bernoulli random variable. In order to explore other properties of this random variable we imagine a sequence of Stern–Gerlach experiments performed in series.



In this first case we first measure the $\hat{z}$ orientation, select those atoms which emerge from the + path after the first instrument and then again the $\hat{z}$ orientation and we obtain that all the atoms emerge from the + path.

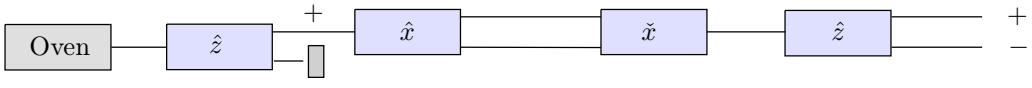


In this second situation we measure a different, orthogonal direction in the second instrument and we obtain that half of the atoms emerge from the + path and half from the − path. This is expected due to the symmetry of the problem.

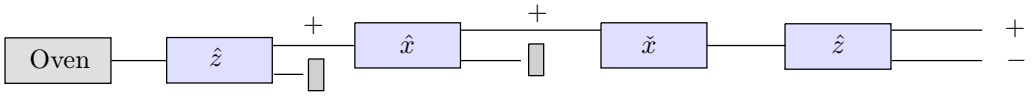


In this third installment we select the atoms which emerge from the + path after the $\hat{x}$ instrument and perform another selection with a $\hat{z}$ instrument. The result is that again half of the atoms emerge from the + path and half from the − path. The interpretation is that the measurement of $\hat{x}$ has completely destroyed the previous measurement of $\hat{z}$.

We now introduce another apparatus which undo the effect of a Stern–Gerlach instrument, this is not difficult to imagine, we just need to produce the opposite magnetic field to undo the effect of the first and arrange appropriately the geometry to recombine the atom beam. We label this instrument $\check{z}$ if it operates in the z direction.



In this first case we use the new instrument to recombine the beams after a $\hat{x}$ beam splitter. If we have selected only atoms with spin in the $\hat{z} = +1$ direction right after the oven, then we will end up with all the atoms in the + beam after the last $\hat{z}$ instrument.



We now block the $\hat{x} = -1$ beam and we observe that atoms exit the instrument with probability $1/2$ in each of the two final beams.

This is quite surprising. Allowing *more* atoms to go through the experiment depletes one of the exit beams! This property is not in agreement with a probabilistic description of the state of the atoms. Removing a conditioning cannot renders impossible events which were possible under the conditioning. This is a manifestation of quantum mechanical interference effects.

Let us use the quantum Bernoulli space to compute the probability that an atom which has been measured in the $\hat{z}=+1$ direction will be, subsequently, measured in the $\hat{v}=+1$ direction where $\hat{v}\in\mathbb{R}^3$ is a unit vector which has an angle θ wrt. $\hat{z}$. We can take σ_z as the r.v. which represents the value of the first measurement and we denote by $|\pm_z\rangle$ the corresponding eigenvector with eigenvalue ±1. After the measurement the state of the system is the pure state $|+_z\rangle\langle+_z|$ since it is the only that satisfies $\mathrm{Tr}(\rho\mathbb{I}_{\sigma_z=+1})=1$. The spin $\sigma(v)$ in the direction $\hat{v}$ has to be represented by a self–adjoint operator with spectrum ±1, therefore

$$\sigma(\hat{v})=c_x\sigma_x+c_y\sigma_y+c_z\sigma_z,$$

with $c_x^2+c_y^2+c_z^2=1$. We need to identify these coefficients. Using a difference reference frame obtained via rotation in $\mathbb{R}^3$ does not affect the state of the atoms but random variables corresponding to measurament in a given direction are transformed into other variables in a different direction. In particular there should exists a unitary operator U which conjugates σ_x, σ_y, σ_z with $\sigma(f_1)$, $\sigma(f_2)$, $\sigma(f_3)$ where $(f_1,\ f_2,\ f_3)$ is any other orthornormal basis of $\mathbb{R}^3$ with positive orientation. We see that rotations $R\in\mathrm{SO}(3)$ have to act as unitary operators U_R on $\mathcal{H}$ such that $U_R^{-1}\sigma(v)U_R=\sigma(Rv)$. As a consequence we observe that the bilinear form $B(\hat{v},\hat{w})=\frac{1}{2}\mathrm{Tr}[\sigma(\hat{v})\sigma(\hat{w})]$ is invariant under rotations and coincide with the standard scalar product of $\hat{v}$ and $\hat{w}$. Therefore we must have

$$\sigma(\hat{v})=v_x\sigma_x+v_y\sigma_y+v_z\sigma_z$$

where $\hat{v}_x,\hat{v}_y,\hat{v}_z$ are the components of $\hat{v}\in\mathbb{R}^3$ in the basis $(\hat{x},\hat{y},\hat{z})$. At this point we are able to carry on our computation

$$\mathrm{Tr}(\rho\mathbb{I}_{\sigma(\hat{v})=+1})=\frac{1}{2}\mathrm{Tr}(\rho\left(\mathbb{I}+\sigma(\hat{v})\right))=\frac{1}{2}(1+\langle+_z|\sigma(\hat{v})|+_z\rangle)=\frac{1}{2}(1+\hat{v}_z)$$

where we used the fact that $\mathbb{I}_{\sigma(\hat{v})=+1}=\left(\mathbb{I}+\sigma(\hat{v})\right)/2$ (by functional calculus, or simply checking that the two operators have the same eigenvalues) and $\langle+_z|\sigma_{x,y}|+_z\rangle=0$ by symmetry. So the probability to observe an atom prepared in the $\sigma(\hat{z})=+1$ state in the $\sigma(\hat{v})=+1$ state is

$$\frac{1}{2}(1+\cos(\theta))$$

where θ is the angle between $\hat{z}$ and $\hat{v}$.

We note that the Pauli matrices are generators for the unitary groups corresponding to rotations around the x,y,z axis respectively. Indeed consider the unitary operator $U_\alpha=\exp(i\sigma_z\alpha/2)$ and let

$$\sigma_x(\alpha)=U_\alpha^{-1}\sigma_xU_\alpha,\quad \sigma_y(\alpha)=U_\alpha^{-1}\sigma_xU_\alpha,\quad \sigma_z(\alpha)=U_\alpha^{-1}\sigma_zU_\alpha=\sigma_z$$

for $\alpha\in\mathbb{R}$. We have $(\sigma_x(0),\sigma_y(0),\sigma_z(0))=(\sigma_x,\sigma_y,\sigma_z)$ and

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\sigma_x(\alpha)=\frac{i}{2}U_\alpha^{-1}(\sigma_x\sigma_z-\sigma_z\sigma_x)U_\alpha=\frac{i}{2}U_\alpha^{-1}[\sigma_x,\sigma_z]U_\alpha=U_\alpha^{-1}\sigma_yU_\alpha=\sigma_y(\alpha)$$

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\sigma_y(\alpha)=\frac{i}{2}U_\alpha^{-1}(\sigma_y\sigma_z-\sigma_z\sigma_y)U_\alpha=\frac{i}{2}U_\alpha^{-1}[\sigma_y,\sigma_z]U_\alpha=-U_\alpha^{-1}\sigma_xU_\alpha=-\sigma_x(\alpha)$$

therefore we must have

$$(\sigma_x(\alpha),\sigma_y(\alpha),\sigma_z(\alpha))=(\cos(\alpha)\sigma_x+\sin(\alpha)\sigma_y,-\sin(\alpha)\sigma_x+\cos(\alpha)\sigma_y,\sigma_z)=R_\alpha(\sigma_x,\sigma_y,\sigma_z)$$

where $R_\alpha \in \mathrm{SO}(3)$ is a rotation of α around $\hat{z}$. For all $v \in \mathbb{R}^3$ we have

$$U_\alpha^{-1}\sigma(v)U_\alpha = U_\alpha^{-1}(v_x\sigma_x + v_y\sigma_y + v_z\sigma_z)U_\alpha = (v_x, v_y, v_z)\cdot R_\alpha(\sigma_x, \sigma_y, \sigma_z)$$

$$= (R_\alpha^{-1}(v_x, v_y, v_z))\cdot(\sigma_x, \sigma_y, \sigma_z) = \sigma(R_\alpha^{-1}v).$$
