

Quantum Probability Theory

HANS MAASSEN

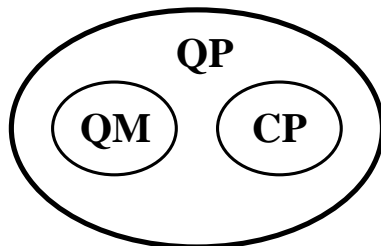
*Caputcollege Kanstheorie
subfaculteit wiskunde, KUN
najaar 1998*

Versie van 27-1-2003

1. The framework of Quantum Probability.

*‘Well! I’ve often seen a cat without a grin,’
thought Alice; ‘but a grin without a cat!’*

Quantum probability theory generalises Kolmogorov’s classical probability theory to make room for quantummechanical models. These models typically contain non-commuting observables. Therefore quantum probability is ‘non-commutative probability’.



1.1. Classical probability.

In Kolmogorov’s ‘Grundbegriffe der Wahrscheinlichkeitsrechnung’ of 1933 a *probability space* is defined as a triple $(\Omega, \Sigma, \mathbb{P})$, where Ω is a set, Σ is a σ -algebra of subsets of Ω , (i.e. a collection of subsets that is closed under the operations of taking unions and complements), containing Ω itself as an element, and \mathbb{P} is a function $\Sigma \rightarrow [0, 1]$ with the properties

(i) *normalisation:*

$$\mathbb{P}(\Omega) = 1 ;$$

(ii) *additivity:*

$$A \cap B = \emptyset \quad \implies \quad \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) ;$$

(iii) *continuity:*

$$A_1 \subset A_2 \subset A_3 \subset \dots \quad \implies \quad \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n) .$$

These objects are to be interpreted as follows. Ω is the set of all possible outcomes of a certain stochastic experiment. The collection Σ consists of ‘events’, statements about the outcome of the experiment that can be tested by observation. They are the ‘yes-no-questions’ that can be asked to the system. When A is such an event, then $\mathbb{P}(A)$ is the *probability* that A will occur (that the answer will be ‘yes’).

Now what does the number $\mathbb{P}(A)$ mean? We take the following point of view: If $\mathbb{P}(A)$ is very close to 1, we may be pretty sure that A will occur. If $\mathbb{P}(A) = 1$, we are certain that it will.

And what about probabilities that are far from 1 (or 0)? The following theorem in probability theory comes to our aid in providing a meaning for these values.

A family $(A_i)_{i \in I}$ of events is called *independent* if for all finite $J \subset I$ we have

$$\mathbb{P} \left(\bigcap_{i \in J} A_i \right) = \prod_{i \in J} \mathbb{P}(A_i) .$$

Theorem (Weak law of large numbers). *If A_1, A_2, A_3, \dots are independent, all having the same probability p , then we have for all $\varepsilon > 0$:*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\left| \frac{1}{n} K_n - p \right| < \varepsilon \right] = 1 .$$

where

$$K_n(\omega) := \# \{ j \in \{1, 2, \dots, n\} \mid \omega \in A_j \} .$$

Interpretation. We think of independent trials of the same experiment. Trial number j is considered a *success* if A_j occurs. The theorem means that we can be pretty sure that after sufficiently many trials the success rate is close to the probability p . This enables us to *measure* probabilities, much in the same way as one measures other physical quantities: with some uncertainty and inaccuracy.

Comment. The limit of the success rate is often taken as a *definition* of the notion of ‘probability’. But this is not really a definition. (Awkward question: what would be the meaning of the law of large numbers if this were the definition of probability?)

1.2. Making probability non-commutative.

We wish to make the mathematical structure $(\Omega, \Sigma, \mathbb{P})$ non-commutative. How is this to be done?

In the last ten years or so a successful strategy has become popular in mathematics. The most widely known example of this strategy is no doubt *non-commutative geometry*, as explained in the imaginative book of Alain Connes (1990,1994). Nowadays we have non-commutative topological spaces, quantum groups and non-commutative differential calculus. But the oldest example is non-commutative integration theory leading to non-commutative probability, which basically goes back to von Neumann (1932) and Segal (1953).

In all these examples the strategy to make some classical mathematical structure non-commutative consists of three steps.

1. Encode the information contained in the classical structure into an appropriate algebra of functions on it.
2. Characterise the resulting algebra axiomatically. One of the axioms will be commutativity.
3. Drop the commutativity axiom.

Let us apply this strategy to the structure of a probability space.

Step 1. Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space. We are looking for an algebra \mathcal{A} of functions

$$f : \Omega \rightarrow \mathbb{C} ,$$

that will reflect in some way Σ and \mathbb{P} .

Let us take the functions *measurable* with respect to Σ , i.e. let us require for all $a, b \in \mathbb{R}$:

$$\{ \omega \in \Omega \mid a \leq \operatorname{Re} f(\omega) \leq b \} \in \Sigma .$$

Furthermore we would like the expectation

$$\varphi(f) := \int_{\Omega} f(\omega) \mathbb{P}(d\omega)$$

to be defined for all $f \in \mathcal{A}$. So let us require f to be a bounded function:

$$\|f\|_{\infty} := \sup_{\omega \in \Omega} |f(\omega)| < \infty .$$

Finally, we do not want to distinguish functions that are equal almost surely:

$$f \sim g \quad \text{if} \quad \mathbb{P}(\{ \omega \in \Omega \mid f(\omega) = g(\omega) \}) = 1 .$$

We have now chosen the algebra

$$\mathcal{A} := L^{\infty}(\Omega, \Sigma, \mathbb{P}) ,$$

of equivalence classes under \sim of bounded measurable functions $f: \Omega \rightarrow \mathbb{C}$.

Our algebraic structure will be the pair (\mathcal{A}, φ) .

Now we must check whether we have faithfully encoded all the relevant information in $(\Omega, \Sigma, \mathbb{P})$. Clearly, the triple $(\Omega, \Sigma, \mathbb{P})$ determines $L^{\infty}(\Omega, \Sigma, \mathbb{P})$ uniquely. In the converse direction, we recover a σ -algebra $\tilde{\Sigma}$ by putting

$$\tilde{\Sigma} := \{ p \in L^{\infty}(\Omega, \Sigma, \mathbb{P}) \mid p = p^2 = p^* \} ,$$

which however is not isomorphic to Σ , since we have identified almost everywhere equal functions. What we find is the so-called *measure algebra*, i.e. the quotient of Σ by the equivalence

$$E \sim F, \text{ meaning: } \mathbb{P}((E \setminus F) \cup (F \setminus E)) = 0 .$$

We regard this simplification as a gain rather than a loss.

We finally (almost) reobtain the probability measure \mathbb{P} by putting

$$\tilde{\mathbb{P}} : \tilde{\Sigma} \rightarrow [0, 1] : p \mapsto \varphi(p) .$$

Step 2.

Characterisation. $\mathcal{A} := L^\infty(\Omega, \Sigma, \mathbb{P})$ is a commutative von Neumann algebra and $\varphi : f \mapsto \int f d\mathbb{P}$ is a faithful normal state on \mathcal{A} .

In order to understand this statement we have to give a few definitions from functional analysis.

A linear operator A on a Hilbert space \mathcal{H} is called *bounded* if

$$\sup\{ \|A\psi\| \mid \psi \in \mathcal{H}, \|\psi\| = 1 \} < \infty.$$

The supremum on the left is called the (*operator*) *norm* of A , written as $\|A\|$.

A *von Neumann algebra* is a collection \mathcal{A} of bounded linear operators on a Hilbert space \mathcal{H} with the following properties.

(i) \mathcal{A} is a linear space:

$$\left. \begin{array}{l} A, B \in \mathcal{A} \\ \lambda \in \mathbb{C} \end{array} \right\} \implies \left\{ \begin{array}{l} A + B \in \mathcal{A}; \\ \lambda A \in \mathcal{A}. \end{array} \right.$$

(ii) \mathcal{A} is a *-algebra:

$$A, B \in \mathcal{A} \implies \left\{ \begin{array}{l} AB \in \mathcal{A}; \\ A^* \in \mathcal{A}. \end{array} \right.$$

(iii) \mathcal{A} is strongly closed:

$$\left. \begin{array}{l} A_i \in \mathcal{A}, \\ \forall \psi \in \mathcal{H} : A_i \psi \longrightarrow A \psi \quad (i \longrightarrow \infty) \end{array} \right\} \implies A \in \mathcal{A}.$$

We shall always assume that \mathcal{A} contains the identity operator $\mathbf{1}$.

A *state* φ on a von Neumann algebra \mathcal{A} is a functional $\varphi : \mathcal{A} \rightarrow \mathbb{C}$ with the properties

(i) *linearity*:

$$\left. \begin{array}{l} A, B \in \mathcal{A} \\ \lambda \in \mathbb{C} \end{array} \right\} \implies \left\{ \begin{array}{l} \varphi(A + B) = \varphi(A) + \varphi(B); \\ \varphi(\lambda A) = \lambda \varphi(A). \end{array} \right.$$

(ii) *positivity*:

$$\forall A \in \mathcal{A} : \varphi(A^*A) \geq 0.$$

(iii) *normalisation*:

$$\varphi(\mathbf{1}) = 1.$$

Moreover, a state φ is called *faithful* if $\varphi(A^*A) = 0$ implies that $A = 0$.

It is called *normal* if for every increasing sequence A_1, A_2, A_3, \dots in \mathcal{A} with strong limit A we have $\lim_{n \rightarrow \infty} \varphi(A_n) = \varphi(A)$.

An element f of $L^\infty(\Omega, \Sigma, \mathbb{P})$ determines an operator M_f on the Hilbert space $L^2(\Omega, \Sigma, \mathbb{P})$ by

$$(M_f \psi)(\omega) := f(\omega) \psi(\omega).$$

Proposition 1.1. *Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space. Then the algebra*

$$\mathcal{A} := \{ M_f \mid f \in L^\infty(\Omega, \Sigma, \mathbb{P}) \}$$

is a (commutative) von Neumann algebra of operators on $\mathcal{H} := L^2(\Omega, \Sigma, \mathbb{P})$, and $\varphi : M_f \mapsto \int f d\mathbb{P}$ is a faithful normal state on \mathcal{A} .

Proof. The point is to show that \mathcal{A} is strongly closed. So let f_1, f_2, f_3, \dots be a sequence of L^∞ -functions such that M_{f_n} tends strongly to some bounded operator X on $\mathcal{H} := L^2(\Omega, \Sigma, \mathbb{P})$, i.e. for all $\psi \in \mathcal{H}$ we have

$$L^2\text{-}\lim_{n \rightarrow \infty} f_n \psi = X \psi .$$

Without loss of generality we may assume that $\|X\| = 1$. We must show that $X = M_f$ for some $f \in L^\infty$. Put $f := X1$. Then for all $g \in L^\infty$ we have

$$Xg = L^2\text{-}\lim_{n \rightarrow \infty} f_n g = L^2\text{-}\lim_{n \rightarrow \infty} M_g f_n = M_g \left(L^2\text{-}\lim_{n \rightarrow \infty} f_n \right) = M_g f = fg .$$

Now let the event E_ε for $\varepsilon > 0$ be defined by

$$E_\varepsilon := \{ \omega \in \Omega \mid |f(\omega)|^2 \geq 1 + \varepsilon \} .$$

Then, since $\|X\| \leq 1$,

$$\mathbb{P}(E_\varepsilon) = \|1_{E_\varepsilon}\|^2 \geq \|X1_{E_\varepsilon}\|^2 = \|f1_{E_\varepsilon}\|^2 = \int_{E_\varepsilon} |f|^2 d\mathbb{P} \geq (1 + \varepsilon)\mathbb{P}(E_\varepsilon) ,$$

and it follows that $\mathbb{P}(E_\varepsilon) = 0$. Since this holds for all $\varepsilon > 0$, we have $|f| \leq 1$ almost everywhere w.r.t. \mathbb{P} . So $f \in L^\infty(\Omega, \Sigma, \mathbb{P})$. Finally, since the operators X and M_f are both bounded and coincide on the dense subspace L^∞ of \mathcal{H} , they are equal. \square

Theorem 1.2 (Gel'fand). *Let \mathcal{A} be a commutative von Neumann algebra and φ a faithful normal state on \mathcal{A} . Then there exists a probability space $(\Omega, \Sigma, \mathbb{P})$ and an linear correspondence $A \mapsto f_A$ from \mathcal{A} to $L^\infty(\Omega, \Sigma, \mathbb{P})$ such that*

$$\begin{aligned} f_{AB} &= f_A f_B , \\ f_{A^*} &= \overline{f_A} , \\ \|f_A\| &= \|A\| , \\ \mathbb{E}(f_A) &:= \int_{\Omega} f_A d\mathbb{P} = \varphi(A) . \end{aligned}$$

So in fact \mathcal{A} and $L^\infty(\Omega, \Sigma, \mathbb{P})$ are isomorphic, and φ corresponds to expectation with respect to \mathbb{P} .

This is Gel'fand's theorem on commutative C*-algebra's, applied to von Neumann algebras (a subclass). We refer to the literature for its complete proof. (Cf., for instance, Douglas (1972) or Pedersen (1979).) Here we shall only treat the finite dimensional case.

The basic idea for this case is that \mathcal{A} is an algebra of $n \times n$ -matrices which can all be simultaneously diagonalised. The diagonal can then be decomposed into blocks, on

which all the matrices in \mathcal{A} are multiples of the identity matrix. We can then choose Ω to be the set of blocks, take Σ to be 2^Ω , and define a probability measure on Ω by

$$\mathbb{P}(\{\omega\}) := \varphi(P_\omega) ,$$

where P_ω is the identity on block ω , and zero everywhere else. Then clearly every element A of \mathcal{A} can be written in the form

$$A = \sum_{\omega \in \Omega} f(\omega) P_\omega ,$$

and the correspondence between A and f is a *-isometry between $l^\infty(\Omega)$ and \mathcal{A} connecting \mathbb{P} and φ in the correct way.

But this approach is useless for the infinite dimensional case. To get some idea how the general proof works, we adopt a more algebraic line of reasoning. We note that in the picture sketched above the number $\omega(A)$ which appears in block ω in the diagonal form of the matrix A , is linear as a function of A and also *multiplicative*:

$$\forall_{A,B \in \mathcal{A}} : \quad \omega(AB) = \omega(A)\omega(B) .$$

Furthermore we observe that the set of matrices which are mapped to 0 by some multiplicative linear functional ω is a so-called *maximal ideal*:

Definition. A subalgebra \mathcal{I} of a commutative algebra \mathcal{A} is called an *ideal* if for all $A \in \mathcal{A}$ and all $B \in \mathcal{I}$ the product AB lies in \mathcal{I} again. An ideal is called *maximal* if it is not included in any larger ideal other than \mathcal{A} itself.

Lemma 1.3. *Let \mathcal{A} be a commutative von Neumann algebra (of finite dimension). Then there is a one-to-one correspondence between multiplicative linear functionals $\omega : \mathcal{A} \rightarrow \mathbb{C}$ and maximal ideals $\mathcal{I} \subset \mathcal{A}$ given by*

$$A - \omega(A) \cdot \mathbf{1} \in \mathcal{I} . \tag{1}$$

Proof. Let $\omega : \mathcal{A} \rightarrow \mathbb{C}$ be linear and multiplicative, and define

$$\mathcal{I}_\omega := \{ A \in \mathcal{A} \mid \omega(A) = 0 \} .$$

Clearly (1) holds for this subalgebra \mathcal{I}_ω . But \mathcal{I}_ω is also an ideal, since for all $B \in \mathcal{I}_\omega$ and all $A \in \mathcal{A}$:

$$\omega(AB) = \omega(A)\omega(B) = 0, \quad \text{so} \quad AB \in \mathcal{I}_\omega .$$

Finally, the ideal \mathcal{I}_ω is maximal, for if we choose any element A outside \mathcal{I}_ω , then by (1) the linear space spanned by \mathcal{I} and A contains $\mathbf{1}$, so it must be all of \mathcal{A} .

Conversely, let \mathcal{I} be some maximal ideal of \mathcal{A} . We claim that for all $A \in \mathcal{A}$ there is a complex number z such that $A - z \cdot \mathbf{1} \in \mathcal{I}$. To see this, choose $A \notin \mathcal{I}$ and consider the ideal

$$\mathcal{I}_A := \{ AB + J \mid B \in \mathcal{A}, J \in \mathcal{I} \} .$$

Choosing $B = 0$ we see that \mathcal{I}_A includes \mathcal{I} , and choosing $B = \mathbf{1}$ and $J = 0$ we see that $A \in \mathcal{I}_A$, so $\mathcal{I}_A \neq \mathcal{I}$. Since \mathcal{I} is maximal, we must have $\mathcal{I}_A = \mathcal{A}$. In particular, $\mathbf{1} \in \mathcal{I}_A$. We conclude that

$$A \notin \mathcal{I} \implies \exists_{B \in \mathcal{A}} : (A + \mathcal{I})(B + \mathcal{I}) = \mathbf{1} + \mathcal{I} .$$

So in the quotient \mathcal{A}/\mathcal{I} every nonzero element is invertible. Let π be the quotient map $\mathcal{A} \rightarrow \mathcal{A}/\mathcal{I} : A \mapsto A + \mathcal{I}$. Then we have for all $z \in \mathbb{C}$:

$$\pi(A - z \cdot \mathbf{1}) = 0 \quad \text{or} \quad \pi(A - z \cdot \mathbf{1}) \text{ is invertible.}$$

But we can not have $\pi(A - z \cdot \mathbf{1})$ invertible for *all* $z \in \mathbb{C}$, for then

$$z \mapsto (\pi(A - z \cdot \mathbf{1}))^{-1}$$

would be a nonzero analytic map $\mathbb{C} \rightarrow \mathcal{A}/\mathcal{I}$ tending to 0 at infinity, which is impossible. So $\pi(A - z \cdot \mathbf{1}) = 0$ for some $z \in \mathbb{C}$, i.e. $A - z \cdot \mathbf{1} \in \mathcal{I}$. Put $\omega(A) := z$. We shall show that ω is multiplicative. Take $A, B \in \mathcal{A}$. By definition of ω , $A - \omega(A) \cdot \mathbf{1}$ and $B - \omega(B) \cdot \mathbf{1}$ are in \mathcal{I} , and it follows that

$$AB - \omega(A)\omega(B) = (A - \omega(A) \cdot \mathbf{1})B + \omega(A)(B - \omega(B) \cdot \mathbf{1})$$

also lies in \mathcal{I} . So $\omega(AB) = \omega(A)\omega(B)$. □

Lemma 1.4. *Let \mathcal{A} be a von Neumann algebra (of finite dimension), and let Ω denote the set of linear multiplicative functionals on \mathcal{A} . Then*

$$\{ \omega(A) \mid \omega \in \Omega \} = \text{sp}(A) .$$

Proof. If $\omega \in \Omega$, then $A - \omega(A) \cdot \mathbf{1}$ is not invertible, so $\omega(A) \in \text{sp}(A)$. Conversely, if $z \in \text{sp}(A)$, i.e. $A - z \cdot \mathbf{1}$ is not invertible, then the set

$$\mathcal{J} := \{ (A - z \cdot \mathbf{1})B \mid B \in \mathcal{A} \}$$

is an ideal that does not contain $\mathbf{1}$. As \mathcal{A} is finite dimensional, there must be some maximal ideal $\mathcal{I} \supset \mathcal{J}$. But then $A - z \cdot \mathbf{1} \in \mathcal{I}$. Let ω be the multiplicative linear functional associated with \mathcal{I} by Lemma 1.3. Then $\omega(A) = z$. □

Proof of Theorem 1.2. Let (\mathcal{A}, φ) be given and let Ω be the set of linear multiplicative functionals on \mathcal{A} . For any $A \in \mathcal{A}$, we define the function f_A by

$$f_A(\omega) := \omega(A) .$$

Then

$$f_{AB}(\omega) = \omega(AB) = \omega(A)\omega(B) = f_A(\omega)f_B(\omega)$$

and

$$f_{A^*}(\omega) = \omega(A^*) = \overline{\omega(A)} = \overline{f_A(\omega)} .$$

Also,

$$\begin{aligned} \|A\|^2 &= \|A^*A\| = \sup\{ z \mid z \in \text{sp}(A^*A) \} = \sup\{ \omega(A^*A) \mid \omega \in \Omega \} \\ &= \sup\{ |\omega(A)|^2 \mid \omega \in \Omega \} = \sup\{ |f_A(\omega)|^2 \mid \omega \in \Omega \} = \|f_A\|_\infty^2 . \end{aligned}$$

From the isometric relation $\|A\| = \|f_A\|$ we may conclude that the correspondence $A \mapsto f_A$ is injective. It is also surjective, since it separates points: if $\omega_1 \neq \omega_2$, then there must be some $A \in \mathcal{A}$ for which $\omega_1(A) \neq \omega_2(A)$.

In particular, let P_ω be such that $f_{P_\omega} = \delta_\omega$, and define a probability measure on Ω by

$$\mathbb{P}(\{\omega\}) := \varphi(P_\omega) .$$

Then, since for all $f \in l^\infty(\Omega)$

$$f = \sum_{\omega \in \Omega} f(\omega) \delta_\omega ,$$

we have that every $A \in \mathcal{A}$ can be written

$$A = \sum_{\omega \in \Omega} f_A(\omega) P_\omega .$$

Therefore for all $A \in \mathcal{A}$:

$$\begin{aligned} \mathbb{E}(f_A) &= \sum_{\omega \in \Omega} f_A(\omega) \mathbb{P}(\{\omega\}) = \sum_{\omega \in \Omega} f_A(\omega) \varphi(P_\omega) \\ &= \varphi \left(\sum_{\omega \in \Omega} f_A(\omega) P_\omega \right) = \varphi(A) . \end{aligned}$$

□

Step 3. We now drop the commutativity requirement to arrive at the following definition.

Definition. By a *non-commutative probability space* we mean a pair (\mathcal{A}, φ) , where \mathcal{A} is a von Neumann algebra of operators on some Hilbert space \mathcal{H} , and φ is a normal state on \mathcal{A} . If φ is faithful, the probability space is called *non-degenerate*.

1.3. Interpretation of quantum probability.

Let us now carry the most important concepts of probability theory over to our new probability space (\mathcal{A}, φ) .

An *event* is an orthogonal projection $E \in \mathcal{A}$, i.e. an element satisfying $E^2 = E = E^*$. Two events E and F will be called *compatible* if EF is also an event, which means that E and F commute:

$$EF = (EF)^* = FE .$$

A stochastic experiment is said to be *modelled* by a non-commutative probability space (\mathcal{A}, φ) if the following is the case.

- The experiment can be repeated arbitrarily often. In contrast to classical stochastic experiments, as envisaged by Kolmogorov, we allow adjustment of the observation equipment between the trials in order to ask different questions.

- Sufficiently many of the orthogonal projections in \mathcal{A} have an interpretation as a statement about the outcome of the experiment that can be tested by observation.
- If E and F are compatible questions, they can be asked together. EF denotes the event that both E and F occur, and $E \vee F := E + F - EF$ is the event that either E or F occurs. So mutually exclusive events correspond to mutually orthogonal subspaces of \mathcal{H} .

Incompatible questions can not be asked together: the instruments needed to measure them obstruct each other. They can be asked in different trials by readjusting the apparatus in between. Inside one single trial it is sometimes possible to ask incompatible questions one *after* the other, but then the fact that one question was posed, may influence the answer to subsequent questions.

- If the questions E_1, E_2, \dots, E_k are asked in the same trial and in this order, the probability that they will all be answered ‘yes’ is

$$\varphi(E_1 E_2 \cdots E_{k-1} E_k E_{k-1} \cdots E_2 E_1) .$$

In quantum probability the weak law of large numbers holds again for independent *and compatible* trials.

It is a surprising fact that nature — at least on a small scale — appears to be governed by a non-commutative probability theory, called quantum mechanics.

2. Examples and motivation.

2.1. The quantum coin toss: ‘spin’.

The simplest non-commutative von Neumann algebra is M_2 , the algebra of all 2×2 matrices with complex entries. And the simplest con-commutative probability space is $(M_2, \frac{1}{2}\text{tr})$, the ‘fair quantum coin toss’.

The events in this probability space are the orthogonal projections in M_2 : the complex 2×2 matrices E satisfying

$$E^2 = E = E^* .$$

Let us see what these projections look like. Since E is self-adjoint, it must have two real eigenvalues, and since $E^2 = E$ these must both be 0 or 1. So we have three possibilities.

- (0) Both are 0; i.e. $E = 0$.
- (1) One of them is 0 and the other is 1.
- (2) Both are 1; i.e. $E = \mathbf{1}$.

In case (1), E is a one-dimensional projection satisfying

$$\text{tr } E = 0 + 1 = 1 \quad \text{and} \quad \det E = 0 \cdot 1 = 0 .$$

As $E^* = E$ and $\text{tr } E = 1$ we may write

$$E = \frac{1}{2} \begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix} .$$

Then $\det E = 0$ implies that

$$\frac{1}{2}((1-z^2) - (x^2 + y^2)) = 0 \quad \implies \quad x^2 + y^2 + z^2 = 1 .$$

So the one-dimensional projections in M_2 are parametrised by the unit sphere S_2 .

Notation. For $a = (a_1, a_2, a_3) \in \mathbb{R}^3$ let us write

$$\sigma(a) := \begin{pmatrix} a_3 & a_1 - ia_2 \\ a_1 + ia_2 & -a_3 \end{pmatrix} = a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3 ,$$

where σ_1, σ_2 and σ_3 are the *Pauli matrices*

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

We note that for all $a, b \in \mathbb{R}^3$ we have

$$\sigma(a)\sigma(b) = \langle a, b \rangle \cdot \mathbf{1} + i\sigma(a \times b) . \tag{2}$$

Let us write

$$E(a) := \frac{1}{2}(\mathbf{1} + \sigma(a)), \quad (\|a\| = 1) . \tag{3}$$

In the same way the possible states on M_2 can be calculated. We find that

$$\varphi(A) = \text{tr}(\rho A) \quad \text{where} \quad \rho = \rho(a) := \frac{1}{2}(\mathbf{1} + \sigma(a)), \quad \|a\| \leq 1 . \tag{4}$$

We summarise:

Proposition 1.5. *The states on M_2 are parametrised by the unit ball in \mathbb{R}^3 , as in (4), and the one-dimensional projections in M_2 are parametrised by the unit sphere as in (3). The probability of the event $E(a)$ in the state $\rho(b)$ is given by*

$$\text{tr}(\rho(b)E(a)) = \frac{1}{2}(1 + \langle a, b \rangle).$$

The events $E(a)$ and $E(b)$ are compatible if and only if $a = \pm b$. Moreover we have for all $a \in S_2$:

$$E(a) + E(-a) = \mathbf{1}, \quad E(a)E(-a) = 0.$$

Proof. Calculate. □

Interpretation. The probability distribution of the quantum coin toss is given by a vector b in the three-dimensional unit ball. For every a on the unit sphere we can say with probability one that of the two events $E(a)$ and $E(-a)$ exactly one will occur, $E(a)$ having probability $\frac{1}{2}(1 + \langle a, b \rangle)$. So we have a classical coin toss (with probability for heads equal to $\frac{1}{2}(1 + \langle a, b \rangle)$) for every direction in \mathbb{R}^3 . The coin tosses in different directions are incompatible.

Particular case: the quantum fair coin is modelled by $(M_2, \frac{1}{2}\text{tr})$.

The quantum coin toss is realised in nature: the spin direction of a particle with total spin $\frac{1}{2}$ behaves in this way.

Photons.

There is a second natural way to parametrise the one-dimensional projections in M_2 , which is closer to the description of polarisation of photons.

A one-dimensional projection corresponds to a (complex) line in \mathbb{C}^2 , and such a line can be characterised by its slope, a number $z \in \mathbb{C} \cup \{\infty\}$.

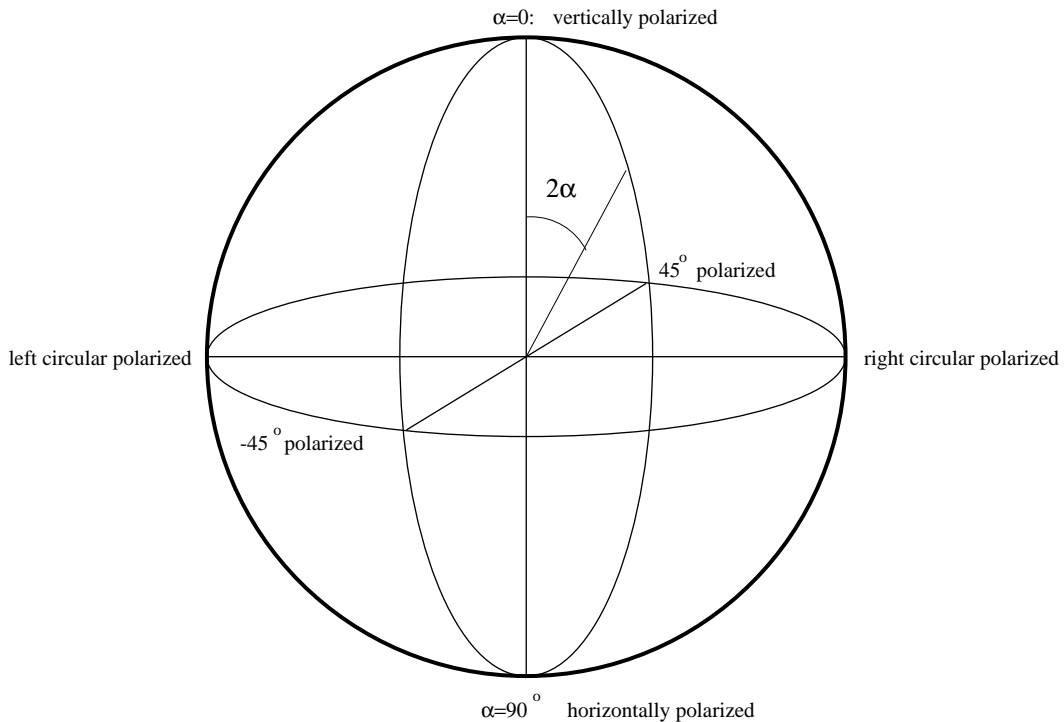


FIG. 2: Polarisation directions of a photon plotted on the sphere

Exercise. Let $f : \mathbb{C} \cup \{\infty\} \rightarrow S_2$ be given by

$$\begin{aligned} f(0) &:= (0, 0, 1) ; \\ f(\infty) &:= (0, 0, -1) ; \\ f(re^{i\varphi}) &:= (\sin \vartheta \cos \varphi, \sin \vartheta \sin \varphi, \cos \vartheta) \\ &\text{with } \vartheta = 2 \arctan r, \quad r \in (0, \infty), \varphi \in [0, \pi) . \end{aligned}$$

Show that $E(f(z))$ is the one-dimensional projection onto the line in \mathbb{C}^2 with slope $z \in \mathbb{C}$.

In particular, the projection $F(\alpha)$ on the line with real slope $\tan \alpha$ with $\alpha \in [0, \pi)$ is given by

$$F(\alpha) = \begin{pmatrix} \cos^2 \alpha & \cos \alpha \sin \alpha \\ \cos \alpha \sin \alpha & \sin^2 \alpha \end{pmatrix} = E(\sin 2\alpha, \sin 2\alpha, \cos 2\alpha) . \quad (5)$$

Finally, any atomic or molecular system, only two energy levels of which are of importance in the experiment, can be described by some (M_2, φ) .

2.2. Two spins and Bell's inequality.

In quantum probability certain strong correlations are possible, called *entanglement*. The simplest case is that of two entangled spins. Entanglement shows itself most dramatically when the two spins are far apart: they show a degree of coordination which classically would only be possible by action at a distance. This has lead many people to believe that in quantum mechanics action at a distance takes place. We claim that there is no reason for this belief.

Let us start by describing the experimental facts.

In 1982 the following experiment was performed by A. Aspect at Orsay (near Paris). An optical device, containing Ca atoms and a pumping laser, produced pairs of photons which flew apart in opposite directions. In two such directions, say 'on the left' and 'on the right' the polarisations of the photons was measured using polarisation filters. When the polarisation filter on the left made an angle α with the vertical direction, and on the right an angle β , the probability for two photons of the same pair to pass through both filters was found to be

$$\frac{1}{2} \sin^2(\alpha - \beta) .$$

So when α and β were the same, no pair of photons passed, but when they were at right angles to each other, half of the pairs passed through both filters. (The photons are said to have 'opposite polarisations', an expression which is made clear by fig. 2.)

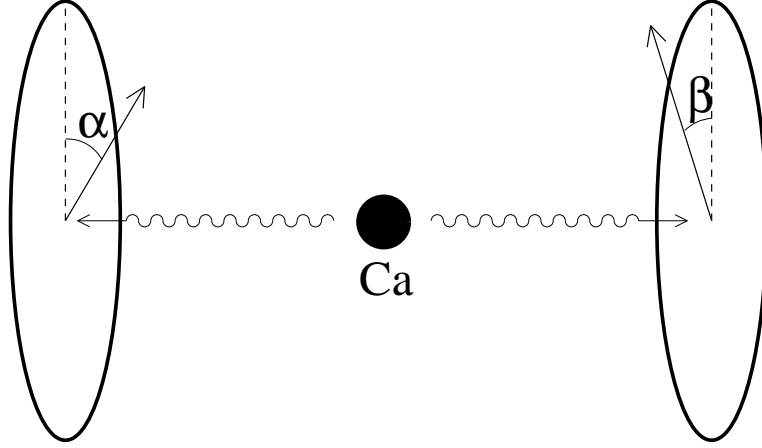


FIG 3: The experiment of A. Aspect

In the experiment care was taken to prevent the polariser setting on the left from influencing the outcome on the right and vice versa. This was done by making the choice of the polariser direction *during the flight of the photons*, so that any influence of the said kind would have to travel faster than light, and would therefore be impossible according to special theory of relativity.

Models that explain the experiment.

Let us now see how this situation can be modelled stochastically.

We need some probability space (\mathcal{A}, φ) and for every pair (α, β) of polariser settings we need a pair of events

$$P(\alpha, \beta) \quad \text{and} \quad Q(\alpha, \beta) .$$

$P(\alpha, \beta)$ says that, with the settings α and β , the photon on the left passes the filter on the left; $Q(\alpha, \beta)$ says that the photon on the right passes the filter on the right. These events must be compatible since both of them are measured. We require that each has probability $\frac{1}{2}$ and that the probability for both to occur is

$$\varphi(P(\alpha, \beta)Q(\alpha, \beta)) = \frac{1}{2} \sin^2(\alpha - \beta) . \quad (6)$$

Definition. We shall call a model of the Aspect experiment *local* if $P(\alpha, \beta)$ does not depend on β and $Q(\alpha, \beta)$ does not depend on α :

$$P(\alpha, \beta) = P(\alpha) \quad \text{and} \quad Q(\alpha, \beta) = Q(\beta) .$$

We call the model *classical* if all the $P(\alpha, \beta)$ and $Q(\alpha, \beta)$ for different values of α and β commute.

Theorem 2.1. *There exists no local classical model of the Aspect experiment.*

To prove this theorem, we employ the following version of Bell's inequality.

Proposition 2.2. (Bell 1965; Clauser and Shimony 1978; Kümmerner and Maassen 1996.) *Let P_1, P_2, Q_1 and Q_2 be classical 0-1-valued random variables on a probability space $(\Omega, \Sigma, \mathbb{P})$. Then*

$$\mathbb{P}[P_1 = Q_1] \leq \mathbb{P}[P_1 = Q_2] + \mathbb{P}[P_2 = Q_2] + \mathbb{P}[P_2 = Q_1] .$$

Proof. For every $\omega \in \Omega$ we have

$$\begin{aligned} P_1(\omega) = Q_1(\omega) &\implies \\ P_1(\omega) = Q_2(\omega) &\text{ or } P_2(\omega) = Q_2(\omega) \text{ or } P_2(\omega) = Q_1(\omega). \end{aligned} \quad (7)$$

□

Now suppose that we have a local classical model (\mathcal{A}, φ) with events $P(\alpha)$ and $Q(\beta)$, $(\alpha, \beta \in [0, \pi))$ satisfying $\varphi(P(\alpha)) = \varphi(Q(\beta)) = \frac{1}{2}$ and (6). By Gel'fand's theorem they can all be represented as 0-1-valued functions on some probability space $(\Omega, \Sigma, \mathbb{P})$. Then since

$$\mathbb{P}[P(\alpha) = Q(\beta) = 0] = \frac{1}{2} - \mathbb{P}[P(\alpha) = 1 \text{ and } Q(\beta) = 0] = \mathbb{P}[P(\alpha) = Q(\beta) = 1],$$

we have

$$\mathbb{P}[P(\alpha) = Q(\beta)] = \mathbb{P}[P(\alpha) = Q(\beta) = 0] + \mathbb{P}[P(\alpha) = Q(\beta) = 1] = \sin^2(\alpha - \beta).$$

Now choose two polariser settings α_1 and α_2 on the left, and two settings β_1 and β_2 on the right, and apply Bell's inequality to $P_1 := P(\alpha_1)$, $P_2 := P(\alpha_2)$, $Q_1 := Q(\beta_1)$, and $Q_2 := Q(\beta_2)$ to find that

$$\sin^2(\alpha_1 - \beta_1) \leq \sin^2(\alpha_1 - \beta_2) + \sin^2(\alpha_2 - \beta_2) + \sin^2(\alpha_2 - \beta_1).$$

This inequality is violated for the choice $\alpha_1 = 0^\circ$, $\alpha_2 = 60^\circ$, $\beta_1 = 90^\circ$, and $\beta_2 = 30^\circ$, in which case it reads

$$1 \leq \frac{1}{4} + \frac{1}{4} + \frac{1}{4}.$$

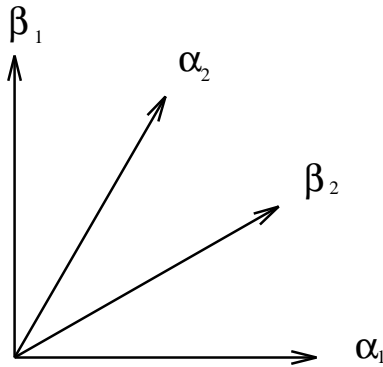


FIG. 4: Polariser settings.

This contradiction shows that no local classical model of the Aspect experiment exists.

Theorem2.2. *There exists a local quantum model of the Aspect experiment.*

Proof. Let

$$\mathcal{A} := M_2 \otimes M_2 \quad \text{and} \quad \varphi(A) := \text{tr}(\rho A)$$

with

$$\rho := \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

For $\alpha, \beta \in [0, \pi)$, put

$$P(\alpha) := F(\alpha) \otimes \mathbf{1} \quad \text{and} \quad Q(\beta) := \mathbf{1} \otimes F(\beta) ,$$

where $F(\alpha)$ is given by (5). Then

$$\begin{aligned} \varphi(P(\alpha)) &= \varphi \left(\begin{array}{cccc} \cos^2 \alpha & 0 & \cos \alpha \sin \alpha & 0 \\ 0 & \cos^2 \alpha & 0 & \cos \alpha \sin \alpha \\ \cos \alpha \sin \alpha & 0 & \sin^2 \alpha & 0 \\ 0 & \cos \alpha \sin \alpha & 0 & \sin^2 \alpha \end{array} \right) \\ &= \frac{1}{2}(\cos^2 \alpha + \sin^2 \alpha) = \frac{1}{2} , \end{aligned}$$

and also $\varphi(Q(\beta)) = \frac{1}{2}$ for all β . For all $\alpha, \beta \in [0, \pi)$ the projections $P(\alpha)$ and $Q(\beta)$ commute and the probability for both to occur is

$$\begin{aligned} \varphi(P(\alpha)Q(\beta)) &= \varphi(F(\alpha) \otimes F(\beta)) \\ &= \varphi \left(\begin{array}{cccc} \cos^2 \alpha \cos^2 \beta & \cos^2 \alpha \cos \beta \sin \beta & \cos \alpha \sin \alpha \cos^2 \beta & \cos \alpha \sin \alpha \cos \beta \sin \beta \\ \cos^2 \alpha \cos \beta \sin \beta & \cos^2 \alpha \sin^2 \beta & \cos \alpha \sin \alpha \cos \beta \sin \beta & \cos \alpha \sin \alpha \sin^2 \beta \\ \cos \alpha \sin \alpha \cos^2 \beta & \cos \alpha \sin \alpha \cos \beta \sin \beta & \sin^2 \alpha \cos^2 \beta & \sin^2 \alpha \cos \beta \sin \beta \\ \cos \alpha \sin \alpha \cos \beta \sin \beta & \cos \alpha \sin \alpha \sin^2 \beta & \sin^2 \alpha \cos \beta \sin \beta & \sin^2 \alpha \sin^2 \beta \end{array} \right) \\ &= \frac{1}{2}(\cos^2 \alpha \sin^2 \beta + \sin^2 \alpha \cos^2 \beta - 2 \cos \alpha \sin \alpha \cos \beta \sin \beta) \\ &= \frac{1}{2}(\cos \alpha \sin \beta - \sin \alpha \cos \beta)^2 \\ &= \frac{1}{2} \sin^2(\alpha - \beta) . \end{aligned}$$

□

Comment. Many people are disturbed by the lack of a ‘realistic’ (i.e. classical) stochastic interpretation of quantum mechanics. They prefer to give up Einstein’s locality requirement, and try to construct classical (non-local) models, not only for a pair of entangled spins, but for all of quantum mechanics. (Bohm 1952 was a pioneering paper.) Although these attempts are to be taken serious, so far they have not arisen above the level of a classical, rather arbitrary add-on to the quantum theory.

Let us finally note that there exist plenty of *non-local* classical models for the Aspect experiment. Here is one:

$$\Omega := [0, 1]; \quad \Sigma := \text{Borel sets}; \quad \mathbb{P} := \text{Lebesgue measure.}$$

$$P(\alpha, \beta) := 1_{[\frac{1}{2}, 1]}; \quad Q(\alpha, \beta) := 1_{[\frac{1}{2} \sin^2(\alpha - \beta), \frac{1}{2}(1 + \sin^2(\alpha - \beta))]} .$$

About φ .

The state φ occurring in the quantum model $(M_2 \otimes M_2, \varphi)$ is called an *entangled* state because it is not a convex combination of product states. It *is* an affine combination, though.

Exercise. Show that

$$\varphi = 3\gamma - 2\tau ,$$

where τ is the normalised trace $\frac{1}{4}\text{tr}$ and γ is given by

$$\gamma(A) := \int_{S_2} \text{tr} ((E(a) \otimes E(-a))A) da .$$

Here da denotes normalised Lebesgue measure on the unit sphere S_2 .

2.3. The Bell game.

To illustrate the main point of the previous section, we shall again present the experiment, but this time in the form of a card game. Nature can win this game. Can you?

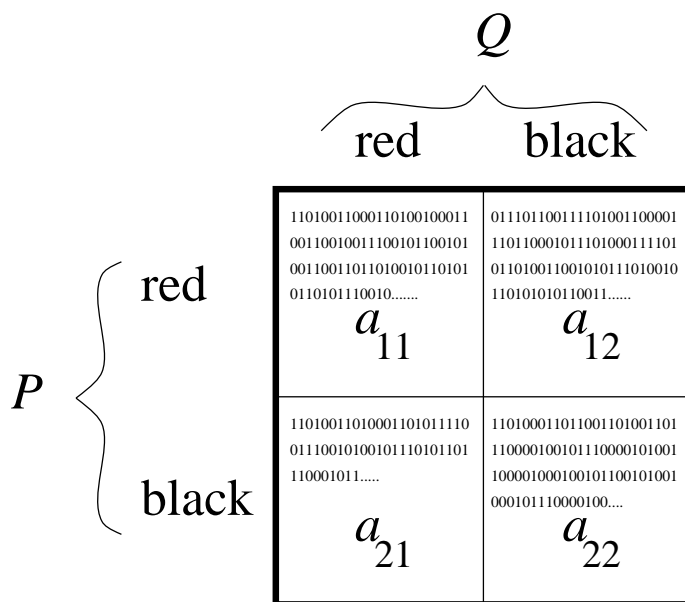


FIG.5: The Bell game.

Two players, P and Q , are sitting at a table. They are cooperating to achieve a single goal. There is an arbiter present to deal cards and to count points. On the table there is a board consisting of four squares as drawn in Fig. 5. There are dice and an ordinary deck of playing cards. The deck of cards is shuffled well. (In fact we shall assume that the deck of cards is an infinite sequence of independent cards, chosen fully at random.) First the players are given some time to make agreements on the strategy they are going to follow. Then the game starts, and from this moment on they are no longer allowed to communicate. The following sequence of actions is then repeated many times.

1. The dealer hands a card to P and a card to Q . Both look at their own card, but not at the other one's. (The only feature of the card that matters is its colour: red or black.)
2. The dice are thrown.

3. P and Q simultaneously say ‘yes’ or ‘no’, according to their own choice. They are free to make their answer depend on any information they possess, such as the color of their own card, the agreements made in advance, the numbers shown by the dice, the weather, the time, et cetera.
4. The cards are laid out on the table. The pair of colours of the cards determines one of the four squares on the board: these are labelled (red,red), (red,black), (black,red) and (black,black).
5. In the square so determined a 0 or a 1 is written: a 0 when the answers of P and Q have been different, a 1 if they have been the same.

In the course of time, the squares on the board get filled with 0’s and 1’s. The arbiter keeps track of the percentage of 1’s in proportion to the total number of digits in each square; we shall call the limits of these percentages as the game stretches out to infinity: a_{11} , a_{12} , a_{21} , and a_{22} . The aim of the game, for both P and Q , is to get a_{11} larger than the sum of the other three limiting percentages. So P and Q must try to give identical answers as often as they can when both their cards are red, but different answers otherwise.

‘PROPOSITION’. (Bell’s inequality for the game) P and Q cannot win the game by classical means, namely:

$$a_{11} \leq a_{12} + a_{21} + a_{22}.$$

‘Proof’.

The best P and Q can do, in order to win the game, is to agree upon some (possibly random) strategy for each turn. For instance, they may agree that P will always say ‘yes’ (i.e., $P_{\text{red}} = P_{\text{black}} = \text{‘yes’}$) and that Q will answer the question ‘Is my card red?’ (i.e., $Q_{\text{red}} = \text{‘yes’}$ and $Q_{\text{black}} = \text{‘no’}$). This will lead to a 1 in the (red,red) square or the (black,red) square or to a 0 in one of the other two. So if we would repeat this strategy very often, then on the long run we would get $a_{11} = a_{12} = 1$ and $a_{21} = a_{22} = 0$, disappointingly satisfying Bell’s inequality.

The above example is an extremal strategy. There are many (in fact, sixteen) strategies like this. By the pointwise version (7) of Bell’s inequality, none of these sixteen extremal strategies wins the game. Inclusion of the randomness coming from the dice yields a full polytope of random strategies, having the above sixteen as its extremal points. But since the inequalities are linear, this averaging procedure does not help. This ‘proves’ our ‘proposition’. Disbelievers are challenged to find a winning strategy. \square

Strangely enough, however, Nature does provide us with a strategy to win the game. Instead of the dice, put a Calcium atom on the table. When the cards have been dealt, P and Q put their polarizers in the direction indicated by their cards. If P has a red card, then he chooses the direction $\alpha_1 = 0$ (cf. Fig. 4). If his card is black, then he chooses $\alpha_2 = 60^\circ$. If Q has a red card, then he chooses $\beta_1 = 90^\circ$. If his card is black, then he chooses $\beta_2 = 30^\circ$. No information on the colours of the cards needs to be exchanged. When the Calcium atom has produced its photon pair, each player looks whether his own photon passes his own polarizer, and then says ‘yes’ if it does,

'no' if it does not. On the long run they will get $a_{11} = 1$, $a_{12} = a_{21} = a_{22} = \frac{1}{4}$, and thus they win the game.

So the Calcium atom, the quantummechanical die, makes possible what could not be done with the classical die.

3. Hilbert spaces, observables and von Neumann algebras.

3.1. Hilbert spaces.

A *Hilbert space* is a complex linear space \mathcal{H} with a function

$$\mathcal{H} \times \mathcal{H} \rightarrow \mathbb{C} : (\psi, \chi) \mapsto \langle \psi, \chi \rangle$$

having the following properties:

- (i) $\langle \psi, \chi_1 + \chi_2 \rangle = \langle \psi, \chi_1 \rangle + \langle \psi, \chi_2 \rangle$ for all $\psi, \chi_1, \chi_2 \in \mathcal{H}$;
- (ii) $\langle \psi, \lambda \chi \rangle = \lambda \langle \psi, \chi \rangle$ for all $\psi, \chi \in \mathcal{H}$ and all $\lambda \in \mathbb{C}$;
- (iii) $\langle \psi, \chi \rangle = \langle \chi, \psi \rangle$ for all $\psi, \chi \in \mathcal{H}$;
- (iv) $\langle \psi, \psi \rangle \geq 0$ for all $\psi \in \mathcal{H}$;
- (v) $\langle \psi, \psi \rangle = 0$ implies that $\psi = 0$;
- (vi) \mathcal{H} is complete in the norm $\psi \mapsto \|\psi\| := \langle \psi, \psi \rangle^{\frac{1}{2}}$,
i.e. if $\psi_1, \psi_2, \psi_3, \dots$ is a *Cauchy sequence*:

$$\lim_{n \rightarrow \infty} \sup_{m \geq n} \|\psi_n - \psi_m\| = 0 ,$$

then there is a vector $\psi \in \mathcal{H}$ such that

$$\lim_{n \rightarrow \infty} \|\psi_n - \psi\| = 0 .$$

The function $\langle \cdot, \cdot \rangle$ is called the *inner product* of the Hilbert space. If the conditions (v) and (vi) are not required, we call \mathcal{H} a *pre-Hilbert space*. Every Hilbert space is a Banach space, in particular for all $\psi, \chi \in \mathcal{H}$ the *triangle inequality* is valid:

$$\|\psi + \chi\| \leq \|\psi\| + \|\chi\| .$$

In a Hilbert space we have the *Cauchy-Schwarz inequality*:

$$|\langle \psi, \chi \rangle| \leq \|\psi\| \|\chi\| .$$

Let \mathcal{S} be a subset of \mathcal{H} . By \mathcal{S}^\perp we mean the closed linear subspace of \mathcal{H} given by

$$\mathcal{S}^\perp := \{ \psi \in \mathcal{H} \mid \forall \chi \in \mathcal{S} : \langle \chi, \psi \rangle = 0 \} .$$

By the *linear span* of \mathcal{S} , written as $\sqrt{\mathcal{S}}$, we mean the space of all finite linear combinations of elements of \mathcal{S} . Its closure $\overline{\sqrt{\mathcal{S}}}$ is the smallest closed subspace of \mathcal{H} which contains \mathcal{S} .

Proposition 3.1. *Let \mathcal{S} be a subset of a Hilbert space \mathcal{H} . Then every element ψ of \mathcal{H} can be written in a unique way as $\psi_1 + \psi_2$, where*

$$\psi_1 \in \overline{\sqrt{\mathcal{S}}} \quad \text{and} \quad \psi_2 \in \mathcal{S}^\perp .$$

In particular,

$$\overline{\sqrt{\mathcal{S}}} = \mathcal{S}^{\perp\perp} .$$

Proof. Choose $\psi \in \mathcal{H}$ and let

$$d := \inf_{\vartheta \in \bigvee \mathcal{S}} \|\psi - \vartheta\| .$$

Then d is the distance from ψ to the span of \mathcal{S} .

Let $\vartheta_1, \vartheta_2, \vartheta_3, \dots$ be a sequence in $\bigvee \mathcal{S}$ with

$$\lim_{n \rightarrow \infty} \|\vartheta_n - \psi\| = d .$$

For all $n, m \in \mathbb{N}$ we have by the parallelogram law

$$\|\vartheta_n + \vartheta_m - 2\psi\|^2 + \|\vartheta_n - \vartheta_m\|^2 = 2 \left(\|\vartheta_n - \psi\|^2 + \|\vartheta_m - \psi\|^2 \right) .$$

As $n, m \rightarrow \infty$, the right hand side tends to $4d^2$. Since $\|\frac{1}{2}(\vartheta_n + \vartheta_m) - \psi\| \geq d$ we must have $\|\vartheta_n - \vartheta_m\| \rightarrow 0$. So $\vartheta_1, \vartheta_2, \vartheta_3, \dots$ is a Cauchy sequence; let ψ_1 be its limit. Then $\psi_1 \in \overline{\bigvee \mathcal{S}}$. Finally we have for all $\chi \in \mathcal{S}$ and all $t \in \mathbb{R}$:

$$\|(\psi_1 + t\chi) - \psi\|^2 = \|\psi_1 - \psi\|^2 + 2t\operatorname{Re} \langle \psi_1 - \psi, \chi \rangle + t^2 \|\chi\|^2 ,$$

and since the left hand side must always be at least d^2 , it follows that $\psi_2 := \psi_1 - \psi$ is orthogonal to χ . This proves the first part of the theorem. To prove the second, note that $\mathcal{S}^{\perp\perp}$ is a closed subspace containing \mathcal{S} . So $\overline{\bigvee \mathcal{S}} \subset \mathcal{S}^{\perp\perp}$. Conversely suppose that $\psi \in \mathcal{S}^{\perp\perp}$. Then

$$\psi_2 = \psi - \psi_1 \in \mathcal{S}^{\perp\perp} \cap \mathcal{S}^\perp = \{0\} ,$$

so $\psi = \psi_1 \in \overline{\bigvee \mathcal{S}}$. □

3.2. Kolmogorov dilations of positive definite kernels.

Let \mathcal{S} be a set and let K be a *kernel* on \mathcal{S} , i.e. a function $\mathcal{S} \times \mathcal{S} \rightarrow \mathbb{C}$. Then K is called *positive definite* if for all $n \in \mathbb{N}$ and all n -tuples $(\lambda_1, \dots, \lambda_n) \in \mathbb{C}^n$ we have

$$\sum_{i=1}^n \sum_{j=1}^n \overline{\lambda_i} \lambda_j K(x_i, x_j) \geq 0 .$$

Theorem 3.2. (Kolmogorov's dilation theorem) *Let K be a positive definite kernel on a set \mathcal{S} . Then up to unitary equivalence there exists a unique Hilbert space \mathcal{H} and a unique embedding $V : \mathcal{S} \rightarrow \mathcal{H}$ such that*

$$\forall_{x, y \in \mathcal{S}} : \langle V(x), V(y) \rangle = K(x, y) \quad \text{and} \quad \overline{\bigvee V(\mathcal{S})} = \mathcal{H} .$$

Proof. Consider the space \mathcal{L} of all functions $\mathcal{S} \rightarrow \mathbb{C}$ with finite support. Then \mathcal{L} becomes a pre-Hilbert space if we define the (pre-)inner product

$$\langle \lambda, \mu \rangle := \sum_{x \in \mathcal{S}} \sum_{y \in \mathcal{S}} \overline{\lambda(x)} K(x, y) \mu(y).$$

To make a Hilbert space out of \mathcal{L} , we divide out the null space

$$\mathcal{N} := \{ \lambda \in \mathcal{L} \mid \langle \lambda, \lambda \rangle = 0 \}$$

to obtain the quotient space

$$\mathcal{L}/\mathcal{N} := \{ \lambda + \mathcal{N} \mid \lambda \in \mathcal{L} \},$$

and we form the completion \mathcal{H}_K of \mathcal{L}/\mathcal{N} . Now let $V_K : \mathcal{S} \rightarrow \mathcal{H}_K$ be given by

$$V_K(x) := \delta_x + \mathcal{N}.$$

Then \mathcal{H}_K is a Hilbert space and for all $x, y \in \mathcal{S}$:

$$\begin{aligned} \langle V_K(x), V_K(y) \rangle &= \langle \delta_x + \mathcal{N}, \delta_y + \mathcal{N} \rangle_{\mathcal{L}/\mathcal{N}} = \langle \delta_x, \delta_y \rangle_{\mathcal{L}} \\ &= \sum_{x' \in \mathcal{S}} \sum_{y' \in \mathcal{S}} \delta_x(x') K(x', y') \delta_y(y') = K(x, y). \end{aligned}$$

Now let $V : \mathcal{S} \rightarrow \mathcal{H}$ be another Kolmogorov dilation of K . Then we define a map

$$U_0 : \mathcal{L} \rightarrow \mathcal{H} : \quad \lambda \mapsto \sum_{x \in \mathcal{S}} \lambda(x) V(x).$$

This map is constant on the classes $\lambda + \mathcal{N}$: for $\lambda - \mu \in \mathcal{N}$ we have

$$\begin{aligned} \|U_0\lambda - U_0\mu\|^2 &= \left\| \sum_{x \in \mathcal{S}} (\lambda(x) - \mu(x)) V(x) \right\|^2 \\ &= \sum_{x \in \mathcal{S}} \sum_{y \in \mathcal{S}} \overline{(\lambda(x) - \mu(x))} K(x, y) (\lambda(y) - \mu(y)) = \langle \lambda - \mu, \lambda - \mu \rangle_{\mathcal{L}} = 0. \end{aligned}$$

So U_0 is in fact a map $\mathcal{L}/\mathcal{N} \rightarrow \mathcal{H}$. By the same calculation we find that U_0 is isometric. Since $\bigvee V(\mathcal{S})$ is dense in \mathcal{H} and $\bigvee V_K(\mathcal{S})$ is dense in \mathcal{H}_K , U_0 extends to a unitary map $U : \mathcal{H}_K \rightarrow \mathcal{H}$ mapping $V_K(x)$ to $V(x)$. \square

Examples.

1. Let \mathcal{S} be any set and let $K(x, y) := \delta_{x, y}$. Then $\mathcal{H} = l^2(\mathcal{S})$ and V maps the elements of \mathcal{S} to an orthonormal basis of \mathcal{H} .
2. Let $\mathcal{S} := \mathcal{H}_1 \times \mathcal{H}_2$, the Cartesian product of two Hilbert spaces \mathcal{H}_1 and \mathcal{H}_2 . Let

$$K((\psi_1, \psi_2), (\chi_1, \chi_2)) := \langle \psi_1, \chi_1 \rangle \cdot \langle \psi_2, \chi_2 \rangle.$$

Then $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, the tensor product of \mathcal{H}_1 and \mathcal{H}_2 , and $V(\psi_1, \psi_2) = \psi_1 \otimes \psi_2$.

3. Let \mathcal{S} be a Hilbert space; call it \mathcal{K} for the occasion. Let $K(\psi, \chi) := \langle \psi, \chi \rangle^2$. Then \mathcal{H} is the *symmetric* tensor product $\mathcal{K} \otimes_s \mathcal{K}$ and $V(\psi) = \psi \otimes \psi$.
4. Let $\mathcal{S} := \mathcal{K}$ as in example 3. Let K be the positive definite kernel

$$K(\psi, \chi) := e^{\langle \psi, \chi \rangle} .$$

Then the Kolmogorov dilation is the *Fock space* $\Gamma(\mathcal{K})$ over \mathcal{K} , defined as

$$\Gamma(\mathcal{K}) := \mathbb{C} \oplus \mathcal{K} \oplus \frac{1}{2}(\mathcal{K} \otimes_s \mathcal{K}) \oplus \frac{1}{6}(\mathcal{K} \otimes_s \mathcal{K} \otimes_s \mathcal{K}) \oplus \dots$$

and $V(\psi)$ is the so-called *exponential vector* or *coherent vector*

$$\text{Exp}(\psi) := 1 \oplus \psi \oplus (\psi \otimes \psi) \oplus (\psi \otimes \psi \otimes \psi) \oplus \dots$$

5. Let $\mathcal{S} = \mathbb{R}$ and let $K : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{C}$ be given by

$$K(s, t) := e^{-\gamma|s-t|} , \quad (\text{Re } \gamma > 0).$$

The Kolmogorov dilation of this kernel can be cast in the form

$$\mathcal{H} = L^2(\mathbb{R}, (2\text{Re } \gamma)dx) ; \quad V : t \mapsto v_t \in L^2(\mathbb{R}) : \quad v_t(x) := \begin{cases} e^{\gamma(x-t)} & \text{if } x \leq t; \\ 0 & \text{if } x > t. \end{cases}$$

3.3. Observables.

In quantum mechanics the notion of an *observable* is usually identified with that of a *self-adjoint* or *Hermitian* operator. Indeed such an operator has a good interpretation in terms of events.

The spectral theorem.

Every bounded operator A on a Hilbert space \mathcal{H} has an *adjoint* A^* defined by

$$\forall \vartheta, \psi \in \mathcal{H} : \quad \langle \vartheta, A^* \psi \rangle = \langle A \vartheta, \psi \rangle .$$

So a bounded operator is self-adjoint if it is *symmetric*, i.e.,

$$\forall \vartheta, \psi \in \mathcal{H} : \quad \langle \vartheta, A \psi \rangle = \langle A \vartheta, \psi \rangle .$$

However, many of the self-adjoint operators of physical interest are unbounded. Such operators can not be defined on all of \mathcal{H} , and therefore their domain of definition has to be taken into account.

Let A be a linear operator defined on a dense domain $\text{Dom}(A) \subset \mathcal{H}$. By the *adjoint* A^* of A we mean the operator with domain

$$\text{Dom}(A^*) := \{ \psi \in \mathcal{H} \mid \exists \psi^* \in \mathcal{H} \forall \vartheta \in \text{Dom}(A) : \langle \vartheta, \psi^* \rangle = \langle A \vartheta, \psi \rangle \}$$

which sends the vector ψ to the corresponding vector ψ^* .

Self-adjointness of the operator A not only means that A is symmetric, but moreover that $\text{Dom}(A^*) = \text{Dom}(A)$. In particular it implies that the graph

$$\text{Graph}(A) := \{ (\psi, A\psi) \in \mathcal{H} \times \mathcal{H} \mid \psi \in \text{Dom}(A) \}$$

is a closed set, the orthogonal complement of $\{ (A\psi, -\psi) \mid \psi \in \text{Dom}(A) \}$.

A self-adjoint operator defined on a dense domain in \mathcal{H} is loosely called a ‘self-adjoint operator on \mathcal{H} ’.

Definition. A *projection-valued measure* on a measure space (Ω, Σ) is a mapping E from Σ to the projections on some Hilbert space \mathcal{H} satisfying the following requirements:

- (i) $E(\Omega) = \mathbf{1}$;
- (ii) $A \cap B = \emptyset \implies E(A \cup B) = E(A) + E(B)$;
- (iii) $A_1 \subset A_2 \subset A_3 \subset \dots \implies E(\bigcup_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} E(A_n)$.

Theorem 3.3. (Spectral Theorem) *Let A be a self-adjoint operator on a Hilbert space \mathcal{H} . Then there exists a projection-valued measure E on \mathbb{R} such that*

$$A = \int_{-\infty}^{\infty} x E(dx) .$$

The integral is meant in the sense of matrix elements, i.e. for all $\psi \in \text{Dom}(A)$: $\langle \psi, A\psi \rangle = \int_{-\infty}^{\infty} x \langle \psi, E(dx)\psi \rangle$.

The operator A is said to be *affiliated* with the von Neumann algebra \mathcal{M} if $E(S) \in \mathcal{M}$ for all measurable subsets S of \mathbb{R} .

About the proof of the spectral theorem. This theorem is in fact a weak form of Gel'fand's theorem (Theorem 1.2). For the full proof we again refer to the literature. See for instance [ReS], [vRo2]. We only give a sketch here, making life (too!) easy by leaning on the strength of Gel'fand's theorem.

If A is bounded, A generates a commutative von Neumann algebra \mathcal{A} . Gel'fand's theorem then gives us a *-isomorphism between \mathcal{A} and the bounded measurable functions on some measure space (Ω, Σ) associating a function f_B to every $B \in \mathcal{A}$. In particular A itself gets sent to a function f_A . Now A and f_A have the same spectrum, so the values of f_A are the spectral points of A . But from Lemma 1.3 we may conclude that f_A is injective. So we may take $\Omega = \text{sp}(A)$ and $f_A(x) = x$. Now for every Borel subset S of $\text{sp}(A)$ let $E(S)$ be the projection for which $f_{E(S)} = 1_S$. Then E does the job.

If A is unbounded, we first map it to a unitary (and therefore bounded and normal) operator W by putting

$$W := (A + i)(A - i)^{-1} ,$$

(where we write i for $i \cdot \mathbf{1}$). This only works if both $A + i$ and $A - i$ are invertible. So we are done if we prove the following lemma.

Lemma 3.4. (The basic criterion for self-adjointness) *Let A be a symmetric operator whose domain is dense in \mathcal{H} and whose graph is closed in $\mathcal{H} \times \mathcal{H}$. Then A is self-adjoint if and only if $A + i$ and $A - i$ are invertible operators $\text{Dom}(A) \rightarrow \mathcal{H}$.*

Proof. Suppose that $A + i$ and $A - i$ are invertible. Then the null space \mathcal{N} of $A^* + i$ vanishes:

$$\mathcal{N} = ((A - i)\mathcal{H})^{\perp} = \mathcal{H}^{\perp} = \{0\} ,$$

so that $A^* + i$ is injective. But since A is symmetric, $A^* + i$ is an extension of the operator $A + i$. And since the latter is surjective, the extension must be a trivial one, i.e. $A^* + i = A + i$. It follows that $A^* = A$.

Conversely, suppose that $A^* = A$. Then for all $\psi \in \text{Dom}(A)$:

$$\|(A - i)\psi\|^2 = \|A\psi\|^2 + \|\psi\|^2 \geq \|\psi\|^2 . \quad (8)$$

So $A - i$ is injective. We claim that its range is *closed*. Indeed, suppose $\vartheta_n = (A - i)\psi_n$, ($n \in \mathbb{N}$) is a sequence in the range of $A - i$ tending to a limit ϑ . Then by (8) $\psi_1, \psi_2, \psi_3, \dots$ is Cauchy as well, say $\psi_n \rightarrow \psi$. So $(\psi_n, \vartheta_n)_{n \in \mathbb{N}}$ is a sequence in the graph of $A - i$ with limit (ψ, ϑ) . As this graph is closed, the point (ψ, ϑ) also lies on it, and it follows that $\vartheta = (A - i)\psi$. So the range of $A - i$ is closed.

To show that the range of $A - i$ is also *dense* in \mathcal{H} , suppose that $\chi \in \mathcal{H}$ is orthogonal to the range of $A - i$. Then for all $\psi \in \text{Dom}(A)$,

$$\langle \chi, (A - i)\psi \rangle = 0 \quad \text{hence} \quad \langle -i\chi, \psi \rangle = \langle \chi, A\psi \rangle .$$

So $\chi \in \text{Dom}(A^*)$ and $A^*\chi = -i\chi$; and since A is self-adjoint, $(A + i)\chi = 0$. But $A + i$ is injective just like $A - i$, so $\chi = 0$. We conclude that the range of $A - i$ is dense in \mathcal{H} . Since it is also closed, $A - i$ is surjective, and hence invertible. Replace i by $-i$ in the above to obtain a proof of the invertibility of $i + A$. \square

Functional calculus.

When Gel'fand's theorem associates a function $f = f_B : \text{sp}(A) \rightarrow \mathbb{C}$ to the operator B in the von Neumann algebra generated by A , we write

$$B = f(A) = j_A(f) = \int_{-\infty}^{\infty} f(x) E(dx) .$$

For instance, when f is the function $x \mapsto x^n$, then $f(A) = A^n$. In the usual approach to spectral theory this is the starting point. Here we have chosen to start from Gel'fand's theorem.

Interpretation of the spectral measure.

The spectral theorem gives for every measurable subset S of the spectrum of a self-adjoint operator A an orthogonal projection $E(S)$. We interpret $E(S)$ as the event

‘the random variable described by A takes a value in S ’

This event can be tested by measuring the observable and looking if the outcome lies in S .

So a self-adjoint operator A can be interpreted as a random variable with values in the spectrum of A , a subset of the real line.

Definition of a random variable.

Now what should be a random variable with values (say) on the unit sphere? In analogy with the case of real-valued random variables we want to have an event for every (measurable) subset of the sphere. But since points on the sphere cannot be added, this does not amount to any operator on \mathcal{H} . We do not bother about this, and we define:

Definition. By a *random variable* on a quantum probability space (\mathcal{A}, φ) with values in a measure space (Ω, Σ) we mean a projection-valued measure on $E : \Sigma \rightarrow \mathcal{A}$. The *probability distribution* \mathbb{P} of E is defined by

$$\mathbb{P}(S) := \varphi(E(S)) .$$

Alternatively we may describe the random variable by the operation j of integration with respect to this measure E , which is the inverse of the Gel'fand map $B \mapsto f_B$:

$$j : L^\infty(\Omega, \Sigma, \mathbb{P}) \rightarrow (\mathcal{A}, \varphi) : f \mapsto \int_{-\infty}^{\infty} f(x) E(dx) .$$

So we may define a random variable as an isometric *-homomorphism from a classical probability space into a quantum probability space.

Both definitions are equivalent.

Now we carry this idea one step further:

Definition. By a *generalised random variable* we mean an isometric *-homomorphism (an *embedding*) of one quantum probability space in another:

$$j : (\mathcal{B}, \psi) \rightarrow (\mathcal{A}, \varphi) .$$

3.4. The momentum observable in one dimension: an example.

Let $\mathcal{H} := L^2(\mathbb{R})$. We want to apply the spectral theorem to the differential operator

$$P : \psi \mapsto -i\psi' .$$

To do this, we must specify a domain and show that P is self-adjoint. Now the natural domain of P consists of those $\psi \in \mathcal{H}$ for which ψ' also lies in \mathcal{H} . However, this formulation is not precise enough. (What do we mean by ψ' when ψ is not continuous? Do we only want ψ to be differentiable almost everywhere? Etc.) Following van Rooij [vRo3] we take the converse approach.

Definition. We say that a function $\psi : \mathbb{R} \rightarrow \mathbb{C}$ is an *indefinite integral* of a function $\vartheta \in L^2(\mathbb{R})$ if for all $a, b \in \mathbb{R}$ with $a < b$:

$$\int_a^b \vartheta(x) dx = \psi(b) - \psi(a) .$$

We define the (one-dimensional) *momentum operator* P by

$$\begin{aligned} \text{Dom}(P) &:= \{ \psi \in L^2(\mathbb{R}) \mid \psi \text{ is indefinite integral of some } \vartheta \in L^2(\mathbb{R}) \} \\ P\psi &:= -i\vartheta . \end{aligned}$$

We note that indefinite integrals of L^2 -functions are continuous:

$$|\psi(b) - \psi(a)| = |\langle 1_{[a,b]}, \vartheta \rangle| \leq \sqrt{b-a} \cdot \|\vartheta\| .$$

Proposition 3.5. $P^* = P$.

From this proposition (and the spectral theorem) we may conclude that there exists a projection-valued measure $S \mapsto E(S)$ such that

$$P = \int_{-\infty}^{\infty} \lambda E(d\lambda) .$$

In fact, we know from Fourier theory that such a measure exists. Take

$$E(S) = \mathcal{F}M_{1_S}\mathcal{F}^{-1} ,$$

where \mathcal{F} is the Fourier transform. So in the example of the momentum operator the spectral theorem is no news. However, for many situations the existence of a spectral measure and a functional calculus is nontrivial and relevant. See Section 3.4 on Stone's Theorem.

Proof of the self-adjointness of P . According to Lemma 3.4 it suffices to show that

- (a) P is symmetric;
- (b) the graph of P is closed;
- (c) $P + i$ and $P - i$ are invertible operators $\text{Dom}(P) \rightarrow L^2(\mathbb{R})$.

It is convenient to write $P = -iD$, where $\text{Dom}(D) = \text{Dom}(P)$ and $D : \psi \mapsto \vartheta$.

(a): This amounts to the partial integration formula

$$\forall_{\psi, \vartheta \in \text{Dom}(D)} : \int_{-\infty}^{\infty} \overline{\psi} \cdot D\vartheta + \overline{D\psi} \cdot \vartheta = 0 ,$$

which regrettably we have to prove anew due to our unusual definition of the differentiation operator.

So let $a, b \in \mathbb{R}$ with $a < b$. Then

$$\begin{aligned} \int_a^b \overline{\psi} \cdot D\vartheta + \overline{D\psi} \cdot \vartheta &= \int_a^b \left(\overline{\psi(a)} + \int_a^y \overline{D\psi(x)} \right) D\vartheta(y) dy \\ &\quad + \int_a^b \overline{D\psi(x)} \left(\vartheta(b) - \int_x^b D\vartheta(y) dy \right) dx \\ &= \overline{\psi(a)}(\vartheta(b) - \vartheta(a)) + (\overline{\psi(b)} - \overline{\psi(a)})\vartheta(b) \\ &= \overline{\psi(b)}\vartheta(b) - \overline{\psi(a)}\vartheta(a) , \end{aligned}$$

where the double integrals cancel due to Fubini's theorem: they are both integrals of the L^2 -function $(x, y) \mapsto \overline{D\psi(x)} \cdot D\vartheta(y)$ over the triangle $\{(x, y) \mid a \leq x \leq y \leq b\}$. Now, since ψ , ϑ , $D\psi$ and $D\vartheta$ are square integrable, the limit of the right hand side as $a \rightarrow -\infty, b \rightarrow \infty$ must exist. But then it must be 0, since any other value would lead to non-integrable ψ and ϑ .

(b): Suppose that $\psi_1, \psi_2, \psi_3, \dots$ is a sequence in $\text{Dom}(D)$ such that $\psi_n \rightarrow \psi$ and $A\psi_n \rightarrow \vartheta$. We must show that $\psi \in \text{Dom}(D)$ and $D\psi = \vartheta$, i.e. for all $a < b$:

$$\int_a^b \vartheta(x) dx = \psi(b) - \psi(a) .$$

Now, since $\|\psi_n - \psi\| \rightarrow 0$, it is possible to choose a sequence n_1, n_2, n_3, \dots such that for almost all $x \in \mathbb{R}$ we have $\psi_{n_k}(x) \rightarrow \psi(x)$. In particular, for almost all $a, b \in \mathbb{R}$ with $a < b$:

$$\begin{aligned} \psi(b) - \psi(a) &= \lim_{k \rightarrow \infty} (\psi_{n_k}(b) - \psi_{n_k}(a)) = \lim_{k \rightarrow \infty} \langle 1_{[a,b]}, \vartheta_{n_k} \rangle \\ &= \langle 1_{[a,b]}, \vartheta \rangle = \int_a^b \vartheta(x) dx . \end{aligned}$$

Since ψ is continuous, this relation holds for *all* $a < b$. □

(c): We show that $\mathbf{1} - D : \text{Dom}(D) \rightarrow L^2(\mathbb{R})$ is invertible by constructing the inverse. For $\vartheta \in L^2(\mathbb{R})$, let

$$(T\vartheta)(x) := \int_{-\infty}^0 e^y \vartheta(x-y) dy .$$

Then by the Cauchy-Schwarz inequality in $L^2((-\infty, 0], e^y dy)$ we have

$$|(T\vartheta)(x)|^2 \leq \int_{-\infty}^0 e^y |\vartheta(x-y)|^2 dy .$$

So by Fubini's theorem,

$$\begin{aligned} \|T\vartheta\|^2 &= \int_{-\infty}^{\infty} |(T\vartheta)(x)|^2 dx \leq \int_{-\infty}^{\infty} \left(\int_{-\infty}^0 e^y |\vartheta(x-y)|^2 dy \right) dx \\ &= \int_{-\infty}^0 \left(\int_{-\infty}^{\infty} |\vartheta(x-y)|^2 dx \right) e^y dy = \|\vartheta\|^2 , \end{aligned}$$

i.e. T is a contraction. We shall show that T is the inverse we are looking for. Let $\vartheta \in L^2(\mathbb{R})$. We must prove:

$$(\mathbf{1} - D)T\vartheta = \vartheta .$$

Let

$$e_a(x) := \begin{cases} e^{a-x} & \text{for } x \geq a; \\ 0 & \text{for } x < a. \end{cases} .$$

Then it is a simple matter to check that

$$\int_a^b e_y dy = e_b - e_a + 1_{[a,b]} .$$

By Fubini's theorem it follows that, since $(T\vartheta)(y) = \langle e_y, \vartheta \rangle$,

$$\int_a^b T\vartheta(y) dy = T\vartheta(b) - T\vartheta(a) + \int_a^b \vartheta(y) dy .$$

In other words, $T\vartheta$ is an indefinite integral of $T\vartheta - \vartheta$:

$$DT\vartheta = T\vartheta - \vartheta ,$$

or briefly,

$$(\mathbf{1} - D)T\vartheta = \vartheta .$$

A similar calculation shows that $\mathbf{1} + D$ is invertible. □

3.5. One-parameter unitary groups.

Let A be a self-adjoint operator on a Hilbert space \mathcal{H} , and let for all $t \in \mathbb{R}$,

$$U_t := e^{itA} := \int_{-\infty}^{\infty} e^{it\lambda} E_A(d\lambda) .$$

Then the U_t are unitary operators satisfying

$$U_{t+s} = U_t U_s \quad \text{and} \quad \text{strong-}\lim_{s \rightarrow t} U_s = U_t .$$

Such a family of operators is called a *strongly continuous one-parameter unitary group*.

Theorem 3.6. (Stone's Theorem) *Let $(U_t)_{t \in \mathbb{R}}$ be a strongly continuous one-parameter unitary group on a Hilbert space \mathcal{H} . Then there is a self-adjoint operator A on \mathcal{H} such that $U_t = e^{itA}$.*

We omit the proof. (Cf. [ReS].)

So altogether we can describe a real-valued observable or random variable in four ways:

1. as a self-adjoint operator A ;
2. as a projection-valued measure $E : \Sigma(\mathbb{R}) \rightarrow \mathcal{A}$;
3. as an isometric *-homomorphism $j : L^\infty(\mathbb{R}, \Sigma(\mathbb{R}), \mathbb{P}) \rightarrow \mathcal{A}$;
4. and as a one-parameter group $t \mapsto U_t \in \mathcal{A}$.

The momentum operator on an interval..

Exercise. Let $\mathcal{H} := L^2([0, 1])$. Let $\text{Dom}(D)$ be the subspace of those functions $\psi \in \mathcal{H}$ that are the indefinite integral of some $\vartheta \in \mathcal{H}$ and that satisfy the boundary condition

$$\psi(0) = \psi(1) = 0 .$$

- (a) Show that $P := -iD$ is symmetric, but *not* self-adjoint.
- (b) Find the null spaces of $P^* + i$ and $P^* - i$.
- (c) Show that for every $\alpha \in \mathbb{C}$ with $|\alpha| = 1$ a self-adjoint extension P_α of P is obtained by weakening the boundary condition to

$$\psi(1) = \alpha\psi(0) .$$

- (d) Determine the unitary group e^{-itP_α} .

3.6. von Neumann algebras.

The double commutant theorem.

Let \mathcal{S} be a set of bounded operators on \mathcal{H} such that $X \in \mathcal{S} \implies X^* \in \mathcal{S}$. The algebra generated by $\mathbf{1}$ and \mathcal{S} we denote by $\text{alg}(\mathcal{S})$, the von Neumann algebra by $\text{vN}(\mathcal{S})$. Note that $\text{vN}(\mathcal{S})$ is the strong closure of $\text{alg}(\mathcal{S})$. The *commutant* \mathcal{S}' of \mathcal{S} is defined by

$$\mathcal{S}' := \{ Y \in \mathcal{B}(\mathcal{H}) \mid \forall X \in \mathcal{S} : XY = YX \} .$$

We note that the commutant is a von Neumann algebra. Indeed,

$$Y, Z \in \mathcal{S}' \implies \begin{cases} \lambda Y + \mu Z \in \mathcal{S}' , \\ YZ \in \mathcal{S}' , \\ Y^* \in \mathcal{S}' . \end{cases}$$

Also, if $Y_i \longrightarrow Y$ strongly with $Y_i \in \mathcal{S}'$, then for all $X \in \mathcal{S}$ and all $\psi \in \mathcal{H}$:

$$XY\psi = X \left(\lim_{i \rightarrow \infty} Y_i \psi \right) = \lim_{i \rightarrow \infty} XY_i \psi = \lim_{i \rightarrow \infty} Y_i X \psi = YX\psi .$$

So $Y \in \mathcal{S}'$, and \mathcal{S}' is a von Neumann algebra.

Double commutant theorem. *For any self-adjoint subset \mathcal{S} of $\mathcal{B}(\mathcal{H})$ we have*

$$\text{vN}(\mathcal{S}) = \mathcal{S}'' .$$

Proof. Clearly $\mathcal{S} \subset \mathcal{S}''$, and since \mathcal{S}'' is a von Neumann algebra, we have $\text{vN}(\mathcal{S}) \subset \mathcal{S}''$. We shall now prove the converse inclusion. Let $B \in \mathcal{S}''$, and let $\mathcal{A} := \text{alg}(\mathcal{S})$. We must show that B can be strongly approximated by elements of \mathcal{A} .

STEP 1. Choose $\psi \in \mathcal{H}$, and let P be the orthogonal projection onto $\overline{\mathcal{A}\psi}$. Then for all $X \in \mathcal{S}$ and $A \in \mathcal{A}$:

$$XPA\psi = XA\psi \in \mathcal{A}\psi \implies XPA\psi = PXA\psi .$$

So XP and PX coincide on the space $\overline{\mathcal{A}\psi}$. But if $\vartheta \perp \overline{\mathcal{A}\psi}$, then $P\vartheta = 0$ and for all $A \in \mathcal{A}$:

$$\langle X\vartheta, A\psi \rangle = \langle \vartheta, X^*A\psi \rangle = 0 ,$$

so $X\vartheta \perp \overline{\mathcal{A}\psi}$ as well. So $PX\vartheta = 0 = XP\vartheta$, and the operators XP and PX also coincide on the orthogonal complement of $\overline{\mathcal{A}\psi}$. We conclude that $XP = PX$, i.e. $P \in \mathcal{S}'$. But then we also have $BP = PB$, since $B \in \mathcal{S}''$. So

$$B\psi = BP\psi = PB\psi \in \overline{\mathcal{A}\psi} ,$$

and $B\psi$ can be approximated arbitrarily closely by vectors of the form $A\psi$ with $A \in \mathcal{A}$.

STEP 2. But this is not sufficient: we must show that $B\psi = \lim_{i \rightarrow \infty} A_i\psi$ for all $\psi \in \mathcal{H}$ at once! Equivalently, we must show that in all *strong environments* of B :

$$U_{(\psi_1, \dots, \psi_n)} := \{ X \in \mathcal{B}(\mathcal{H}) \mid \| (X - B)\psi_j \| < 1, \quad j = 1, \dots, n \}$$

there lies some element of \mathcal{A} .

So choose $\psi_1, \dots, \psi_n \in \mathcal{H}$. We define:

$$\begin{aligned}\tilde{\mathcal{H}} &:= \mathcal{H} \oplus \mathcal{H} \oplus \dots \oplus \mathcal{H} = \mathbb{C}^n \otimes \mathcal{H}, \\ \tilde{\mathcal{A}} &:= \{ A \oplus A \oplus \dots \oplus A \mid A \in \mathcal{A} \} = \mathcal{A} \otimes \mathbf{1}, \\ \tilde{\psi} &:= \psi_1 \oplus \psi_2 \oplus \dots \oplus \psi_n.\end{aligned}$$

Then $(\tilde{\mathcal{A}})' = (\mathcal{A} \otimes \mathbf{1})' = \mathcal{A}' \otimes M_n$ and $(\tilde{\mathcal{A}})'' = (\mathcal{A}' \otimes M_n)' = \mathcal{A}'' \otimes \mathbf{1}$. So $B \otimes \mathbf{1} \in (\tilde{\mathcal{A}})''$. By step 1 we find an element \tilde{A} of $\tilde{\mathcal{A}}$, such that

$$\left\| \left(\tilde{A} - (B \otimes \mathbf{1}) \right) \tilde{\psi} \right\| < 1.$$

But $\tilde{A} \in \tilde{\mathcal{A}}$ must be of the form $A \otimes \mathbf{1}$ with $A \in \mathcal{A}$, so

$$\sum_{j=1}^n \|(A - B)\psi_j\|^2 = \left\| ((A \otimes \mathbf{1}) - (B \otimes \mathbf{1})) \tilde{\psi} \right\|^2 < 1.$$

Hence $\|A\psi_j - B\psi_j\| < 1$ for each j , in other words:

$$A \in U_{(\psi_1, \dots, \psi_n)}.$$

□

What kinds of von Neumann algebras are there?.

Von Neumann algebras \mathcal{A} and \mathcal{B} on Hilbert spaces \mathcal{H} and \mathcal{K} are called *isomorphic* if there exists a linear bijection

$$i : \mathcal{A} \rightarrow \mathcal{B}$$

which preserves product and conjugation. They are called *spatially isomorphic* or *unitarily equivalent* if moreover there exists a unitary map

$$U : \mathcal{H} \rightarrow \mathcal{K} \quad \text{such that} \quad \forall_{A \in \mathcal{A}} : \quad i(A) = UAU^{-1}.$$

Isomorphic algebras which are not spatially isomorphic are considered as different *representations* of the same algebra.

For example, the algebra M_n of all $n \times n$ -matrices can be represented on \mathbb{C}^n in the natural way, but also on all Hilbert spaces of the form

$$\mathcal{H} = \mathcal{L} \otimes \mathbb{C}^n,$$

by acting trivially on \mathcal{L} . In particular, we may put $\mathcal{L} = \mathbb{C}^n$ and obtain the so-called *tracial* representation of M_n . This name refers to the fact that there is a unit vector $\psi \in \mathcal{H}$ with the property that

$$\forall_{A, B \in M_n} : \quad \langle \psi, AB\psi \rangle = \langle \psi, BA\psi \rangle.$$

Indeed, this *trace vector* ψ is given by

$$\psi = \frac{1}{\sqrt{n}} \sum_{j=1}^n e_j \otimes e_j,$$

and we have that $\langle \psi, A\psi \rangle = \frac{1}{n} \text{tr}(A)$.

Finite dimensional von Neumann algebras.

If \mathcal{A} is of finite dimension, then it can be written in the form

$$\mathcal{A} = M_{n_1} \oplus M_{n_2} \oplus \cdots \oplus M_{n_k} . \quad (9)$$

It can be represented on Hilbert spaces of the form

$$\mathcal{H} = (\mathbb{C}_{m_1} \otimes \mathbb{C}_{n_1}) \oplus (\mathbb{C}_{m_2} \otimes \mathbb{C}_{n_2}) \oplus \cdots \oplus (\mathbb{C}_{m_k} \otimes \mathbb{C}_{n_k}) ,$$

by acting on the second space in the tensor product each time. The commutant \mathcal{A}' in this representation is isomorphic to

$$M_{m_1} \oplus M_{m_2} \oplus \cdots \oplus M_{m_k} .$$

The *centre* \mathcal{Z} of \mathcal{A} is the abelian von Neumann algebra defined as

$$\mathcal{Z} := \mathcal{A} \cap \mathcal{A}' ,$$

which in the present case is given by

$$\mathcal{Z} = \{ c_1 \cdot \mathbf{1} \oplus c_2 \cdot \mathbf{1} \oplus \cdots \oplus c_k \cdot \mathbf{1} \mid c_1, \dots, c_k \in \mathbb{C} \} .$$

If $\mathcal{Z} = \mathbb{C} \cdot \mathbf{1}$ then \mathcal{A} is called a *factor*. Every finite dimensional factor is equal to M_n for some $n \in \mathbb{N}$.

A vector $\psi \in \mathcal{H}$ is called *cyclic* for \mathcal{A} if every vector in \mathcal{H} can be written as $A\psi$ for some $A \in \mathcal{A}$. (In the infinite-dimensional case: if it can be *approximated* by such vectors.) There is a cyclic trace vector in \mathcal{H} if and only if $m_j = n_j$ for $j = 1, \dots, k$. So in the trace representation the algebra and its commutant are isomorphic.

General von Neumann algebras.

Every von Neumann algebra can be decomposed into factors like in (9), but the decomposition is generally not a sum, but an integral. Also, the factors themselves may be of very different types.

Type I: All bounded operators on some Hilbert space \mathcal{H} . If $\mathcal{H} = \mathbb{C}^n$, this is called type I_n ; if \mathcal{H} is infinite dimensional, type I_∞ .

Type II: A factor \mathcal{M} is of type II_1 if it is infinite dimensional, yet every isometry in \mathcal{M} is unitary. \mathcal{M} is of type II_∞ if it is not of type II_1 , but contains projections P such that PMP is of type II_1 .

Type III: None of the above, but still a factor. In contrast to the types I and II, on a factor of type III there is no trace.

Example of a factor of type II_1 :

$$\mathcal{H} := L^2 \left(\{-1, 1\}^{\mathbb{N}}, \bigotimes_{\mathbb{N}} \left(\frac{1}{2}, \frac{1}{2} \right) \right) ,$$

\mathcal{A} is generated by $\mathbf{1} \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1} \otimes M_2 \otimes \mathbf{1} \otimes \cdots$

Here \mathcal{H} is the L^2 -space of an infinite sequence of tosses of a fair coin.

Example of a factor of type III: take a biased coin.

4. Some Quantum Mechanics.

Quantum mechanics are a physical theory that fits in the framework of non-commutative probability, but which has much more structure. It deals with particles and fields, using observables like position, momentum, angular momentum, energy, charge, spin, isospin, etc. All these observables develop in time according to a certain dynamical rule, the Schrödinger equation.

In this section we shall pick out a few elements of this theory that are of particular interest to our main example: the damped harmonic oscillator as a quantum Markov chain.

4.1. Position and momentum.

Let us start with a simple example: a particle on a line. This particle must have a *position* observable, a projection valued measure on the Borel σ -algebra $\Sigma(\mathbb{R})$ of the real line \mathbb{R} :

$$E : \Sigma(\mathbb{R}) \rightarrow \mathcal{B}(\mathcal{H}) .$$

The easiest choice (when the particle is alone in the world and has no further degrees of freedom) is

$$\mathcal{H} := L^2(\mathbb{R}) ;$$

$$E(S) : \psi \mapsto 1_S \cdot \psi .$$

In this example the Hilbert space \mathcal{H} naturally carries a second real-valued random variable in the form of the group $(T_t)_{t \in \mathbb{R}}$ of the spatial translations:

$$(T_t \psi)(x) := \psi(x - \hbar t) , \tag{10}$$

according to the remark made in Section 3.5. This second observable is called the *momentum* of the particle. The associated self-adjoint operators are Q and P given by

$$\begin{aligned} (Q\psi)(x) &= x\psi(x) ; \\ (P\psi)(x) &= -i\hbar \frac{\partial}{\partial x} \psi(x) . \end{aligned}$$

Just like we have $T_t = e^{-itP}$, it is natural to introduce $S_s := e^{isQ}$ whose action on \mathcal{H} is given by

$$S_s \psi(x) := e^{isx} \psi(x) . \tag{11}$$

The operators P and Q satisfy Heisenberg's *canonical commutation relation* (CCR)

$$[P, Q] = -i\hbar \cdot \mathbf{1} . \tag{12}$$

A pair of self-adjoint operators (P, Q) satisfying (12) is called a *canonical pair*.

Representations of the canonical commutation relations.

What kinds of canonical pairs do there exist?

Before this question can be answered, it has to be reformulated. Relation (12) is not satisfactory as a definition of a canonical pair since the domains on the left and on the right are not the same. Worse than that, quite pathological examples can be constructed, even if (12) is postulated to hold on a dense stable domain, with the property that P and Q admit only unique self-adjoint extensions [ReS].

In order to circumvent domain questions, Hermann Weyl proposed to replace (12) by a relation between the associated unitary groups T_t and S_s , namely:

$$T_t S_s = e^{-i\hbar st} S_s T_t, \quad (s, t \in \mathbb{R}). \quad (13)$$

It was von Neumann's idea to mix the two unitary representations (T_t) and (S_s) of \mathbb{R} into a two-parameter family

$$W(t, s) := e^{\frac{i\hbar}{2} st} T_t S_s, \quad (14)$$

forming a 'twisted' representation of \mathbb{R}^2 , as expressed by the 'Weyl relation': for all $s, t, u, v \in \mathbb{R}$,

$$W(t, s)W(u, v) = e^{-\frac{i\hbar}{2}(tv-su)} W(t+u, s+v). \quad (15)$$

This relation captures the group property of T_t and S_s together with the relation (13). Formally,

$$W(t, s) = e^{i(sQ-tP)}.$$

We shall call the representation on $L^2(\mathbb{R})$ of the CCR given by (10), (11) and (14) the *Schrödinger representation* of the CCR.

Here and in the rest of the text we shall follow the Quantum Probabilist's convention that

$$\hbar = 2.$$

Theorem 2.1. (von Neumann's Uniqueness Theorem) *Let $(W(t, s))_{t, s \in \mathbb{R}}$ be a strongly continuous family of unitary operators on some Hilbert space \mathcal{H} satisfying the Weyl relation (15). Then \mathcal{H} is unitarily equivalent with $L^2(\mathbb{R}) \otimes \mathcal{K}$, such that $W(t, s)$ corresponds to $W_S(t, s) \otimes \mathbf{1}$, where W_S is the Schrödinger representation (14) of the CCR.*

Proof. Let $W : \mathbb{R}^2 \rightarrow \mathcal{U}(\mathcal{H})$ satisfy the Weyl relation (15). For $f : \mathbb{R}^2 \rightarrow \mathbb{C}$ with $\int \int |f(t, s)| dt ds < \infty$, define a bounded operator $A(f)$ on \mathcal{H} by

$$A(f) := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t, s) W(t, s) dt ds.$$

We find the following calculating rules for such operators $A(f)$ and their *kernels* f :

$$\begin{aligned} A(f) + A(g) &= A(f + g); \\ A(f)^* &= A(\tilde{f}), \quad \text{where } \tilde{f}(t, s) := \overline{f(-t, -s)}; \\ A(f)A(g) &= A(f * g). \end{aligned}$$

Here the ‘twisted convolution product’ $*$ is defined by

$$(f * g)(t, s) := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i(tv-su)} f(t-u, s-v)g(u, v)du dv .$$

Moreover we claim that an operator (on a nontrivial Hilbert space) can have at most one kernel:

$$A(f) = 0 \implies \mathcal{H} = \{0\} \text{ or } f = 0 . \quad (16)$$

Indeed, if $A(f) = 0$ then we have for all $a, b \in \mathbb{R}$,

$$0 = W(a, b)^* A(f) W(a, b) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{2i(as-bt)} f(t, s) W(t, s) dt ds .$$

Taking matrix elements, we find that for all $\psi, \vartheta \in \mathcal{H}$ the function

$$(t, s) \mapsto f(t, s) \langle \varphi, W(t, s) \psi \rangle$$

has Fourier transform 0. So either $W(t, s) = 0$ (i.e. $\mathcal{H} = \{0\}$), or $f(t, s) = 0$ almost everywhere.

The key idea to the proof of uniqueness is to consider the operator

$$E := \frac{1}{\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(t^2+s^2)} W(t, s) dt ds .$$

This operator has the remarkable property that for all $a, b \in \mathbb{R}$, $EW(a, b)E$ differs from E only by a scalar factor:

$$EW(a, b)E = e^{-\frac{1}{2}(a^2+b^2)} E . \quad (17)$$

Indeed, E has kernel $g(t, s) := \frac{1}{\pi} e^{-\frac{1}{2}(t^2+s^2)}$, and the product $W(a, b)E$ has kernel

$$h(u, v) := \frac{1}{\pi} e^{-i(av-bu)} \cdot e^{-\frac{1}{2}((a-u)^2+(b-v)^2)} .$$

So $EW(a, b)E$ has kernel

$$\begin{aligned} (g * h)(t, s) &= \int \int e^{-i(tv-su)} g(t-u, s-v) h(u, v) du dv \\ &= \frac{1}{\pi^2} \int \int e^{-i(tv-su)} e^{-\frac{1}{2}((t-u)^2+(s-v)^2)} e^{-i(av-bu)} e^{-\frac{1}{2}((a-u)^2+(b-v)^2)} du dv \\ &= \frac{1}{\pi^2} e^{-\frac{1}{2}(a^2+b^2)} e^{-\frac{1}{2}(t^2+s^2)} \int \int e^{-(u^2+v^2)} e^{(u-iv)(t+is+a+ib)} du dv \\ &= \frac{1}{\pi^2} e^{-\frac{1}{2}(a^2+b^2)} e^{-\frac{1}{2}(t^2+s^2)} \int \int e^{-(u-\frac{1}{2}(t+is+a+ib))^2 - (v-\frac{1}{2}(t+is+a+ib))^2} du dv \\ &= \frac{1}{\pi} e^{-\frac{1}{2}(a^2+b^2)} \cdot e^{-\frac{1}{2}(t^2+s^2)} \\ &= e^{-\frac{1}{2}(a^2+b^2)} g(t, s) , \end{aligned}$$

which proves (17).

We conclude that $E^* = E$ (since $\tilde{g} = g$), $E^2 = E$ (putting $a = b = 0$ in (17)), and that $E\mathcal{A}E = \mathbb{C}E$, where \mathcal{A} is the von Neumann algebra generated by the Weyl operators. So E is a minimal projection in \mathcal{A} . Denote its range by \mathcal{K} . Then we have for all $\psi, \vartheta \in \mathcal{K}$ and all $t, s, u, v \in \mathbb{R}$:

$$\begin{aligned} \langle W(t, s)\psi, W(u, v)\vartheta \rangle &= \langle W(t, s)E\psi, W(u, v)E\vartheta \rangle \\ &= \langle \psi, EW(-t, -s)W(u, v)E\vartheta \rangle \\ &= e^{i(tv-su)} \langle \psi, EW(u-t, v-s)E\vartheta \rangle \\ &= e^{i(tv-su)} e^{-\frac{1}{2}((u-t)^2+(v-s)^2)} \langle \psi, E\vartheta \rangle \\ &= e^{(t-is)(u+iv)} e^{-\frac{1}{2}(t^2+s^2+u^2+v^2)} \langle \psi, \vartheta \rangle, \end{aligned}$$

Therefore the map

$$V : \mathbb{R}^2 \times \mathcal{K} \rightarrow \mathcal{H} : ((t, s), \psi) \mapsto e^{\frac{1}{2}(t^2+s^2)} W(t, s)\psi$$

is a Kolmogorov dilation of the positive definite kernel (cf. Section 3.2.)

$$K : (\mathbb{R}^2 \times \mathcal{K}) \times (\mathbb{R}^2 \times \mathcal{K}) \rightarrow \mathbb{C} : ((t, s), \psi; (u, v), \vartheta) \mapsto e^{(t-is)(u+iv)} \langle \psi, \vartheta \rangle. \quad (18)$$

Since the Schrödinger representation of the CCR is irreducible, the associated minimal projection E_S must be one-dimensional. By explicit calculation we find that E_S is the orthogonal projection onto the one-dimensional subspace spanned by the unit vector $\Omega(x) := \sqrt{\gamma(x)}$, where γ is the density of the standard normal distribution:

$$\gamma(x) := \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}.$$

So in the Schrödinger case the dilation is

$$V_S : \mathbb{R}^2 \rightarrow L^2(\mathbb{R}) : (t, s) \mapsto e^{\frac{1}{2}(t^2+s^2)} e^{-\frac{i}{2}ts} e^{isx} \Omega(x - 2t).$$

By Kolmogorov's Dilation Theorem (Section 3.2.), if the linear span of the range of V is dense in \mathcal{H} (and in particular the range of V_S is dense in $L^2(\mathbb{R})$), then there exists a unitary equivalence $U : L^2(\mathbb{R}) \otimes \mathcal{K} \rightarrow \mathcal{H}$ such that for all $a, b \in \mathbb{R}$ and $\psi \in \mathcal{K}$:

$$U(W_S(a, b)\Omega \otimes \psi) = W(a, b)\psi,$$

connecting our operators $W(a, b)$ to the Schrödinger representation $W_S(a, b)$ of the CCR. It is then easy to check that for all $a, b \in \mathbb{R}$ and $\psi \in \mathcal{K}$:

$$W(a, b) = U(W_S(a, b) \otimes \mathbf{1})U^{-1}.$$

So it remains to show that the range of V is dense in \mathcal{H} . Let \mathcal{L} denote the orthogonal complement of this range. Then \mathcal{L} is invariant for the Weyl operators; let $W_0(t, s)$ be the restriction of $W(t, s)$ to \mathcal{L} . Construct $E_0 := A_0(g)$ in terms of W_0 in the same way as E was constructed from W . Then clearly $E_0 \leq E$, but also $E_0 \perp E$. So $E_0 = A_0(g) = 0$ and by (16) we have $\mathcal{L} = \{0\}$. \square

Exercise. Calculate the minimal projection E_S in the Schrödinger representation.

4.2. Energy and time evolution.

The evolution in time of a closed quantum system is given by a pointwise strongly continuous one-parameter group $(\alpha_t)_{t \in \mathbb{R}}$ of $*$ -automorphisms of the observable algebra \mathcal{A} .

Like in the case of a particle on a line, in the case of a finite number n of (distinguishable) particles in d -dimensional space we take $\mathcal{A} = \mathcal{B}(\mathcal{H})$ with $\mathcal{H} = L^2(\mathbb{R}^{nd})$. Since all automorphisms of this algebra are implemented by unitary transformations of \mathcal{H} , our group (α_t) is of the form

$$\alpha_t(A) = U_t A U_t^{-1} ,$$

where $t \mapsto U_t$ is strongly continuous $\mathbb{R} \rightarrow \mathcal{U}(\mathcal{H})$. We denote the generator by H/\hbar :

$$U_t = e^{itH/\hbar} .$$

The self-adjoint operator H corresponds to an observable of the system of particles, called its *energy*. The operator H itself is known as the *Hamilton operator* or *Hamiltonian* of the system. As the Hamiltonian always commutes with the time evolution operator, energy is a conserved quantity:

$$\alpha_t(H) = U_t H U_t^{-1} = H .$$

The *dynamical law* (as it was called by Hermann Weyl [Wey]), is an equation which expresses the Hamiltonian in terms of other observables. This equation depends on the nature of the interaction between our particles. In the absence of magnetic fields the dynamical law takes the form

$$H = \sum_{j=1}^{nd} \frac{1}{2m_{k(j)}} P_j^2 + V(Q_1, Q_2, \dots, Q_{nd})$$

for some function $V : (\mathbb{R}^d)^n \rightarrow \mathbb{R}$, called the *potential*. The positive constants m_k , $k = 1, \dots, n$ are the *masses* of the particles. (We put $k(j) := 1 + [(j-1)/d]$ in order to attach the same mass to the coordinates of the same particle.)

Free particles.

If $V = 0$, then U_t factorises into a tensor product of nd one-dimensional evolution operators, all of the form

$$U_t = e^{itH/\hbar} = e^{i\frac{t}{2m\hbar} P^2} .$$

Since the Hamiltonian $H = P^2/2m$ now commutes with P , momentum is conserved:

$$\alpha_t(P) = P .$$

On a formal level the time development of the operator Q is easily found by solving the differential equation

$$\frac{d}{dt}\alpha_t(Q) = \frac{d}{dt}U_tQU_t^{-1} = \frac{i}{2m\hbar}[P^2, \alpha_t(Q)].$$

A solution is

$$\alpha_t(Q) = Q + \frac{t}{m}P.$$

According to the uniqueness theorem the canonical pairs (P, Q) and $(P, Q + \frac{t}{m}P)$ are indeed unitarily equivalent. So we expect that the evolution of the Weyl operators will be the following:

$$\begin{aligned}\alpha_t(W(x, y)) &= \alpha_t(e^{-ixP+iyQ}) = e^{-ixP+iy(Q+\frac{t}{m}P)} \\ &= e^{-i(x-\frac{t}{m}y)P+iyQ} = W\left(x - \frac{t}{m}y, y\right).\end{aligned}$$

Proposition 2.2. Let $P := -i\hbar\frac{\partial}{\partial x}$ denote the momentum operator on $\mathcal{H} := L^2(\mathbb{R})$, and let $W : \mathbb{R}^2 \rightarrow \mathcal{U}(\mathcal{H})$ be given by (15). Let

$$U_t := e^{i\frac{t}{2m\hbar}P^2}.$$

Then

$$U_tW(x, y)U_t^{-1} = W\left(x - \frac{t}{m}y, y\right).$$

Proof. From the definitions of T_t and E_Q it follows that for all measurable sets $B \subset \mathbb{R}$ and all $t \in \mathbb{R}$:

$$T_tE(B)T_t^{-1} = E(B + \hbar t).$$

From the uniqueness theorem it follows that irreducible representations of the CCR have the symmetry $Q \rightarrow P, P \rightarrow -Q$. So we also have the exchanged imprimitivity relation

$$\forall_{B \in \Sigma(\mathbb{R})} \forall_{y \in \mathbb{R}} : S_yE_P(B)S_y^{-1} = E_P(B + \hbar y).$$

It follows that for all $y, t \in \mathbb{R}$,

$$\begin{aligned}S_yU_{-t}S_y^{-1} &= S_y\left(\int_{-\infty}^{\infty} e^{-i\frac{t}{2m\hbar}\lambda^2} E_P(d\lambda)\right)S_y^{-1} \\ &= \int_{-\infty}^{\infty} e^{-i\frac{t}{2m\hbar}(\lambda - \hbar y)^2} E_P(d\lambda) \\ &= U_{-t} \cdot T_{-\frac{t}{m}y} \cdot e^{-i\frac{\hbar t}{2m}y^2}.\end{aligned}$$

Multiplying by U_t on the left and by S_y on the right we find

$$U_tW(0, y)U_t^{-1} = U_tS_yU_t^{-1} = e^{-i\frac{\hbar t}{2m}y^2}T_{-\frac{t}{m}y}S_y = W\left(-\frac{t}{m}y, y\right).$$

As T_x commutes with U_t we may freely add $(x, 0)$ to the argument of W , and the proposition is proved. \square

By imposing some state φ on $\mathcal{A} = \mathcal{B}(L^2(\mathbb{R}))$, all stochastic information on the model $(\mathcal{A}, \varphi, \alpha_t)$ can be obtained from the simple evolution equation $\alpha_t(Q) = Q + \frac{t}{m}P$. For example, for large times the random variable $\frac{1}{t}\alpha_t(Q)$ approaches $\frac{1}{m}P$ in distribution, provided that φ does not favour large Q values too much. So a position measurement can serve as a measurement of momentum, despite their well-known relative uncertainty.

The Schrödinger picture and the Schrödinger equation.

The above type of description of a system, namely with random variables moving in time, and the state φ given once and for all, is called the *Heisenberg picture* in the physical literature. In probability theory this is the general custom, and we take it over in quantum probability.

However, physicists often think in another picture, where the state is supposed to move, and the operators are fixed objects. This is called the *Schrödinger picture* of quantum mechanics.

In particular, if we take for φ a pure (i.e. extremal) state on the algebra $\mathcal{A} = \mathcal{B}(\mathcal{H})$ where \mathcal{H} is, say, $L^2(\mathbb{R}^{nd})$:

$$\varphi(A) = \langle \psi, A\psi \rangle, \quad (\psi \in \mathcal{H}, \|\psi\| = 1),$$

then we can express all probabilities at later times t in terms of the *wave function*

$$\psi(x_1, \dots, x_{nd}; t) := (U_t^{-1}\psi)(x_1, \dots, x_{nd}).$$

This wave function satisfies the *Schrödinger equation*, a partial differential equation:

$$\begin{aligned} & -i\hbar \frac{\partial}{\partial t} \psi(x_1, \dots, x_{nd}; t) \\ &= \sum_{j=1}^{nd} -\frac{1}{2m_{k(j)}\hbar^2} \frac{\partial^2}{\partial x_j^2} \psi(x_1, \dots, x_{nd}; t) + V(x_1, \dots, x_{nd})\psi(x_1, \dots, x_{nd}; t). \end{aligned}$$

If E is an orthogonal projection in \mathcal{H} , then the probability of the associated event can be calculated in the Schrödinger picture by

$$\begin{aligned} \varphi(\alpha_t(E)) &= \langle \psi, U_t E U_t^{-1} \psi \rangle = \langle U_t^{-1} \psi, E U_t^{-1} \psi \rangle \\ &= \int_{\mathbb{R}^{nd}} \overline{\psi_t(x_1, \dots, x_{nd})} (E\psi_t)(x_1, \dots, x_{nd}) dx_1, \dots, dx_{nd}. \end{aligned}$$

4.3. The harmonic oscillator.

A harmonic oscillator is a canonical pair of observables that under time evolution $(\alpha_t)_{t \in \mathbb{R}}$ performs a rotation such as

$$\begin{aligned} \alpha_t(Q) &= Q \cos t + P \sin t; \\ \alpha_t(P) &= -Q \sin t + P \cos t. \end{aligned}$$

Since rotation in the plane is symplectic (has determinant 1), this evolution respects the canonical commutation relation $QP - PQ = i\hbar \cdot \mathbf{1}$, so by the uniqueness theorem it determines (up to a time-dependent phase) a group of unitary transformations $(U_t)_{t \in \mathbb{R}}$ of the Hilbert space on which it is represented. (In particular, $U_{\frac{\pi}{2}}$ is a unitary transformation of $L^2(\mathbb{R})$ that sends Q into P and P into $-Q$: this is the Fourier transformation.)

Making a formal calculation again, differentiating the equality

$$\alpha_t(A) = e^{itH/\hbar} A e^{-itH/\hbar}$$

we find that a Hamiltonian of the form

$$H = \frac{1}{2}(P^2 + Q^2) \tag{19}$$

can be expected to generate such a rotating evolution.

The standard treatment of the harmonic oscillator follows the elegant algebraic reasoning of Dirac, who factorised the Hamiltonian (19) as

$$H = \frac{1}{2}(Q - iP)(Q + iP) + \frac{1}{2}i[P, Q] =: \hbar a^* a + \frac{1}{2}\hbar \cdot \mathbf{1} .$$

The operators a and a^* are then seen to lower and raise the eigenvalue of H , and are called the *annihilation* and *creation* operators.

Here we choose to proceed more analytically, seizing the opportunity to introduce techniques which will be useful again later on for the treatment of free quantum fields and the damped oscillator.

Our goal is to describe H and U_t explicitly.

Heisenberg's matrix representation.

First we note that, since α_t has period 2π , the differences between points in the spectrum of H must be multiples of \hbar . On the grounds of (19) we suspect that H is bounded from below, so let us try

$$\text{sp}(H) = \hbar\mathbb{N} + c .$$

We take as our Hilbert space $\mathcal{H}_H := l^2(\mathbb{N}, \frac{1}{n!})$ with the Hamiltonian given by

$$(H\vartheta)(n) = (\hbar n + c)\vartheta(n) .$$

The subscript ' H ' indicates that on this space we wish to stage matrix mechanics of the Heisenberg type. If we define on \mathcal{H}_H the 'product' or 'coherent' vectors

$$\pi(z) := (1, z, z^2, z^3, \dots), \quad (z \in \mathbb{C}) ,$$

then our intended time evolution takes the simple form

$$U_t^H \pi(z) = e^{itc/\hbar} \pi(e^{it} z) . \tag{20}$$

Now we want to represent a canonical pair (P, Q) in this space, or equivalently, Weyl operators $W(z)$, that rotate in the same way: $U_t W(z) U_t^{-1} = W(e^{it} z)$. We note that

$$\langle \pi(u), \pi(v) \rangle = \sum_{n=0}^{\infty} \frac{\bar{u}^n v^n}{n!} = e^{\bar{u}v} ,$$

so that we have here another dilation of the positive definite kernel (18) in the proof of the Uniqueness Theorem. An irreducible representation of the CCR is close at hand. Put:

$$W_H(z)\pi(u) = e^{-\bar{z}u - \frac{1}{2}|z|^2} \pi(u+z), \quad (z, u \in \mathbb{C}).$$

These operators satisfy the Weyl relation

$$W_H(w)W_H(z) = e^{-i\text{Im}(\bar{w}z)} W_H(w+z), \quad (21)$$

the same as (15) if we identify $W(t, s)$ with $W_H(t+is)$. Clearly we have also obtained

$$U_t W(z) U_t^{-1} = W(e^{it}z). \quad (22)$$

Let us summarise.

Proposition 2.3. *The Heisenberg representation of the Harmonic oscillator is given by*

$$\begin{aligned} \mathcal{H}_H &= l^2(\mathbb{N}, \frac{1}{n!}); \\ (H_H \vartheta)(n) &= (2n+1)\vartheta(n); \quad U_t^H \pi(z) = e^{\frac{i}{2}t} \pi(e^{it}z); \\ W_H(z)\pi(u) &= e^{-\bar{z}u - \frac{1}{2}|z|^2} \pi(u+z). \end{aligned}$$

In concrete terms, we have on an orthonormal basis,

$$Q = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \dots \\ 1 & 0 & \sqrt{2} & 0 & 0 & \dots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & 0 & \dots \\ 0 & 0 & \sqrt{3} & 0 & \sqrt{4} & \dots \\ 0 & 0 & 0 & \sqrt{4} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad P = \frac{1}{i} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \dots \\ -1 & 0 & \sqrt{2} & 0 & 0 & \dots \\ 0 & -\sqrt{2} & 0 & \sqrt{3} & 0 & \dots \\ 0 & 0 & -\sqrt{3} & 0 & \sqrt{4} & \dots \\ 0 & 0 & 0 & -\sqrt{4} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

These matrices satisfy

$$QP - PQ = 2i \cdot \mathbf{1} \quad \text{and} \quad \frac{1}{2}(Q^2 + P^2) = H,$$

where

$$H = \begin{pmatrix} 1 & & & & & \\ & 3 & & & & \\ & & 5 & & & \\ & & & 7 & & \\ & & & & 9 & \\ & & & & & \ddots \end{pmatrix}.$$

Proof. It only remains to check the matrices for Q and P . We note that

$$e^{iyQ} \pi(u) = W_H(iy)\pi(u) = e^{iyu - \frac{1}{2}y^2} \pi(u+iy),$$

and we find by differentiation

$$Q\pi(u) = u\pi(u) + \pi'(u).$$

Taking the coefficient of u^n the matrix of Q is found. The matrix for P is found in the same way. The choice of the ground state energy $c = \frac{1}{2}\hbar = 1$ in the definition of H fixes the relation with Q and P correctly. \square

The Gaussian representation.

Here is another useful representation of the harmonic oscillator algebra on a Hilbert space.

Let $\mathcal{H}_G := L^2(\mathbb{R}, \gamma)$, where γ is the standard Gauss measure on \mathbb{R} :

$$\gamma(dx) := \gamma(x)dx := \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx .$$

Define for $z \in \mathbb{C}$ the vector $\varepsilon(z)$ by

$$\varepsilon(z) : x \mapsto e^{zx - \frac{1}{2}z^2} .$$

Then $\varepsilon(z)$ with $z \in \mathbb{C}$ is a total set in \mathcal{H}_G . (Actually, $z \in i\mathbb{R}$ is already sufficient by the uniqueness of the Fourier transform.) Again we find

$$\langle \varepsilon(z), \varepsilon(u) \rangle = e^{\bar{z}u} , \quad (z, u \in \mathbb{C}) .$$

Proposition 2.4. *There exists a unitary map $U_{HG} : \mathcal{H}_H \rightarrow \mathcal{H}_G$ such that for all $z \in \mathbb{C}$*

$$U_{HG} \pi(z) = \varepsilon(z) .$$

This map sends the vector $e_n := (0, \dots, 0, 1, 0, \dots)$ into the n -th Hermite polynomial, where these polynomials are given by the generating function

$$\sum_{n=0}^{\infty} z^n h_n(x) = e^{zx - \frac{1}{2}z^2} .$$

Consequently, this version of the Hermite polynomials satisfies

$$\int_{-\infty}^{\infty} h_n(x) h_m(x) \gamma(dx) = \frac{1}{n!} \delta_{nm} .$$

Proof. The map $\pi(z) \mapsto \varepsilon(z)$ extends to a unitary map since the linear spans of the ranges of π and ε are dense and both π and ε are dilations of the positive definite kernel $(z, u) \mapsto e^{\bar{z}u}$. \square

Let us carry over some more operators with this unitary transformation. We find:

$$(e^{isQ_G} \psi)(x) = e^{isx} \psi(x), \quad (Q_G \psi)(x) = x\psi(x) ;$$

$$(e^{-itP_G} \psi)(x) = \psi(x - 2t) \left(\frac{\gamma(x - 2t)}{\gamma(x)} \right)^{1/2}, \quad (P_G \psi)(x) = ix\psi(x) - 2\psi'(x) ;$$

$$H_G \psi = (2N_G + \mathbf{1})\psi = -2 \frac{\partial^2}{\partial x^2} \psi + 2x \frac{\partial}{\partial x} \psi + \psi .$$

The Schrödinger representation.

Finally we get to the standard Schrödinger representation of the harmonic oscillator by dividing away a factor $\sqrt{\gamma(x)}$. Let $\mathcal{H}_S := L^2(\mathbb{R})$ and define

$$U_{GS} : \mathcal{H}_G \rightarrow \mathcal{H}_S : (U_{GS} \psi)(x) := \sqrt{\gamma(x)} \psi(x) .$$

Then we are back in the Schrödinger representation of the CCR.

4.4. The problem of damping.

A damped harmonic oscillator is an evolution $(T_t)_{t \geq 0}$ on the real-linear span of the canonical pair (P, Q) that has the form

$$\begin{aligned} T_t(Q) &= e^{-\gamma t} (Q \cos \omega t + P \sin \omega t) ; \\ T_t(P) &= e^{-\gamma t} (-Q \sin \omega t + P \cos \omega t) . \end{aligned} \tag{23}$$

(We apologise for the clash of notation: T_t is not related to translations.) This spiralling motion in the plane compresses areas by a factor $e^{-\gamma t}$, so that for $t > 0$ the operators $T_t(Q)$ and $T_t(P)$ disobey the canonical commutation relation, and T_t cannot be extended to an automorphism of $\mathcal{B}(\mathcal{H})$.

Yet this damped oscillatory behaviour occurs in nature, for instance when an atom is losing its energy to its surroundings by emission of light. So it would be worthwhile to make sense of it. There are two basic questions related to this model.

1. How should T_t be extended to $\mathcal{B}(\mathcal{H})$?
2. Can $(T_t)_{t \geq 0}$ be explained as part of a larger whole that evolves according to a Schrödinger equation?

Spirals and jumps.

In Heisenberg's matrix mechanics atoms were supposed to move in a mixture of two ways. Most of the time they were thought to rotate according to the evolution U_t^H as described above, but occasionally they made random jumps down the ladder of eigenvalues of the energy operator H . Each time an atom made such a jump, it emitted a quantum of light whose (angular) frequency ω was related to the size E of the jump by

$$E = \hbar \omega .$$

The probability per unit time for the atom to jump was given by Fermi's 'Golden Rule', formulated in terms of the coupling between the atom and its surroundings, and it is proportional to the damping rate γ .

In the following sections we shall describe this behaviour as a quantum Markov process. Both jumps and spirals will be visible in the extension of our T_t to the atom's full observable algebra. This will be our answer to question 1, for which we shall need the notion of completely positive operators.

Our answer to question 2 will be a reconstruction of the atom's surroundings. There we will see how the atom can absorb and emit quanta.

5. Conditional expectations and operations.

5.1. Conditional expectations in finite dimension.

In this section we choose for definiteness: $\mathcal{A} := M_n$, the algebra of all complex $n \times n$ matrices, and

$$\varphi : \mathcal{A} \rightarrow \mathbb{C} : \quad A \mapsto \text{tr}(\rho A),$$

where ρ is a symmetric $n \times n$ matrix with strictly positive eigenvalues and trace 1, so that φ is faithful.

Let A be a symmetric $n \times n$ matrix with the spectral decomposition

$$A = \sum_{\alpha \in \text{sp}(A)} \alpha E_\alpha.$$

The orthogonal projections E_α , $\alpha \in \text{sp}(A)$, form a partition of unity. *Measuring the observable A* means asking all the compatible, but mutually exclusive questions E_α at the same time. Precisely one of the answers will be ‘yes’, as stipulated in the interpretation rules. If the answer to E_α is ‘yes’, then A is said to take the value α . It is natural to define the *expectation* of A as

$$\sum_{\alpha \in \text{sp}(A)} \alpha \varphi(E_\alpha) = \varphi \left(\sum_{\alpha \in \text{sp}(A)} \alpha E_\alpha \right) = \varphi(A).$$

So our state φ not only plays the role of a probability measure, but naturally also as an expectation.

Now let $B = B^* \in \mathcal{A}$ be a second observable with spectral decomposition

$$B = \sum_{\beta \in \text{sp}(B)} \beta F_\beta.$$

If we first measure B and then A in each trial, we obtain a probability measure \mathbb{P} on $\text{sp}(A) \times \text{sp}(B)$ given by

$$\mathbb{P}(\{(\alpha, \beta)\}) = \varphi(F_\beta E_\alpha F_\beta).$$

For a probabilist it is then natural to define the conditional probability

$$\mathbb{P}[A = \alpha | B = \beta] := \frac{\mathbb{P}(\{(\alpha, \beta)\})}{\sum_{\alpha \in \text{sp}(A)} \mathbb{P}(\{(\alpha, \beta)\})} = \frac{\varphi(F_\beta E_\alpha F_\beta)}{\varphi(F_\beta)}.$$

The associated conditional expectation is naturally defined as

$$\mathbb{E}(A | [B = \beta]) := \sum_{\alpha \in \text{sp}(A)} \alpha \mathbb{P}[A = \alpha | B = \beta] = \frac{\varphi(F_\beta A F_\beta)}{\varphi(F_\beta)},$$

Note that this is a function, $f(\beta)$ say, of β . Seen as a quantum random variable this conditional expectation is described by the matrix $f(B)$:

$$\mathbb{E}(A|B) := f(B) = \sum_{\beta \in \text{sp}(B)} f(\beta)F_\beta = \sum_{\beta \in \text{sp}(B)} \frac{\varphi(F_\beta A F_\beta)}{\varphi(F_\beta)} F_\beta. \quad (24)$$

Note that

$$\varphi(\mathbb{E}(A|B)) = \sum_{\beta \in \text{sp}(B)} \varphi(F_\beta A F_\beta).$$

Remark. In general we do *not* have

$$\varphi(\mathbb{E}(A|B)) = \varphi(A). \quad (25)$$

The left hand side is the expectation of A after measuring B . The right hand side is the expectation of A without any previous operation. The fact that these two expectation values can differ is a quantummechanical phenomenon.

Let us give a simple counterexample to (25) here: Let $\mathcal{A} := M_2$, choose $\lambda \in (0, 1)$, and put

$$\rho = \begin{pmatrix} \lambda & 0 \\ 0 & 1 - \lambda \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

It is readily checked that

$$\varphi(A) = \lambda, \quad \varphi(\mathbb{E}(A|B)) = \frac{1}{2},$$

so that the equality (25) holds iff φ is the trace state.

The conditional expectation given a discrete random variable.

Let \mathcal{B} denote the (abelian) subalgebra of M_n generated by $\mathbf{1}$ and B . In quantum probability theory it is accepted practice to call $\mathbb{E}(A|B)$ the *conditional expectation of A given the algebra \mathcal{B}* , written $P_{\mathcal{B}}(A)$ *only* if the equality (25) *does* hold. The reason is that the definition as it stands does not generalise to observables B with continuous spectrum, let alone to non-commutative subalgebras \mathcal{B} of \mathcal{A} . (The value of $\varphi(\mathbb{E}(A|B))$ changes if the possible values of B are not all distinguished while measuring B : if β_1 and β_2 are not distinguished, their eigenspaces group together into a single subspace, and in (24) the projections F_{β_1} and F_{β_2} are replaced by the projection $F_{\beta_1} + F_{\beta_2}$.) Note that the projections in \mathcal{B} are labeled by *subsets* of $\text{sp}(B)$:

$$\mathcal{E}(\mathcal{B}) = \left\{ \sum_{\beta \in V} F_\beta \mid V \subset \text{sp}(B) \right\},$$

and that \mathcal{B} is the linear span of the projections F_β .

The following is a toy version of Takesaki's theorem [Tak73] on the existence of conditional expectations onto von Neumann subalgebras.

Theorem 5.1. *Let $B = B^* \in M_n$ and let \mathcal{B} be the $*$ -algebra generated by $\mathbf{1}$ and B . Let $\rho : M_n \rightarrow \mathbb{C} : A \mapsto \text{tr}(\rho A)$ with ρ strictly positive and $\text{tr}(\rho) = 1$. Then the following are equivalent.*

(a) *There exists a linear map $P : M_n \rightarrow \mathcal{B}$ such that*

$$\forall_{F \in \mathcal{E}(\mathcal{B})} : \quad \varphi(FAF) = \varphi(FP(A)F) . \quad (26)$$

(b) *There exists a linear map P from M_n onto \mathcal{B} such that*

- (i) *P maps positive definite matrices to positive definite matrices.*
- (ii) *$P(\mathbf{1}) = \mathbf{1}$;*
- (iii) *$\varphi \circ P = \varphi$;*
- (iv) *$P^2 = P$.*

(c) *$B\rho = \rho B$.*

If these equivalent conditions hold, then $P =: P_{\mathcal{B}}$ will be called the *conditional expectation onto \mathcal{B} compatible with φ* .

Proof. (a) \implies (b): suppose $P : \mathcal{A} \rightarrow \mathcal{B}$ is such that (26) holds.

Let $A \geq 0$ and put $P(A) =: \sum_{\beta \in \text{sp}(B)} a_{\beta} F_{\beta}$. Then $a_{\beta} \varphi(F_{\beta}) = \varphi(F_{\beta} P(A)) = \varphi(F_{\beta} P(A) F_{\beta}) = \varphi(F_{\beta} A F_{\beta}) \geq 0$. So $a_{\beta} \geq 0$ for all β and $P(A) \geq 0$.

Putting $A = \mathbf{1}$ we find that for all $\beta \in \text{sp}(B)$: $\varphi(F_{\beta}) = \varphi(F_{\beta} P(\mathbf{1}) F_{\beta}) = \varphi(F_{\beta} P(\mathbf{1}))$.

Writing $P(\mathbf{1}) = \sum_{\beta \in \text{sp}(B)} e_{\beta} F_{\beta}$, we see that $e_{\beta} = 1$, hence $P(\mathbf{1}) = \mathbf{1}$.

By putting $F = \mathbf{1}$ in (26), (iii) is obtained.

Finally, given A , the element $P(A)$ of \mathcal{B} is obviously uniquely determined by (26). But if $A \in \mathcal{B}$, then $P(A) := A$ clearly satisfies (26). It follows that P is an idempotent with range \mathcal{B} .

(b) \implies (c): Make a Hilbert space out of $\mathcal{A} = M_n$ by defining the inner product

$$\langle X, Y \rangle_{\varphi} := \varphi(X^* Y) .$$

We claim that on this Hilbert space P is an orthogonal projection. For this, it suffices to show that P is a contraction:

$$\|P(A)\|_{\varphi} \leq \|A\|_{\varphi} . \quad (27)$$

Define numbers $a_{\beta} \in \mathbb{C}$ and $b_{\beta} \geq 0$ by

$$P(A) = \sum_{\beta \in \text{sp}(B)} a_{\beta} F_{\beta} ; \quad P(A^* A) = \sum_{\beta \in \text{sp}(B)} b_{\beta} F_{\beta} .$$

Then from the positivity property it follows that

$$\forall_{\lambda \in \mathbb{C}} : \quad P((\lambda \cdot \mathbf{1} - A)^*(\lambda \cdot \mathbf{1} - A)) \geq 0 .$$

This implies that for all $\beta \in \text{sp}(B)$ and all $\lambda \in \mathbb{C}$,

$$|\lambda|^2 - (\bar{\lambda} a_{\beta} + \lambda \bar{a}_{\beta}) + b_{\beta} \geq 0 ,$$

from which it follows that

$$|a_{\beta}|^2 \leq b_{\beta} , \quad \text{i.e.} \quad P(A)^* P(A) \leq P(A^* A) .$$

Applying φ to the last inequality and using (ii) yields the statement (27). So P is an orthogonal projection, i.e.

$$A - P(A) \perp_{\varphi} \mathcal{B} .$$

Therefore for all $A \in M_n$ and all $F \in \mathcal{E}(\mathcal{B})$,

$$\varphi(FAF) - \varphi(FA) = \varphi(FP(A)) - \varphi(FA) = \langle F, P(A) \rangle_{\varphi} - \langle F, A \rangle_{\varphi} = 0 .$$

So

$$\mathrm{tr}(F\rho FA) = \mathrm{tr}(\rho FA) ,$$

hence $\rho F = F\rho F = (F\rho F)^* = (\rho F)^* = F\rho$ for all $F \in \mathcal{E}(\mathcal{B})$, and (c) follows.

(c) \implies (a): Suppose that $B\rho = \rho B$. Then for all $F \in \mathcal{E}(\mathcal{B})$ and all $A \in \mathcal{M}_n$,

$$\varphi(FAF) = \mathrm{tr}(\rho FAF) = \mathrm{tr}(F\rho FA) = \mathrm{tr}(\rho F^2 A) = \mathrm{tr}(\rho FA) = \varphi(FA) .$$

Therefore, defining $P(A)$ by the r.h.s. of (24), and putting $F = \sum_{\beta \in V} F_{\beta}$ with $V \subset \mathrm{sp}(B)$:

$$\begin{aligned} \varphi(FP(A)F) &= \sum_{\beta \in \mathrm{sp}(B)} \frac{\varphi(F_{\beta} A F_{\beta})}{\varphi(F_{\beta})} \varphi(F F_{\beta} F) = \sum_{\beta \in V} \varphi(F_{\beta} A F_{\beta}) \\ &= \sum_{\beta \in V} \varphi(F_{\beta} A) = \varphi(FA) = \varphi(FAF) . \end{aligned}$$

□

5.2. Operations in finite dimension.

Let \mathcal{A} and \mathcal{B} be finite dimensional von Neumann algebras, and let \mathcal{A}^* and \mathcal{B}^* denote their duals. A linear map $T : \mathcal{A} \rightarrow \mathcal{B}$ defines by duality a linear map $T^* : \mathcal{B}^* \rightarrow \mathcal{A}^*$. The map T is said to be *positive* if it maps positive elements of \mathcal{A} to positive elements of \mathcal{B} . In that case T^* maps states into states.

The map T is said to be *n-positive* if it maps positive elements of $\mathcal{A} \otimes M_n$ to positive elements of $\mathcal{B} \otimes M_n$:

$$(A_{ij})_{i,j=1}^n \geq 0 \implies (T(A_{ij}))_{i,j=1}^n \geq 0.$$

T is called *completely positive* if it is n -positive for all $n \in \mathcal{N}$. In that case $T^* \otimes \text{id}$ maps states on $\mathcal{B}^* \otimes M_n$ to states on $\mathcal{A}^* \otimes M_n$. T is called *identity preserving* if $T(\mathbf{1}_{\mathcal{A}}) = \mathbf{1}_{\mathcal{B}}$.

Definition. An *operation* $T : \mathcal{A} \rightarrow \mathcal{B}$ is a completely positive identity preserving map. Adjoints of operations will also be called operations.

The idea is that any physical procedure which takes as an input a state on some quantum system described by \mathcal{B} , and which turns out a state on a quantum system described by \mathcal{A} must necessarily be of the above kind. Not all operations in the sense of the definition can actually be performed, but certainly nothing else is physically possible. The reason to require complete positivity rather than just positivity is that any physical operation on a quantum system \mathcal{A} should also define a physical operation on $\mathcal{A} \otimes \mathcal{R}$, where \mathcal{R} stands for some quantum system not affected by the operation. The existence of such an ‘innocent bystander’ outside our quantum system \mathcal{A} should under no circumstances lead to the prediction by quantum theory of negative probabilities.

Substituting M_n for \mathcal{R} turns out to be sufficient.

The following example shows that complete positivity is strictly stronger than positivity. Let

$$T : M_2 \rightarrow M_2 : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

Then $T(A^*A) = T(A)T(A)^* \geq 0$ for all A , but

$$T \otimes \text{id} : M_2 \otimes M_2 \rightarrow M_2 \otimes M_2 : \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

maps the one-dimensional projection on the left to the matrix on the right with eigenvalues 1 and -1 .

5.3. Operations on quantum probability spaces.

A quantum probability space (\mathcal{A}, φ) has a canonical representation on a Hilbert space, called the GNS representation after Gel'fand, Naimark and Segal. It is the representation of \mathcal{A} on \mathcal{H}_φ , the Kolmogorov dilation of the positive definite kernel

$$\mathcal{A} \times \mathcal{A} \rightarrow \mathbb{C} : (A, B) \mapsto \varphi(A^*B) .$$

States which are given by density matrices on this space are called *normal states* on \mathcal{A} , and the set of all normal states is denoted by \mathcal{A}_* .

When we write $T : (\mathcal{A}, \varphi) \rightarrow (\mathcal{B}, \psi)$, we mean that T is a completely positive operator $\mathcal{A} \rightarrow \mathcal{B}$ such that $T(\mathbf{1}_\mathcal{A}) = \mathbf{1}_\mathcal{B}$ and *also* $\psi \circ T = \varphi$. The latter condition, which can equivalently be written as

$$T^*\psi = \varphi ,$$

ensures that T^* maps normal states to normal states. This property is only relevant for infinite dimensional von Neumann algebras.

5.4. Quantum Stochastic Processes.

Let us now consider the category QP whose objects are quantum probability spaces and whose morphisms are operations.

Lemma 5.2. (*Schwartz's inequality for completely positive operators*) Let $T : (\mathcal{A}, \varphi) \rightarrow (\mathcal{B}, \psi)$. Then for all $A \in \mathcal{A}$,

$$T(A^*A) \geq T(A)^*T(A) .$$

Proof. Let \mathcal{A} be represented on \mathcal{H} . By the positivity of $T \otimes \text{id}_{M_2}$ we have for all $A \in \mathcal{A}$,

$$(\psi, T(A)\psi)(T \otimes \text{id}) \left(\left(\begin{pmatrix} A & -\mathbf{1} \\ 0 & 0 \end{pmatrix} \right)^* \left(\begin{pmatrix} A & -\mathbf{1} \\ 0 & 0 \end{pmatrix} \right) \right) \begin{pmatrix} \psi \\ T(A)\psi \end{pmatrix} \geq 0 .$$

Writing this out we obtain

$$\langle \psi, (T(A^*A) - T(A)^*T(A))\psi \rangle \geq 0 .$$

□

Corollary 5.2. If $T : (\mathcal{A}, \varphi) \rightarrow (\mathcal{B}, \psi)$ then for all $A \in \mathcal{A}$

$$\varphi(T(A)^*T(A)) \leq \varphi(A^*A) .$$

This inequality states that T is a contraction on the GNS Hilbert space related to (\mathcal{A}, φ) .

Lemma 5.3. $T : (\mathcal{A}, \varphi) \rightarrow (\mathcal{B}, \psi)$ is an isomorphism in the category QP iff $T : \mathcal{A} \rightarrow \mathcal{B}$ is a *-isomorphism and $\psi \circ T = \varphi$.

Proof. (Exercise:) Apply Schwartz's inequality to T and T^{-1} . □

A *random variable* is an injective *-homomorphism

$$j : (\mathcal{A}, \varphi) \rightarrow (\widehat{\mathcal{A}}, \widehat{\varphi}) .$$

A *quantum stochastic process* is a family $(j_t)_{t \in \mathbb{T}}$ of random variables indexed by time. Here, \mathbb{T} is a linearly ordered set such as \mathbb{Z} , \mathbb{R} , \mathbb{N} or \mathbb{R}_+ . If $\mathbb{T} = \mathbb{R}$ or \mathbb{R}_+ we require that for all $A \in \mathcal{A}$ the curve $t \mapsto j_t(A)$ is strongly continuous.

If \mathbb{T} is a group, say \mathbb{Z} or \mathbb{R} , then the process is called *stationary* provided $j_t = j_0 \circ \widehat{T}_t$ for some representation $t \mapsto \widehat{T}_t$ of \mathbb{T} into the automorphisms of $(\widehat{\mathcal{A}}, \widehat{\varphi})$.

Open system interpretation.

We are observing a subsystem with observable algebra \mathcal{A} of a larger environment with algebra $\widehat{\mathcal{A}}$ that we cannot see. According to the Heisenberg picture, the smaller algebra is moving inside the larger one. A question $E \in \widehat{\mathcal{A}}$ can be asked at time t only if $E \in j_t(\mathcal{A})$. If $t_0 \leq t_1 \leq \dots \leq t_n$ is a sequence of times, and E_1, E_2, \dots, E_n a sequence of events in \mathcal{A} , then

$$\widehat{\varphi}(j_{t_1}(E_1)j_{t_2}(E_2) \cdots j_{t_{n-1}}(E_{n-1})j_{t_n}(E_n)j_{t_{n-1}}(E_{n-1}) \cdots j_{t_2}(E_2)j_{t_1}(E_1))$$

is the probability that E_1 occurs at time t_1 , E_2 at time t_2 , \dots , and E_n at time t_n . Note the double role played here by the time ordering: Unless some of the questions $j_{t_k}(E_k)$ recur, i.e. they lie in $j_t(\mathcal{A})$ for different values of t , they must be asked in the order dictated by the times t_k .

Stochastic process interpretation.

In a classical stochastic process $(X_t)_{t \in \mathbb{T}}$ the random variable X_t is a different one for different times t , so the events concerning X_t change in time accordingly. If the process is stationary, X_t and X_s differ by an automorphism of the underlying probability space. These remarks generalise to the non-commutative situation.

5.5. Conditional expectations and transition operators.

If we are to describe an open quantum system such as the damped harmonic oscillator by an internal dynamics, say $T_t : \mathcal{A} \rightarrow \mathcal{A}$, without reference to its surroundings, we need to be able to keep track of an observable A which starts in \mathcal{A} at time zero, during its motion away from the algebra \mathcal{A} at positive times. So we need its conditional expectation.

In view of Theorem 5.1. we give the following general definition.

Definition. . Let $j : (\mathcal{A}, \varphi) \rightarrow (\widehat{\mathcal{A}}, \widehat{\varphi})$ be a random variable. The *conditional expectation* (if it exists) is the unique morphism $P : (\widehat{\mathcal{A}}, \widehat{\varphi}) \rightarrow (\mathcal{A}, \varphi)$ for which

$$P \circ j = \text{id}_{\mathcal{A}} .$$

Without proof we state some properties.

Proposition 3.4. *If $P : (\widehat{\mathcal{A}}, \widehat{\varphi}) \rightarrow (\mathcal{A}, \varphi)$ is the conditional expectation with respect to $j : (\mathcal{A}, \varphi) \rightarrow (\widehat{\mathcal{A}}, \widehat{\varphi})$, then*

$$\forall_{B_1, B_2 \in \mathcal{A}} \forall_{A \in \widehat{\mathcal{A}}} : B_1 P(A) B_2 = P(j(B_1) A j(B_2)) .$$

In particular

$$\forall_{F \in \mathcal{E}(\mathcal{A})} \forall_{A \in \mathcal{A}} : \widehat{\varphi}(j(F) A j(F)) = \varphi(F P(A) F) .$$

The second line indicates the connection with Theorem 5.1.

5.7. Markov processes.

Let us now apply the above notion to an open quantum system.

two-time-probabilities.

Suppose that for all $s \in \mathbb{T}$ there exists a conditional expectation P_s with respect to j_s . Then the probability for F to occur at time s and E at time $t > s$ can be written as

$$\widehat{\varphi}(j_s(F) j_t(E) j_s(F)) = \varphi(F P_s(j_t(E)) F) = \varphi(F T_{s,t}(E) F) ,$$

where $T_{s,t} = P_s \circ j_t$ is an endomorphism of (\mathcal{A}, φ) , the *transition operator* from time s to time t .

multi-time-probabilities.

This reduction to the subsystem succeeds for more than two time points if there also exist conditional expectations onto the algebras

$$\mathcal{A}_{(-\infty, t]} := \text{vN}\{ j_s(\mathcal{A}) \mid s \leq t \} .$$

and moreover the *Markov property* holds:

$$s \leq t \implies P_{(-\infty, s]}(j_t(\mathcal{A})) \subset j_s(\mathcal{A}) . \quad (28)$$

Proposition 5.4. *Let $(j_t : (\mathcal{A}, \varphi) \rightarrow (\widehat{\mathcal{A}}, \widehat{\varphi}))_{t \in \mathbb{T}}$ be a Markov process with conditional expectations P_t . Then the transition operators form a monoid:*

$$0 \leq s \leq t \leq u \implies T_{s,t} T_{t,u} = T_{s,u} .$$

In particular, if the process is stationary, then $T_t := T_{0,t} = T_{s,s+t}$ satisfies

$$T_s T_t = T_{s+t} \quad (s, t \geq 0) .$$

In the latter case, $(T_t)_{t \geq 0}$ is known as the *dynamical semigroup* induced by the stationary Markov process. Conversely, the process $(j_t)_{t \in \mathbb{T}}$ is called the *Markov dilation* of the dynamical semigroup $(T_t)_{t \in \mathbb{T}}$.

The situation is symbolised by the diagram

$$\begin{array}{ccc} (\mathcal{A}, \varphi) & \xrightarrow{T_t} & (\mathcal{A}, \varphi) \\ j \downarrow & & \uparrow P \\ (\widehat{\mathcal{A}}, \widehat{\varphi}) & \xrightarrow{\widehat{T}_t} & (\widehat{\mathcal{A}}, \widehat{\varphi}) \end{array}$$

Our goal is to describe a Markov dilation of the damped harmonic oscillator.