

# Elements of Quantum Probability

B. KÜMMERER AND H. MAASSEN

This is an introductory article presenting some basic ideas of quantum probability. From a discussion of simple experiments with polarized light and a card game we deduce the necessity of extending the body of classical probability theory. For a class of systems, containing classical systems with finitely many states, a probabilistic model is developed. It can describe, in particular, the polarization experiments. Some examples of “quantum coin tosses” are discussed, closely related to V.F.R. Jones’ approach to braid group representations, to spin relaxation, and to nuclear magnetic resonance. In an appendix we indicate the steps which lead to the full mathematical model of quantum probability.

## INTRODUCTION

The aim of quantum probability is to provide a mathematical scheme in which classical probability theory and quantum mechanics can be discussed together.

Although quantum mechanics by its very nature always has been a probabilistic theory, it took a very long time for the typical concepts of classical probability theory to get a foothold in this area. Concepts such as random variables, stochastic processes, etc., were lacking. In the seventies, through developments in quantum optics, quantum field theory, and quantum statistical mechanics, a variety of probabilistic terminology began to invade quantum theory and the need for a common conceptual foundation became apparent. The efforts in this direction constitute the subject of quantum probability. The literature on this subject is a blend of functional analysis, quantum mechanics, and probability theory, mixed in different ratios and therefore hardly accessible to outsiders. The present paper is an attempt to present a motivation for the field and some of its basic ideas in an elementary way. For further reading we suggest the books of P. A. Meyer [Mey] and K. R. Parthasarathy [Par] and references therein.

The first section discusses polarization experiments which show the need to extend classical probability theory. In particular, we discuss Aspect’s experiment [Asp] and the violation of Bell’s inequality [Bel]. The experiment is illustrated by a card game. In the second section we show how the simple considerations about polarizers are paradigmatic for quantum probability. We develop its elementary building blocks in finite dimensions. To illustrate these ideas we consider some

---

Acknowledgment. The authors gratefully acknowledge the possibility of writing this article during a stay at the “Mathematisches Forschungsinstitut Oberwolfach”, financed also from the European Human Capital and Mobility Program on Quantum Probability.

finite coin tosses in the third section. Suprisingly, they lead quite quickly to recent developments. In an appendix we indicate the steps which lead to the full mathematical model of quantum probability.

## 1. WHY CLASSICAL PROBABILITY DOES NOT SUFFICE

### 1.1 *An experiment with polarisers*

To start with, we consider a simple experiment. In a beam of light of a fixed colour we put a pair of polarizing filters, each of which can be rotated around the axis formed by the beam. As is well known, the light which comes through both filters differs in intensity when the filters are rotated relative to each other. If we fix the first filter and rotate the second, then we see that there is a direction where the resulting intensity is maximal. Starting from this position, and rotating the second filter through an angle  $\alpha$ , the light intensity decreases with  $\alpha$  until it vanishes for  $\alpha = \frac{1}{2}\pi$ . Careful measurement shows that the intensity of the light passing the first filter is half the beam intensity (we assume the original beam being completely unpolarized) and that of the light passing the second filter is proportional to  $\cos^2 \alpha$ . If we call the intensity of the beam before the filters  $I_0$ , after the first  $I_1$ , and after the second  $I_2$ , then  $I_1 = \frac{1}{2}I_0$  and

$$I_2 = I_1 \cos^2 \alpha. \tag{1}$$

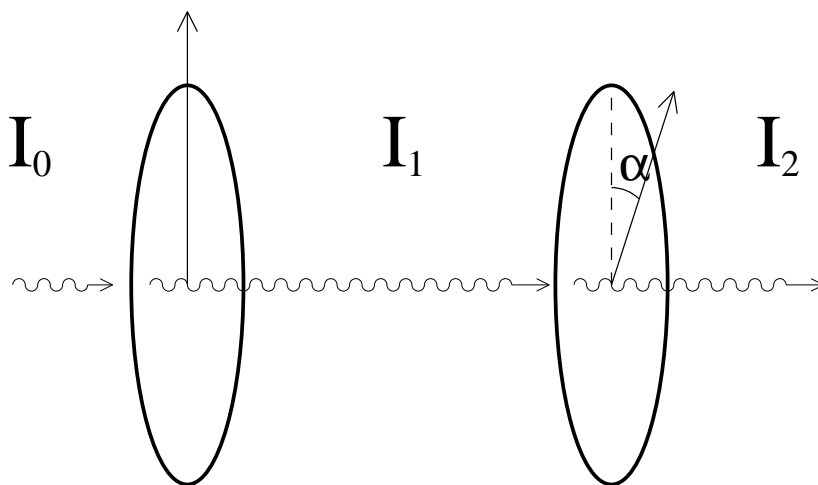


FIG. 1

Now, it is known from the work of Max Planck (1900) and Albert Einstein (1905) that for extremely low intensities monochromatic light comes in small packages, later called photons, all of the same energy, (which is independent of the total

intensity). Nowadays experimentalists are indeed able to produce these low intensities and see these single photons.

So the intensity must be proportional to the number of photons, and formula (1) has to be given a statistical meaning: a photon passing through the first filter has a probability  $\cos^2 \alpha$  to pass through the second. So formula (1) only holds on the average, i.e., for large numbers of photons.

If we think along the lines of classical probability, then we may attach to a polarization filter in the direction  $\alpha$  a random variable  $P_\alpha$ , taking the values 0 and 1, where  $P_\alpha(\omega) = 0$  if the photon  $\omega$  is absorbed by the filter and  $P_\alpha(\omega) = 1$  if it passes through. For two filters in the directions  $\alpha$  and  $\beta$  we may write for their correlation:

$$\mathbb{E}(P_\alpha P_\beta) = \mathbb{P}[P_\alpha = 1 \text{ and } P_\beta = 1] = \frac{1}{2} \cos^2(\alpha - \beta).$$

(Here a common notation from probability theory is used, namely, the expression  $[P_\alpha = 1 \text{ and } P_\beta = 1]$  stands for the set of those  $\omega$  for which  $P_\alpha(\omega) = 1$  and  $P_\beta(\omega) = 1$ .)

The following argument shows that this line of reasoning leads into difficulties. Take three polarizing filters  $\mathcal{F}_1$ ,  $\mathcal{F}_2$ , and  $\mathcal{F}_3$ , having polarization directions  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$  respectively. We put them on the optical bench in pairs. Then they give rise to random variables  $P_1$ ,  $P_2$  and  $P_3$  satisfying

$$\mathbb{E}(P_i P_j) = \frac{1}{2} \cos^2(\alpha_i - \alpha_j).$$

PROPOSITION. (Bell's 3 variable inequality) *For any three 0-1-valued random variables  $P_1$ ,  $P_2$ , and  $P_3$  on a probability space  $(\Omega, \mathbb{P})$  the following inequality holds:*

$$\mathbb{P}[P_1 = 1, P_3 = 0] \leq \mathbb{P}[P_1 = 1, P_2 = 0] + \mathbb{P}[P_2 = 1, P_3 = 0].$$

*Proof.* Write

$$\begin{aligned} \mathbb{P}[P_1 = 1, P_3 = 0] &= \mathbb{P}[P_1 = 1, P_2 = 0, P_3 = 0] + \mathbb{P}[P_1 = 1, P_2 = 1, P_3 = 0] \\ &\leq \mathbb{P}[P_1 = 1, P_2 = 0] + \mathbb{P}[P_2 = 1, P_3 = 0]. \end{aligned}$$

□

In our example, however, we have

$$\begin{aligned} \mathbb{P}[P_i = 1, P_j = 0] &= \mathbb{P}[P_i = 1] - \mathbb{P}[P_i = 1, P_j = 1] \\ &= \frac{1}{2} - \frac{1}{2} \cos^2(\alpha_i - \alpha_j) = \frac{1}{2} \sin^2(\alpha_i - \alpha_j). \end{aligned}$$

Bell's inequality thus reads

$$\frac{1}{2} \sin^2(\alpha_1 - \alpha_3) \leq \frac{1}{2} \sin^2(\alpha_1 - \alpha_2) + \frac{1}{2} \sin^2(\alpha_2 - \alpha_3),$$

which is clearly violated for  $\alpha_1 = 0, \alpha_2 = \frac{1}{6}\pi$  and  $\alpha_3 = \frac{1}{3}\pi$ , where it becomes

$$\frac{3}{8} \leq \frac{1}{8} + \frac{1}{8}.$$

We thus come to the conclusion that classical probability cannot describe this simple experiment!

*Remark*

The above calculation could be summarized as follows: we are in fact looking for a family of 0-1-valued random variables  $(P_\alpha)_{0 \leq \alpha < \pi}$  with  $\mathbb{P}[P_\alpha = 1] = \frac{1}{2}$ , satisfying the requirement that

$$\mathbb{P}[P_\alpha \neq P_\beta] = \sin^2(\alpha - \beta).$$

Now, on the space of 0-1-valued random variables on a probability space the function  $(X, Y) \mapsto \mathbb{P}[X \neq Y]$  equals the  $L^1$ -distance of  $X$  and  $Y$ :

$$\mathbb{P}[X \neq Y] = \int_{\Omega} |X(\omega) - Y(\omega)| \mathbb{P}(d\omega) = \|X - Y\|_1.$$

On the other hand, the function  $(\alpha, \beta) \mapsto \sin^2(\alpha - \beta)$  does not satisfy the triangle inequality for a metric on the interval  $[0, \pi)$ . Therefore no family  $(P_\alpha)_{0 \leq \alpha < \pi}$  exists which meets the above requirement.

## 1.2 An improved experiment

A possible criticism to the above argument runs as follows. Are the random variables  $P_\alpha$  well-defined? Is it indeed true that for each photon  $\omega$  and each filter  $\mathcal{F}_\alpha$  it is determined whether  $\omega$  passes through  $\mathcal{F}_\alpha$  or not? Could not filter  $\mathcal{F}_\alpha$  influence the photon's reaction to filter  $\mathcal{F}_\beta$ ? In fact, it seems quite obvious that it will!

In order to meet this criticism we should do a better experiment. We should let the filters act on each of the photons without influence on each other. At first sight this may seem impossible: one of the filters is bound to act first.

At this stage, however, a clever technique from quantum optics comes to our aid. Namely, it is possible to build a device that produces pairs of photons, such that the members of each pair move in opposite directions and show opposite behaviour towards polarization filters: if one passes the filter, then the other is surely absorbed. The device contains Calcium atoms, which are excited by a laser to a state they can only leave under emission of such a pair.

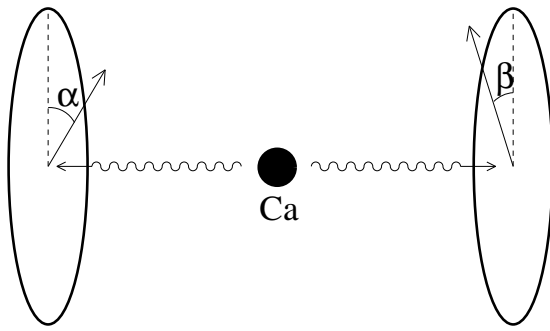


FIG. 2

With these photon pairs, the very same experiment can be performed, but this time the polarizers are far apart, each one acting on its own photon. The same correlations are measured, say first between  $P_\alpha$  on the left and  $P_\beta$  on the right, then between  $P_\alpha$  on the left and  $P_\gamma$  on the right, and finally between  $P_\beta$  on the left and  $P_\gamma$  on the right. The same outcomes are found, violating Bell's three variable inequality, thus strengthening the case against classical probability.

### 1.3 The decisive experiment

Advocates of classical probability could still find serious fault with the argument given so far. Indeed, do we really have to assume that we are measuring the same random variable  $P_\beta$  on the right as later on the left? Is it really true that the polarizations in these pairs are exactly opposite? There could exist a probabilistic explanation of the phenomena without this assumption.

So the argument has to be tightened still further. This brings us to an experiment which was actually performed by A. Aspect in Orsay (near Paris) in 1982 [Asp]. In this experiment a random choice out of two different polarization measurements was performed on each side of the pair-producing device, say in the direction  $\alpha_1$  or  $\alpha_2$  on the left and in the direction  $\beta_1$  or  $\beta_2$  on the right, giving rise to four random variables  $P_1 := P(\alpha_1)$ ,  $P_2 := P(\alpha_2)$  and  $Q_1 := Q(\beta_1)$ ,  $Q_2 := Q(\beta_2)$ , two of which are measured and compared at each trial.

PROPOSITION. (Bell's 4 variable inequality) *For any quadruple  $P_1, P_2, Q_1,$  and  $Q_2$  of 0-1-valued random variables on  $(\Omega, \mathbb{P})$  the following inequality holds:*

$$\mathbb{P}[P_1 = Q_1] \leq \mathbb{P}[P_1 = Q_2] + \mathbb{P}[P_2 = Q_1] + \mathbb{P}[P_2 = Q_2]. \quad (2)$$

(In fact, by symmetry, neither of these four probabilities is larger than the sum of the other three.)

*Proof.* Let  $A_{jk}$ , ( $j, k = 1, 2$ ), denote the random variable which is 1 if  $P_j = Q_k$  and 0 otherwise. Then clearly  $\mathbb{P}[P_j = Q_k] = \mathbb{E}(A_{jk})$ . We shall show that pointwise

$$A_{11} \leq A_{12} + A_{21} + A_{22}, \quad (3)$$

from which (2) is obtained by taking expectations. Indeed, in order to violate (3) one needs at least one point  $\omega \in \Omega$  for which  $A_{11}(\omega) = 1$  and  $A_{12}(\omega) = A_{21}(\omega) = A_{22}(\omega) = 0$ , i.e.,  $P_1(\omega) = Q_1(\omega)$  and  $P_1(\omega) \neq Q_2(\omega) \neq P_2(\omega) \neq Q_1(\omega)$ . But this is a contradiction, because there are an odd number of inequality signs  $\neq$  here.  $\square$

Note that Bell's 4-variable inequality is just the quadrangle inequality with respect to the metric  $(X, Y) \mapsto \|X - Y\|_1$ .

On the other hand, quantum mechanics predicts (cf. Section 2.4 below), and the experiment of Aspect showed, that one has,

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

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

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

So Bell's 4 variable inequality reads in this example:

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

which is clearly violated for the choices  $\alpha_1 = 0$ ,  $\alpha_2 = \frac{\pi}{3}$ ,  $\beta_1 = \frac{\pi}{2}$ , and  $\beta_2 = \frac{\pi}{6}$ , in which case it reads

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

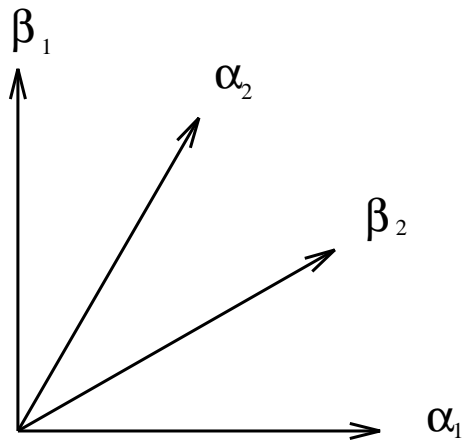


FIG. 3

So there does not exist, on any classical probability space, a quadruple  $P_1, P_2, Q_1$ , and  $Q_2$  of random variables with the correlations measured in this experiment.

*Remarks.*

1. When applying the above Proposition to the Orsay experiment, we should keep in mind that a crucial assumption has to be made. It must be assumed that for each  $\omega \in \Omega$  the values of  $P_j(\omega)$  and  $Q_j(\omega)$  are well-defined. This means that in each imagined realization of the world it is determined how each photon will react to any possible filter, including those it does not actually meet. This assumption is typical for classical probabilistic physical theories, but is abandoned in standard quantum mechanics. (Unmeasured quantities like the ones mentioned above are called ‘hidden variables’ in the literature on the foundations of quantum mechanics.)

2. A second important assumption, also necessary for the applicability of Bell’s inequality, is that the outcome on the right (described by  $Q(\beta)$  for some  $\beta$ ) should not depend on the angle  $\alpha$  of the polarizer on the left. We shall call this assumption ‘locality’. In order to justify this assumption, Aspect has made considerable efforts. In his (third) experiment, the choice of what to measure on the left ( $\alpha_1$  or  $\alpha_2$ ) and on the right ( $\beta_1$  or  $\beta_2$ ) was made *during the flight of the photons*, so that any influence which each of these choices might have on *the outcome* on the opposite end would have to travel faster than light. By the causality principle of Relativity Theory such influences are not possible.

3. Clearly, the above reasoning does not exclude the possibility of an explanation of the experiment in classical probabilistic terms, if one is willing to give up the causality principle. Serious attempts have been made in this direction (e.g. [Boh]). The majority of physicists, however, seem to be convinced of the principle of causality.

1.4 *The Orsay experiment as a card game*

It has now become very difficult for the advocates of classical probability to criticize the experiment. To illustrate this point, we shall again present the experiment, but this time in the form of a card game. Nature can win this game. Can you?

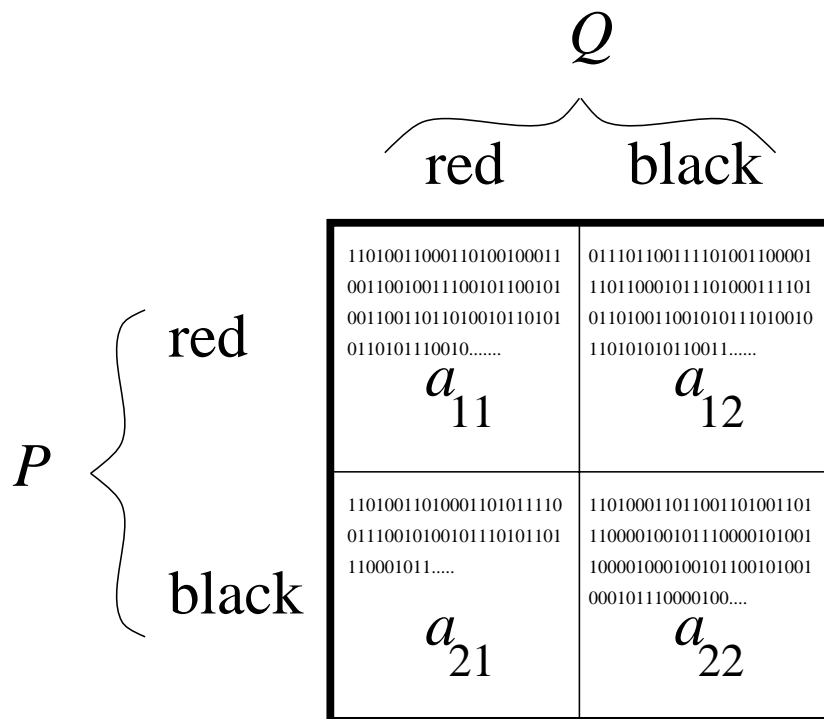


FIG. 4

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. 4. 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 (3) of Bell’s 4-variable inequality (recall Section 1.3), 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, solely based on the  $\cos^2$  law (1) for photon absorption! 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. 3). If his card is black, then he chooses

$\alpha_2 = \frac{\pi}{3}$ . If  $Q$  has a red card, then he chooses  $\beta_1 = \frac{\pi}{2}$ . If his card is black, then he chooses  $\beta_2 = \frac{\pi}{6}$ . 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.

## 2. TOWARDS A MATHEMATICAL MODEL

### 2.1 *A mathematical description of polarization*

Coerced by the foregoing considerations, we give up trying to make a classical probabilistic model in order to explain polarization experiments. Instead, we take these experiments as a paradigm for an alternative type of probability, to be developed now.

We have discussed (linear) polarization of a light beam. This is completely characterized by a direction in the plane perpendicular to the light beam. This suggests that we should describe different directions of polarization by different directions in a two-dimensional real plane  $\mathbb{R}^2$ , or equivalently by unit vectors  $x \in \mathbb{R}^2$ ,  $\|x\| = 1$ , pointing in this direction. Moreover, it appears that we cannot physically distinguish between two states which differ by a rotation of  $\pi$ , so we have to describe states of polarizations by one-dimensional subspaces of  $\mathbb{R}^2$ . (Two unit vectors span the same one-dimensional subspace if they differ only by a sign.) Given two directions of polarization with an angle  $\alpha$  between them, spanned by two unit vectors  $x, y \in \mathbb{R}^2$ , the transition probability  $\cos^2 \alpha$  can be expressed as

$$\cos^2 \alpha = \langle x, y \rangle^2$$

where  $\langle x, y \rangle$  denotes the scalar product between  $x$  and  $y$ . (Since  $\cos^2 \alpha = \cos^2(\pi - \alpha)$ , this expression does not depend on the sign of  $x, y$ .)

Certainly, in order to come to a mathematical model we should distinguish between the physical state of polarization of a photon on the one hand and the filter on the other hand, i.e., the 0-1-valued random variable which asks, whether a photon is polarized in a certain direction. This can be done by identifying the filter, (i.e., the random variable), with the orthogonal projection  $P$  onto the one-dimensional subspace. We can then write

$$\cos^2 \alpha = \langle x, y \rangle^2 = \langle Px, x \rangle \quad .$$

So we arrive at the following mathematical model:

States of polarization of a photon  $\hat{=}$  one-dimensional subspaces of  $\mathbb{R}^2$  described by unit vectors  $x$  spanning the subspace.

Polarization filters, (i.e., random variables measuring polarization)  $\hat{=}$  orthogonal projections  $P$  from  $\mathbb{R}^2$  onto the corresponding one-dimensional subspace.

Probability that a photon, described by  $x$ , passes through the filter described by  $P$   $\hat{=}$   $\langle Px, x \rangle = \cos^2 \alpha$ .

Since  $P$  is 0-1-valued, (i.e., a photon passes or is absorbed), this probability is equal to the expectation of this random variable:

$$\langle Px, x \rangle = \mathbb{E}(P) .$$

It is important to realize that, although we gave a kind of proof (in Section 1) that polarization experiments *cannot* be described by classical random variables on classical probability spaces, there is *no* logical argument that photons must be described by vectors and filters by projections, as we just did. Indeed, since the beginnings of quantum mechanics there have been many efforts to develop alternative mathematical models. We are going to describe here the traditional point of view of quantum mechanics [Neu]. This will lead to a mathematical model which extends classical probability and up until now has described experiments correctly.

## 2.2 *The full quantum mechanical truth about polarization*

In the foregoing description of polarization things were presented somewhat simpler than they are: we considered only linear polarization, thus disregarding circular polarization. The full description of polarization leads to the quantum mechanics of a 2-level system:

State of polarization of a photon  $\hat{=}$  one-dimensional subspace of  $\mathbb{C}^2$ , described by a unit vector  $x$  spanning this subspace (and determined only up to a phase).

Polarization filter or generalized 0-1-valued random variable  $\hat{=}$  orthogonal projection  $P$  onto a complex one-dimensional subspace.

(Also for left- or right-circular polarization do there exist physical filters.)

Probability for a photon, described  $\hat{=} \langle Px, x \rangle$ ,  
 by  $x$ , to pass through a filter, de-  
 scribed by  $P$

The set of all states is conveniently parametrized by the unit vectors of the  
 form

$$(\cos \alpha, e^{i\phi} \sin \alpha) \in \mathbb{C}^2, \quad \frac{-\pi}{2} \leq \alpha \leq \frac{\pi}{2}, \quad 0 \leq \phi \leq \pi.$$

This set can be identified with the points on the unit sphere  $S^2 \in \mathbb{R}^3$  when using  
 the polar coordinates  $\theta = 2\alpha$  and  $\phi$ . Restricting to states with  $\phi = 0$ , (which  
 are parametrized by the points of the circle  $(\cos 2\alpha, \sin 2\alpha)$  in  $\mathbb{R}^2$ ), we retain the  
 foregoing real description when we identify  $\alpha$  with the angle of polarization.

A possible identification of the points of  $S^2 \subseteq \mathbb{R}^3$  with physical states, giving  
 the correct values for all probabilities, is as in the picture below:

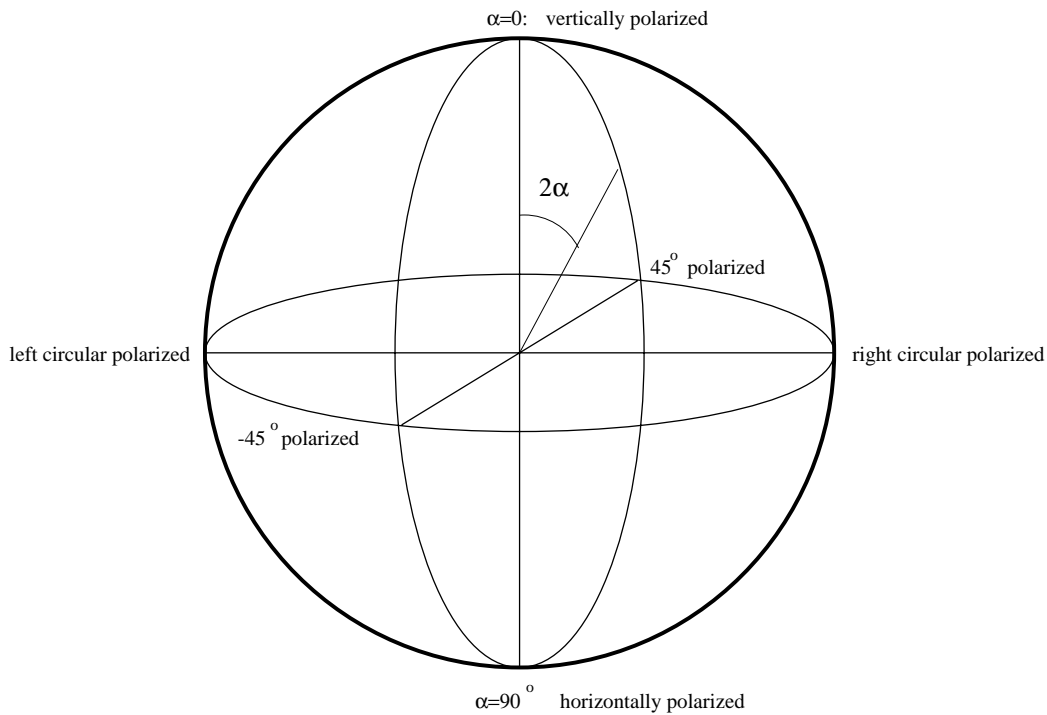


FIG. 5

With these identifications we come to the following mathematical model of  
 polarization of light:

vertically polarized light  $\hat{=} (1, 0) \in \mathbb{C}^2$

$$\begin{aligned} \text{horizontally polarized light} &\hat{=} (0, 1) \in \mathbb{C}^2 \\ \text{light polarized at an angle } \alpha \text{ to the} &\hat{=} (\cos \alpha, \sin \alpha) \in \mathbb{C}^2 \\ \text{vertical direction} & \\ \text{light polarized at an angle } \alpha = \pm \frac{\pi}{4} &\hat{=} \left( \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}} \right) \in \mathbb{C}^2 \\ \text{to the vertical direction} & \\ \text{left-/right-circular polarized light} &\hat{=} \left( \frac{1}{\sqrt{2}}, \pm \frac{i}{\sqrt{2}} \right) \in \mathbb{C}^2 \end{aligned}$$

and correspondingly

$$\begin{aligned} \text{vertical polarizer} &\hat{=} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\ \text{horizontal polarizer} &\hat{=} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ \text{angle-}\alpha\text{-polarizer} &\hat{=} \begin{pmatrix} \cos^2 \alpha & \cos \alpha \sin \alpha \\ \cos \alpha \sin \alpha & \sin^2 \alpha \end{pmatrix} \\ \pm \frac{\pi}{4}\text{-polarizer} &\hat{=} \begin{pmatrix} \frac{1}{2} & \pm \frac{1}{2} \\ \pm \frac{1}{2} & \frac{1}{2} \end{pmatrix} \\ \text{left/right circular polarizer} &\hat{=} \begin{pmatrix} \frac{1}{2} & \mp \frac{i}{2} \\ \pm \frac{i}{2} & \frac{1}{2} \end{pmatrix} \end{aligned}$$

### 2.3 Finite dimensional models

The mathematical model that is used by quantum mechanics is the straightforward generalization of the above description. In order to keep things simple, we restrict ourselves to the quantum mechanics of finite dimensions. It generalizes the probability theory of systems with only finitely many states. As in classical probability, the generalization to systems with a countable number of states or a continuum of states is analytically more involved, though conceptually easy.

The model is as follows:

**States** correspond to one-dimensional subspaces of  $\mathbb{C}^n$ , where the dimension  $n$  is determined by the model. Again, a state is described conveniently by some unit vector spanning this subspace.

**0-1-valued random variables** are described by orthogonal projections onto a linear subspace of  $\mathbb{C}^n$ . If the random variable only asks whether the system is in a certain state, then the subspace is one-dimensional. But also projections onto higher dimensional subspaces  $\mathcal{K}$  make sense. They answer the question whether

the system is in any of the states represented by a unit vector in  $\mathcal{K}$ . Similar questions for other subsets of states are not allowed!

The **probability** that a measurement of a random variable  $P$  on a system in a state  $x$  gives the value 1 is still given by  $\langle Px, x \rangle$ .

Note that we do not assume that every unit vector  $x \in \mathbb{C}^n$  describes a state of the system, nor that every orthogonal projection corresponds to a meaningful random variable. Specializing these two sets is part of the description of the mathematical model for a given system. In a truly quantum mechanical situation, typically all possible vectors and projections are used. In contrast to this, a model from classical probability is incorporated into this description as follows.

A finite probability space is usually described by a finite set  $\Omega = \{\omega_1, \dots, \omega_n\}$  and a probability distribution  $(p_1, \dots, p_n)$ ,  $0 \leq p_i \leq 1$ ,  $\sum_i p_i = 1$ , such that the probability for  $\omega_i$  is  $p_i$ . A 0-1-valued random variable is a 0-1-valued function on  $\Omega$ , i.e., a characteristic function  $\chi_A$  of some subset  $A \subseteq \Omega$ . In order to describe such a system in our model, we think of  $\mathbb{C}^n$  as the space of complex valued functions on  $\Omega$ , and use the functions  $\delta_i$  with  $\delta_i(\omega_j) = \delta_{i,j}$  as basis. The states of the system, i.e., the points  $\omega_i$  of  $\Omega$ , are now represented by the unit vectors  $\delta_i$ ,  $1 \leq j \leq n$ . The random variable  $\chi_A$  is identified with the orthogonal projection  $P_A$  onto the linear span of the vectors  $\{\delta_i : \omega_i \in A\}$ . In our basis  $\chi_A$  becomes a diagonal matrix with a 1 at the  $i$ -th place of the diagonal if  $\omega_i \in A$ , and a 0 otherwise. It is obvious that  $\omega_i \in A$  if and only if  $\chi_A(\omega_i) = 1$  if and only if  $\langle P_A \delta_i, \delta_i \rangle = 1$ .

Conversely, any set of pairwise commuting projections on  $\mathbb{C}^n$  can be diagonalized simultaneously and thus have an interpretation as a set of classical 0-1-valued random variables. Therefore:

*Classical probability corresponds to sets of pairwise commuting projections.*

In the above sketch of classical probability an important point is obviously missing: So far we have only considered pure states of the system, a probability distribution  $(p_1, \dots, p_n)$  did not enter the discussion. How can we describe a situation where a system is in a certain state  $x$  with probability  $q$  and in another state  $y$  with probability  $1 - q$  ( $0 \leq q \leq 1$ ) ?

Obviously, the set of states should be a convex set, containing also the mixed states. In the classical model of probability, the appropriate convex combinations of point measures are taken in order to obtain a new probability measure.

In general, if  $P$  is any 0-1-valued (quantum) random variable and  $x_1, \dots, x_k$  are arbitrary quantum states, each occurring with a probability  $p_i$ ,  $1 \leq i \leq k$ ,  $\sum_i p_i = 1$ ,  $p_i \geq 0$ , then the probability that a measurement of  $P$  gives 1 is clearly given by

$$\sum_i p_i \langle Px_i, x_i \rangle \quad .$$

A more convenient description of mixed states is obtained as follows.

For a unit vector  $x \in \mathbb{C}^n$  denote by  $\Phi_x$  the orthogonal projection onto the one-dimensional subspace generated by  $x$ . In the physics literature,  $\Phi_x$  is frequently denoted by  $|x\rangle\langle x|$ . By  $tr$  denote the trace on the  $n \times n$ -matrices, summing up the diagonal entries of such a matrix. Then one obtains

$$\langle Px, x \rangle = tr(\Phi_x \cdot P) .$$

Hence

$$\sum_i p_i \langle Px_i, x_i \rangle = tr\left(\sum_i p_i \Phi_{x_i} \cdot P\right) = tr(\Phi \cdot P) ,$$

where  $\Phi := \sum_i p_i \Phi_{x_i}$ .

Being a convex combination of 1-dimensional projections,  $\Phi$  obviously is a positive (i.e., self-adjoint positive semidefinite)  $n \times n$ -matrix with  $tr(\Phi) = 1$ . Conversely, from diagonalizing positive matrices it is clear that any such positive matrix  $\Phi$  with  $tr(\Phi) = 1$  can be written as a convex combination of 1-dimensional projections. The set of these matrices forms a closed (even compact) convex set, and its extreme points are precisely the 1-dimensional projections which in turn correspond to pure states, represented also by unit vectors. Therefore it is precisely this class of matrices which represents mixed states. These matrices are frequently called *density matrices*.

Thus, a general mixed state is described by a density matrix  $\Phi$  and the probability for an observation of  $P$  to yield the value 1 is given by  $tr(\Phi \cdot P)$ .

*Remarks*

1. Although in this description also pure states are described by 1-dimensional projections, they are not considered as random variables.

2. The decomposition of a density matrix  $\Phi$  into a convex combination of 1-dimensional projections is by no means unique. The compact convex set of density matrices is far from being a simplex. Indeed, on  $\mathbb{C}^2$  it can be affinely identified with a full ball in  $\mathbb{R}^3$ , by taking in  $\mathbb{R}^3$  the convex hull of the sphere that was described above.

3. In classical probability the convex set of mixed states is the simplex of all probability distributions. In our picture, if we insist on decomposing a mixed state given by  $\Phi = \sum_i p_i P_{\delta_i}$  into a convex combination of pure states (within the convex hull of  $\{P_{\delta_i} : 1 \leq i \leq n\}$  which is a simplex), then it becomes unique.

4. Physically, a state  $\Phi$  is completely described by all of its values  $tr(\Phi \cdot P)$ , where  $P$  runs through the random variables of the model. Thus, if we consider only subsets of projections, then two different density matrices can represent the same physical state of the system. As a drastic example, consider the classical system  $\Omega = \{\omega_1, \dots, \omega_n\}$  with equidistribution, i.e.,  $p_i(\omega_i) = \frac{1}{n}$ , leading to the density matrix  $\Phi = \sum_i \frac{1}{n} P_{\delta_i} = \frac{1}{n} \cdot \mathbb{1}$ . On the other hand, with the unit vector  $x = (\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}) \in \mathbb{C}^n$ , we obtain for any subset  $A \subseteq \Omega$ :  $tr(\Phi \cdot P_A) = \frac{1}{n} \cdot |A| = \langle P_A x, x \rangle$ . Therefore, on the random variables  $\{P_A : A \subseteq \Omega\}$ , the rank-one-density matrix  $P_x$  represents the same state as the density matrix  $\frac{1}{n} \cdot \mathbb{1}$ . Note, however, that  $P_x$  is not in the convex hull of  $\{P_{\delta_i} : 1 \leq i \leq n\}$ .

## 2.4 The mathematical model of Aspect's experiment

As an illustration, we shall now explain the photon correlation in the Orsay experiment, given by the  $\cos^2$ -law. Note that here we cannot simply refer to the basic  $\cos^2$ -law of quantum probability, since the filters are acting on two different photons.

The polarization of a pair of photons is described by a unit vector in the tensor product  $\mathbb{C}^2 \otimes \mathbb{C}^2 = \mathbb{C}^4$ , where we use the basis

$$\begin{aligned} (1, 0, 0, 0) &= e_1 \otimes e_1 =: e_{11}, \\ (0, 1, 0, 0) &= e_1 \otimes e_2 =: e_{12}, \\ (0, 0, 1, 0) &= e_2 \otimes e_1 =: e_{21}, \\ (0, 0, 0, 1) &= e_2 \otimes e_2 =: e_{22}, \end{aligned}$$

with  $e_1 = (1, 0) \in \mathbb{C}^2$  and  $e_2 = (0, 1) \in \mathbb{C}^2$ . For example, in the pure state  $e_{12}$  the left-hand photon is vertically polarized and the right-hand photon horizontally. As it turns out, the state of the pair of photons as produced by the Calcium atom is described by the state

$$x = \frac{1}{\sqrt{2}}(e_{12} - e_{21}).$$

Now, the filters  $P(\alpha)$  on the left and  $Q(\beta)$  on the right, introduced in §1.3, are represented by two-dimensional projection operators on  $\mathbb{C}^4$ , which are the “2-right amplification” and the “2-left-amplification” of the polarization matrix

$$\begin{pmatrix} \cos^2 \alpha & \cos \alpha \sin \alpha \\ \cos \alpha \sin \alpha & \sin^2 \alpha \end{pmatrix},$$

namely

$$\begin{aligned} P(\alpha) &= \begin{pmatrix} \cos^2 \alpha & \cos \alpha \sin \alpha \\ \cos \alpha \sin \alpha & \sin^2 \alpha \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos^2 \alpha & 0 & \cos \alpha \sin \alpha & 0 \\ 0 & \cos^2 \alpha & 0 & \cos \alpha \sin \alpha \\ \cos \alpha \sin \alpha & 0 & \sin^2 \alpha & 0 \\ 0 & \cos \alpha \sin \alpha & 0 & \sin^2 \alpha \end{pmatrix} \\ Q(\beta) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} \cos^2 \beta & \cos \beta \sin \beta \\ \cos \beta \sin \beta & \sin^2 \beta \end{pmatrix} \\ &= \begin{pmatrix} \cos^2 \beta & \cos \beta \sin \beta & 0 & 0 \\ \cos \beta \sin \beta & \sin^2 \beta & 0 & 0 \\ 0 & 0 & \cos^2 \beta & \cos \beta \sin \beta \\ 0 & 0 & \cos \beta \sin \beta & \sin^2 \beta \end{pmatrix}. \end{aligned}$$



(More about such tensor products will be treated in Section 3.)

We note that  $P(\alpha)$  and  $Q(\beta)$  are commuting projections for fixed  $\alpha$  and  $\beta$ . It follows that  $P(\alpha)Q(\beta)$  is again a projection, as well as the products  $P(\alpha)(\mathbb{1} - Q(\beta))$ ,  $(\mathbb{1} - P(\alpha))Q(\beta)$ , and  $(\mathbb{1} - P(\alpha))(\mathbb{1} - Q(\beta))$ . So we obtain the description of a classical probability space with four states, to be interpreted as

(“left photon passes”, “right photon passes”),  
 (“left photon passes”, “right photon is absorbed”),  
 (“left photon is absorbed”, “right photon passes”),  
 (“left photon is absorbed”, “right photon is absorbed”).

The probabilities of these four events are found by the actions on  $x = \frac{1}{\sqrt{2}}(e_{12} - e_{21}) = \frac{1}{2}(0, 1, -1, 0)$  of the four projections. In particular, the probability that both photons pass is given by

$$\begin{aligned}
 & \langle x, P(\alpha)Q(\beta)x \rangle \\
 &= \frac{1}{2}(0, 1, -1, 0) \times \\
 & \quad \times \begin{pmatrix} \cos^2 \alpha \cos^2 \beta & \cos^2 \alpha \cos \beta \sin \beta & \cos \alpha \sin \alpha \cos^2 \beta & \cos \alpha \sin \alpha \cos \beta \sin \beta \\ \cos^2 \alpha \cos \beta \sin \beta & \cos^2 \alpha \sin^2 \beta & \cos \alpha \sin \alpha \cos \beta \sin \beta & \cos \alpha \sin \alpha \sin^2 \beta \\ \cos \alpha \sin \alpha \cos^2 \beta & \cos \alpha \sin \alpha \cos \beta \sin \beta & \sin^2 \alpha \cos^2 \beta & \sin^2 \alpha \cos \beta \sin \beta \\ \cos \alpha \sin \alpha \cos \beta \sin \beta & \cos \alpha \sin \alpha \sin^2 \beta & \sin^2 \alpha \cos \beta \sin \beta & \sin^2 \alpha \sin^2 \beta \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \\
 &= \frac{1}{2}(\cos^2 \alpha \sin^2 \beta + \sin^2 \alpha \cos^2 \beta - 2 \cos \alpha \sin \alpha \cos \beta \sin \beta) \\
 &= \frac{1}{2}(\cos \alpha \sin \beta - \sin \alpha \cos \beta)^2 \\
 &= \frac{1}{2} \sin^2(\alpha - \beta) .
 \end{aligned}$$

### 3. EXAMPLES OF QUANTUM COIN TOSSING

#### 3.1 Model of a quantum coin toss

A fair quantum coin toss should be a (finite or infinite) sequence of 0-1-valued random variables  $(P_i)_i$  such that for each random variable  $P_i$  the probability of measuring 1 or 0 is  $\frac{1}{2}$  each. Moreover, the random variables  $P_i$  and  $P_j$  should be ‘independent’ for  $i \neq j$  in a sense yet to be made precise.

In our mathematical model the above implies that we ought to construct sequences of orthogonal projections such that in a certain mixed state, i.e., for a certain density matrix  $\Phi$ , we have  $tr(\Phi \cdot P_i) = \frac{1}{2}$ . Independence of  $P_i$  and

$P_j$  means that the probability for  $P_i$  and  $P_j$ ,  $i \neq j$ , both yielding 1 is equal to  $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ . However, as the random variables do not commute, one has to decide about the order in which the random variables are measured. Therefore, any of the requirements  $\text{tr}(\Phi \cdot P_i P_j) = \frac{1}{4}$ ,  $\text{tr}(\Phi \cdot P_j P_i) = \frac{1}{4}$ ,  $\text{tr}(\Phi \cdot P_i P_j P_i) = \frac{1}{4}$ ,  $\text{tr}(\Phi \cdot P_j P_i P_j) = \frac{1}{4}$  are good candidates for the definition of independence. In fact, one usually considers situations where all these requirements are equivalent. In the examples to follow we shall consider the simplest of such situations, namely where  $\Phi = \frac{1}{n} \cdot \mathbb{1}$ . Then  $\text{tr}(\Phi \cdot X)$  is the normalized trace of  $X$  and the above conditions are obviously equivalent. This mixed state is often called the *chaotic state* as it is the state of equidistribution. It extends the classical equidistribution on a finite set  $\Omega$ , since it can be written as  $\sum_i \frac{1}{n} P_{x_i}$  where  $x_i$  runs through any orthonormal basis of  $\mathbb{C}^n$ ; in particular,  $x_i$  can be chosen as  $\delta_i$  (cf. §2.3).

Since in our description we restricted to probability for systems on finite dimensional spaces, i.e., projections on  $\mathbb{C}^n$  for some finite  $n$ , one cannot expect to find an infinite number of such projections  $P_i$ . Therefore we shall construct the coin tosses for any finite number of experiments.

In order to be more explicit in the following examples some notation is helpful.

Denote by  $M_n$  the set of all complex  $n \times n$ -matrices, i.e., the set of all (complex-) linear maps from  $\mathbb{C}^n$  to  $\mathbb{C}^n$  in a certain fixed orthonormal basis  $\{e_1, \dots, e_n\}$ .

The tensor product  $\mathbb{C}^m \otimes \mathbb{C}^n$ , which is isomorphic to  $\mathbb{C}^{m \cdot n}$ , will be identified with the direct sum  $\bigoplus_{i=1}^m \mathbb{C}^n$  where the  $i$ -th direct summand contains the vectors  $\{e_i \otimes x : x \in \mathbb{C}^n\} \subseteq \mathbb{C}^m \otimes \mathbb{C}^n$ . With this identification a matrix  $A \in M_{mn}$  can be written as an element in  $M_m(M_n)$ , i.e., an  $m \times m$ -block-matrix with entries in  $M_n$ . In particular, given  $A \in M_m$ ,  $A = (\alpha_{ij})$ , its “ $n$ -right-amplification”  $A \otimes \mathbb{1}_n$  is the block matrix  $(\alpha_{ij} \cdot \mathbb{1}_n)_{1 \leq i, j \leq m}$ , where  $\mathbb{1}_n$  is the identity matrix in  $M_n$ . Similarly, given  $B \in M_n$ , its “ $m$ -left-amplification”  $\mathbb{1}_m \otimes B$  is the  $m \times m$ -block diagonal matrix with all diagonal blocks equal to  $B$ .

Thus, if  $m = 2$ , the  $n$ -right-amplification of

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \quad \text{is} \quad \begin{pmatrix} a_{11} & \mathbf{0} & a_{12} & \mathbf{0} \\ & \ddots & & \ddots \\ \mathbf{0} & & a_{11} & \mathbf{0} & a_{12} \\ a_{21} & \mathbf{0} & a_{22} & \mathbf{0} \\ & \ddots & & \ddots \\ \mathbf{0} & a_{21} & \mathbf{0} & a_{22} \end{pmatrix},$$

if  $B \in M_n$ , its 2-left-amplification is the block matrix

$$\begin{pmatrix} B & \mathbf{0} \\ \mathbf{0} & B \end{pmatrix}.$$

The following observation is helpful: Suppose, we have constructed some random variables  $P_1, \dots, P_k$  on  $\mathbb{C}^n$  which meet our requirements on a coin tossing.

Then all requirements on a coin tossing remain satisfied if we take a left- (or right-) amplification of all of them. In fact, they describe the same physical situation, since all possible correlations are the same.

### 3.2 Classical coin tossing

Tossing a classical coin, say a penny, has in our model the following description (we label the two sides of the penny by 0 and 1):

The random variable for tossing once is

$$P_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} .$$

The random variables for tossing twice are

$$P_1 = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 0 & \\ & & & 0 \end{pmatrix} \text{ and } P_2 = \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & 1 & \\ & & & 0 \end{pmatrix}$$

(In a different basis this example has already occurred in Section 2.3). Here and in the following, we omit 0 entries in matrices, whenever this is convenient. The random variables for tossing three times are

$$P_1 = \begin{pmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & 1 & & & & & \\ & & & 1 & & & & \\ & & & & 0 & & & \\ & & & & & 0 & & \\ & & & & & & 0 & \\ & & & & & & & 0 \end{pmatrix} P_2 = \begin{pmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & 0 & & & & & \\ & & & 0 & & & & \\ & & & & 1 & & & \\ & & & & & 1 & & \\ & & & & & & 0 & \\ & & & & & & & 0 \end{pmatrix}$$

$$P_3 = \begin{pmatrix} 1 & & & & & & & \\ & 0 & & & & & & \\ & & 1 & & & & & \\ & & & 0 & & & & \\ & & & & 1 & & & \\ & & & & & 0 & & \\ & & & & & & 1 & \\ & & & & & & & 0 \end{pmatrix} .$$

Generally, from the random variables  $\bar{P}_1, \dots, \bar{P}_n \in M_k$  describing  $n$  tossings we recursively construct  $P_1, \dots, P_{n+1}$  for  $n+1$  tossings:

$P_{i+1}$  is the 2-left-amplification of  $\bar{P}_i$ ,  $1 \leq i \leq n$ , i.e.,  $P_{i+1} = \mathbb{1}_2 \otimes \bar{P}_i$ ,

$P_1$  is the  $k$ -right-amplification of  $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ , i.e.,  $P_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \mathbb{1}_k$ .

### 3.3 Polarizers lead to quantum coins

In order to construct non-classical coins, we go back to our polarization experiments.

The chaotic state for the polarization of light is just unpolarized light. This is physically defined as light such that any polarization filter decreases its intensity to 50 %.

If such a beam of unpolarized light passes first through a vertically (linearly) polarizing filter and then through a filter rotated against the first through an angle of  $\alpha = \frac{\pi}{4}$ , then its intensity has decreased to  $\frac{1}{4}$ . Hence these two filters correspond to two tossings of a quantum coin, in our mathematical model described by the projections

$$P_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } P_2 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} .$$

Since these projections do not commute, they do not describe a classical situation, cf. Section 1. Instead they satisfy the following relations:

$$P_1 P_2 P_1 = \frac{1}{2} P_1, \quad P_2 P_1 P_2 = \frac{1}{2} P_2 .$$

Can we extend these two tossings to a coin tossing of arbitrary length? The answer is ‘yes’: In fact, there are many ways, even if we require in addition

$$P_i P_{i+1} P_i = \frac{1}{2} P_i, \quad P_{i+1} P_i P_{i+1} = \frac{1}{2} P_{i+1}$$

(“stationarity”).

### 3.4 From coin tossing to braids

One extension can be obtained as follows: Put

$$P := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad Q := \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

then they describe two coin tosses. Four coin tosses are described by

$$P_1 = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}, \quad P_2 = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

$$P_3 = \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & 1 \end{pmatrix}, \quad P_4 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & & \\ \frac{1}{2} & \frac{1}{2} & & \mathbf{0} \\ & & \frac{1}{2} & \frac{1}{2} \\ \mathbf{0} & & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

In general, from  $n$  coin tossing variables  $\bar{P}_1, \dots, \bar{P}_n \in M_k$  we obtain  $n + 2$  coin tossing variables in  $M_{2k}$ , denoted  $P_1, \dots, P_{n+2}$  as follows:

$P_{i+2}$  is the 2-left-amplification of  $\bar{P}_i$  for  $i \geq 2$ , i.e.,  $P_{i+2} = \mathbb{1}_2 \otimes \bar{P}_i$ .  
 $P_1$  and  $P_2$  are the  $k$ -right-amplifications of  $P$  and  $Q$ , i.e.,  $P_1 = P \otimes \mathbb{1}_k$ ,  
 $P_2 = Q \otimes \mathbb{1}_k$   
 $P_3$  is the  $(k - 2)$ -right-amplification of

$$\begin{pmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & 1 \end{pmatrix},$$

i.e.,  $P_3 = [P \otimes P + (\mathbb{1}_2 - P) \otimes (\mathbb{1}_2 - P)] \otimes \mathbb{1}_{k-2}$ .

This example, in a somewhat different notation, had been constructed with the techniques from ([Kue1], 4.3.10) by one of the authors in the middle eighties, and it took him a while to notice that the random variables  $P_j$  satisfy the relations

$$P_j P_{j+1} P_j = \frac{1}{2} P_j, \quad P_{j+1} P_j P_{j+1} = \frac{1}{2} P_{j+1} \quad \text{for all } j$$

$$P_l P_m = P_m P_l \quad \text{for } |l - m| \geq 2.$$

These relations are nowadays famous from the work of V.F.R. Jones (cf. [Jon]), where they first occurred in certain constructions with von Neumann algebras. He recognized that one should consider the matrices

$$\sigma_k := i \cdot P_k - (\mathbb{1} - P_k)$$

thus obtaining from the above the braid group relations

$$\sigma_j \sigma_{j+1} \sigma_j = \sigma_{j+1} \sigma_j \sigma_{j+1} \quad \text{for all } j, \quad \sigma_l \sigma_m = \sigma_m \sigma_l \quad \text{for } |l - m| \geq 2.$$

This bridge between von Neumann algebras and certain braid group representations leads quickly to Jones' polynomial invariants for knots. As is well known, this was the start shot for tremendous developments in mathematics and mathematical physics.

To understand the relation between these developments and quantum probability is a challenge for the near future. Some promising results exist already.

### 3.5 From coin tossing to nuclear magnetic resonance

Another way of extending the two polarizations in 3.3 to a coin tossing of arbitrary length is the following: As before, put

$$P := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad Q := \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

and, in addition,

$$P^\perp := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{1}_2 - P, \quad Q^\perp := \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} = \mathbb{1}_2 - Q.$$

From the  $n$  projections  $P_1, \dots, P_n$  in §3.2 describing the classical coin toss, construct  $2n + 2$  projections  $Q_1, \dots, Q_{2n+2} \in M_{2^{n+1}} = M_{2^n}(M_2)$ :

$Q_1 \in M_{2^n}(M_2)$  is the  $2^n$ -left-amplification of  $P$ , i.e.,  $Q_1 = \mathbb{1}_{2^n} \otimes P$ .

$Q_2 \in M_{2^n}(M_2)$  is the  $2^n$ -left-amplification of  $Q$ , i.e.,  $Q_2 = \mathbb{1}_{2^n} \otimes Q$ .

$Q_{2i+1}$  is the  $2^n \times 2^n$  block diagonal matrix with entries from  $M_2$  which has in its diagonal a  $P$  where  $P_i$  has a 1 and a  $P^\perp$  where  $P_i$  has a 0.

$Q_{2i+2}$  is the  $2^n \times 2^n$  block diagonal matrix with entries from  $M_2$  which has in its diagonal a  $Q$  where  $P_i$  has a 1 and a  $Q^\perp$  where  $P_i$  has a 0.

Thus, when denoting the random variables of this coin toss again by  $P_i$  instead of  $Q_i$ , four tossings are described by the matrices

$$P_1 = \begin{pmatrix} P & \\ & P \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & 1 & \\ & & & 0 \end{pmatrix}, \quad P_2 = \begin{pmatrix} Q & \\ & Q \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & & \\ \frac{1}{2} & \frac{1}{2} & & \\ & & \mathbf{0} & \\ \mathbf{0} & & \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

$$P_3 = \begin{pmatrix} P & \\ & P^\perp \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & 1 \end{pmatrix},$$

$$P_4 = \begin{pmatrix} Q & \\ & Q^\perp \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & & \\ \frac{1}{2} & \frac{1}{2} & & \\ & & \mathbf{0} & \\ \mathbf{0} & & \frac{1}{2} & -\frac{1}{2} \\ & & -\frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

while six tossings are described by the block-diagonal matrices

$$P_1 = \begin{pmatrix} P & & & \\ & P & & \\ & & P & \\ & & & P \end{pmatrix}, \quad P_2 = \begin{pmatrix} Q & & & \\ & Q & & \\ & & Q & \\ & & & Q \end{pmatrix},$$

$$P_3 = \begin{pmatrix} P & & & \\ & P & & \\ & & P^\perp & \\ & & & P^\perp \end{pmatrix}, P_4 = \begin{pmatrix} Q & & & \\ & Q & & \\ & & Q^\perp & \\ & & & Q^\perp \end{pmatrix},$$

$$P_5 = \begin{pmatrix} P & & & \\ & P^\perp & & \\ & & P & \\ & & & P^\perp \end{pmatrix}, P_6 = \begin{pmatrix} Q & & & \\ & Q^\perp & & \\ & & Q & \\ & & & Q^\perp \end{pmatrix}.$$

It is immediate, that these projections satisfy the relations

$$P_i P_j P_i = \frac{1}{2} P_i, \quad \text{if } |i - j| \text{ is odd,}$$

$$P_i P_j = P_j P_i, \quad \text{if } |i - j| \text{ is even.}$$

Also this example can be constructed by systematic methods. It was obtained from the construction scheme developed in ([Kue1], 4.2.2). From this construction a physical interpretation can be deduced:

The family of random variables describes the z-component of a spin- $\frac{1}{2}$ -particle in an exterior magnetic field. The field is stochastic in time and takes one of two possible values at each time unit. For a detailed description of this system we refer to [Kue2]. Models of spins in stochastic magnetic fields are part of the theory of spin relaxation which is at the basis of nuclear magnetic resonance (NMR) (cf. [Sli]).

This model can also be translated into a model for photons: It describes polarization of a photon propagating through a medium, where the medium changes the polarization of the photon at random. An example would be a sugar solution whose concentration varies at random in space.

#### 4. APPENDIX: AN OUTLOOK ONTO THE FULL MATHEMATICAL MODEL

In this section we give a brief indication of the steps necessary in order to extend our considerations so far to the full mathematical model of quantum probability, which is at the basis of most publications in this field. It should be noted, however, that once one has accepted the picture we have developed up to now, the next steps are only straightforward extensions. They are necessary in order to describe more general systems (e.g., the extension to real-valued random variables and infinite dimensions) or convenient in order to obtain a smooth mathematical framework (e.g., using random variables in an operator algebra). A more detailed description of these steps is planned as a continuation of the present paper.

## 4.1 Real-valued random variables

So far we have considered only 0-1-valued random variables. The mathematical body of quantum mechanics suggests to describe real-valued random variables as follows:

A real-valued random variable is a self-adjoint matrix  $A$  on  $\mathbb{C}^n$ . In quantum mechanics this is called an observable. Its eigenvalues are the possible outcomes of a measurement of this random variable. A state is still described by a unit vector or, more generally, a density matrix  $\Phi$ . If  $P_i$  is the orthogonal projection onto the eigenspace of an eigenvalue  $\lambda_i$  of  $A$ , then the probability for a measurement of  $A$  to yield the eigenvalue  $\lambda_i$  is given by  $p_i = \text{tr}(\Phi \cdot P_i)$ . Therefore, attributing this probability to the eigenvalue  $\lambda_i$ , the state  $\Phi$  induces a probability distribution on the set of all eigenvalues:

$$\sum_i p_i = \sum_i \text{tr}(\Phi \cdot P_i) = \text{tr}(\Phi \cdot \sum_i P_i) = \text{tr}(\Phi) = 1.$$

It follows that the expectation of a measurement of  $A$  on a system in the state  $\Phi$  is given by

$$\mathbb{E}(A) = \text{tr}(\Phi \cdot A).$$

As in §2.3, the set of real-valued random variables on a (finite) classical probability space is identified with the set of all diagonal matrices with real entries. Conversely, given any set of pairwise commuting self-adjoint matrices, they can be diagonalized simultaneously and thus have an interpretation as a set of classical random variables.

## 4.2 Towards infinite dimensions

Already from classical probability it is clear, that models on finite dimensional spaces are not sufficient. For many asymptotic considerations one should be able to handle infinitely many different (e.g., independent) random variables simultaneously. Moreover, random variables allowing an infinite number of possible outcomes are necessary. Therefore, we must take models into consideration on Hilbert spaces  $\mathcal{H}$  of infinite dimension. Then, pure states are still described by unit vectors. For mixed states, however, it is convenient to allow infinite convex combinations of pure states. They are described by positive trace class operators on  $\mathcal{H}$  with trace 1 (normalization). These are the self-adjoint operators on  $\mathcal{H}$  with discrete spectrum, whose eigenvalues are non-negative and sum up to 1 when counted with multiplicity. Trace class operators are alternatively called ‘nuclear operators’.

Real-valued random variables are now self-adjoint operators on  $\mathcal{H}$ . For their theory the spectral theorem is crucial. The spectrum of a self-adjoint operator  $A$  is real and we still identify the real numbers in the spectrum with the possible



outcomes of a measurement. With any unit vector and, more generally, with any positive normalized trace class operator  $\Phi$ , the spectral theorem associates a probability measure  $\mu_\Phi$  on the spectrum of  $A$ . Namely, if  $I$  is any measurable subset of the spectrum of  $A$ , and  $P_I$  is the spectral projection of  $A$  corresponding to this subset (i.e.,  $P_I = \chi_I(A)$  by the functional calculus, where  $\chi_I$  is the characteristic function of  $I$ ), then the probability for a measurement of  $A$  to yield a value in  $I$  is  $\mu_\Phi(I) = \text{tr}(\Phi \cdot P_I)$ . By the properties of trace class operators, the right hand side of this equation gives a well defined finite value. Again, the expectation of the outcome of a measurement of  $A$  on a system in the state  $\Phi$  has the expectation  $\text{tr}(\Phi \cdot A)$ , but now this may be infinite or even ill-defined. Clearly, this model generalizes the finite dimensional model.

Classical probability is incorporated into this model as follows: Given a classical probability space  $(\Omega, \Sigma, \mu)$  and a real-valued random variable  $X$  on it, we form the Hilbert space  $\mathcal{H} := L^2(\Omega, \Sigma, \mu)$ . The probability measure  $\mu$  can be translated into the rank-one trace class operator  $\Phi_{\mathbb{1}}$  (orthogonal projection onto the subspace spanned by the vector  $\mathbb{1}$ ), where  $\mathbb{1}$  is the constant function on  $\Omega$  with value 1. The random variable  $X$  is turned into the self-adjoint (possibly unbounded) operator multiplying each function in  $L^2(\Omega, \Sigma, \mu) = \mathcal{H}$  pointwise by  $X$ . All probabilistic considerations can be easily transferred into this mathematically equivalent model.

Conversely, given a positive normalized trace class operator  $\Phi$ , on any Hilbert space  $\mathcal{H}$ , and a set of pairwise commuting self-adjoint operators, there always exists, although this is a non-trivial matter to prove, a probability space  $(\Omega, \Sigma, \mu)$  and a unitary map between  $\mathcal{H}$  and  $L^2(\Omega, \Sigma, \mu)$  carrying the self-adjoint operators into multiplication by real functions in such a way that all probabilities are preserved. Therefore, it remains true that classical probability corresponds to sets of pairwise commuting random variables.

### 4.3 Towards an algebraic description

Although many considerations in quantum probability can be done and are done within the model described so far, there are reasons to come to a more algebraic description of quantum probability. We mention two of them:

1. It is often difficult to find naively a realization of an infinite system on some Hilbert space. As a typical example, the reader may try to extend the coin tossing model in Section 3.4 to an infinite number of tosses.

2. So far we only mentioned that specification of the set of random variables is part of the description of a mathematical model (Section 2.3). However we did not comment on the question, on what kinds of sets of random variables mathematical models should be based. Without any further specification the mathematics would become extremely cumbersome, and indeed to a great extent it is just mathematical convenience, which leads to certain additional requirements on the set of all random variables of a model.

To start with, it is generally assumed that the set of all random variables of a model forms a linear space. Second, almost everyone assumes that it is — or can be realized as — the set of all self-adjoint operators lying in (or being attached to) some  $*$ -algebra of operators on a Hilbert space. (Clearly, domain questions have to be taken into consideration). It is often convenient to assume, that this  $*$ -algebra of operators consists of bounded operators and is closed in the operator norm, i.e., is a concrete  $C^*$ -algebra; or that it is even closed in the strong operator topology, so that it becomes a von Neumann algebra. This makes available the rich theory of these algebras. Unbounded random variables can still be realized as unbounded functions of bounded self-adjoint operators.

Von Neumann algebras are particularly useful, when one wants to make use of the spectral theorem for selfadjoint operators, since all spectral projections of such an operator lie in the same von Neumann algebra. From the foregoing it should be clear, that models from classical probability correspond to commutative algebras and indeed, every commutative von Neumann algebra is isomorphic to some  $L^\infty(\Omega, \Sigma, \mu)$ . Traditional quantum mechanics uses the algebra  $B(\mathcal{H})$  of all bounded operators on some Hilbert space  $\mathcal{H}$ . Finite dimensional  $C^*$ -algebras (which are also von Neumann algebras), are automatically isomorphic to a direct sum of copies of full matrix algebras. The infinite version of the example in Section 3.4, however, leads to a different type of von Neumann algebras, as it lives naturally in the hyperfinite factor of type  $II_1$ . The example in Section 3.5 can be realized in the algebra of all functions on a classical probability space with values in the  $2 \times 2$ -matrices.

In this algebraic picture, a (mixed) state on a  $C^*$ -algebra  $\mathcal{A}$  can be abstractly characterized as a linear functional  $\phi : \mathcal{A} \rightarrow \mathbb{C}$  which is positive, i.e., takes positive values on the positive selfadjoint operators in  $\mathcal{A}$  and is normalized, i.e., has norm 1. If  $\mathcal{A}$  is a von Neumann algebra, it is convenient to assume in addition that  $\phi$  is even continuous in the ultrastrong operator topology. It is then normal, i.e., it is completely additive on sets of pairwise orthogonal projections. Such a state on a von Neumann algebra  $\mathcal{A} \subseteq B(\mathcal{H})$  can always be realized as  $\phi(x) = \text{tr}(\Phi \cdot x)$  for  $x \in \mathcal{A}$  and  $\Phi$  some positive normalized trace class operator on  $\mathcal{H}$ . More generally, if  $\phi$  is a state on a  $C^*$ -algebra  $\mathcal{A}$ , there is a representation of  $\mathcal{A}$  as a concrete  $C^*$ -algebra of operators on a — possibly different — Hilbert space, such that the state  $\phi$  can be realized by a unit vector in this Hilbert space (GNS-representation).

Finally, the notion of a random variable is to be generalized. Although in many situations it is enough to consider real-valued random variables, and a large part of the literature concentrates on this case, there is also a need for considering random variables with a different state space. In classical probability, a Brownian particle in three dimensions would be a typical example. It is described by three random variables, one for each coordinate. In the mathematics of classical probability, there is not much difference between random variables with different state spaces. In general, however, a set of self-adjoint operators describing different quantities of a physical system under observation, do not need to commute

with each other. The polarization of a photon in different directions, which we discussed extensively in this paper, or the spin in different directions of a particle, are typical examples of such a situation. In these situations one is led to consider not only one self-adjoint operator at a time, but instead a set of possibly non-commuting self-adjoint operators, constituting the description of a single random variable. The same considerations as above led people to assume, that this set comes from a  $*$ -algebra. If one takes this point of view, then even one self-adjoint operator is already replaced by the commutative algebra of functions of this operator, closed in a suitable topology. This leads to the notion of a random variable as a subalgebra of operators contained in the larger algebra which is generated by all random variables of the model. Finally, in order to keep track of the variation in time of a single operator with a particular interpretation — say the horizontal polarization of a photon — the following formulation is more convenient:

The quantities of an observed system form a  $*$ -algebra  $\mathcal{A}_0$  and a random variable identifies  $\mathcal{A}_0$  with a subalgebra of the algebra  $\mathcal{A}$  generated by all real-valued random variables of the model. Thus a random variable is an injective  $*$ -homomorphism from  $\mathcal{A}_0$  into  $\mathcal{A}$ . A stochastic process is a family of such random variables [AFL]. This is the starting point of many papers in quantum probability.

For more information we refer to the two books [Mey] and [Par], already mentioned and also to our forthcoming paper, developing these ideas further and illustrating them by more examples.

#### REFERENCES

- [AFL] L. Accardi, F. Frigerio, J.T. Lewis: Quantum stochastic processes. Publ. RIMS 18 (1982), 97 - 133.
- [Asp] A. Aspect, J. Dalibard, G. Roger: Experimental test of Bell's inequalities using time-varying analyzers. Phys. Rev. Lett. 49 (1982), 1804 - 1807.
- [Bel] J.S. Bell: On the Einstein-Podolsky-Rosen paradox. Physics 1 (1965) 195.
- [Boh] D. Bohm: A suggested interpretation of quantum theory in terms of 'hidden variables'. Phys. Rev. 85 (1952) 180.
- [Jon] V.F.R. Jones: "Subfactors and Knots". CBMS-Lectures 80. American Mathematical Society 1991.
- [Kue1] B. Kümmerer: Markov dilations on  $W^*$ -algebras. J. Funct. Anal. 63 (1985), 139 - 177.
- [Kue2] B. Kümmerer: Examples of Markov dilations over the  $2 \times 2$ -matrices. In L. Accardi, A. Frigerio, V. Gorini (Eds), "Quantum Probability and Applications to the Quantum Theory of Irreversible

- Processes". Lecture Notes in Mathematics 1055, Springer-Verlag, Berlin-Heidelberg-New York 1984, 228 – 244.
- [Mey] P. A. Meyer: "Quantum Probability for Probabilists". Lecture Notes in Mathematics 1538, Springer-Verlag, Berlin Heidelberg 1993.
- [Neu] J. v. Neumann: "Mathematische Grundlagen der Quantenmechanik", Reprinted by Springer-Verlag, Berlin-Heidelberg-New York 1981.
- [Par] K.R. Parthasarathy: "An Introduction to Quantum Stochastic Calculus". Birkhäuser Verlag, Basel 1992.
- [Sli] C.P. Slichter: "Principles of Magnetic Resonance". Springer-Verlag Berlin Heidelberg New York 1980.

Burkhard Kümmeler  
 Mathematisches Institut A  
 Universität Stuttgart  
 Pfaffenwaldring 57  
 D - 70 569 Stuttgart  
 Germany

Hans Maassen  
 Mathematisch Instituut  
 Katholieke Universiteit  
 Toernooiveld 1  
 6525 ED Nijmegen  
 Netherlands