Linear representations of probabilistic transformations induced by context transitions

2001 J. Phys. A: Math. Gen. 34 9965

(http://iopscience.iop.org/0305-4470/34/47/304)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **34** (2001) 9965–9981 PII: S0305-4470(01)25808-9

Linear representations of probabilistic transformations induced by context transitions

Andrei Khrennikov

International Centre for Mathematical Modelling in Physics and Cognitive Sciences, MSI, University of Växjö, S-35195 Sweden

E-mail: Andrei.Khrennikov@msi.vxu.se

Received 30 May 2001, in final form 14 September 2001
Published 16 November 2001
Online at stacks.iop.org/JPhysA/34/9965

Abstract

By using straightforward frequency arguments we classify transformations of probabilities which can be generated by transition from one preparation procedure (context) to another. There are three classes of transformations corresponding to statistical deviations of different magnitudes: (a) trigonometric; (b) hyperbolic; (c) hyper-trigonometric. It is shown that not only quantum preparation procedures can have trigonometric probabilistic behaviour. We propose generalizations of C-linear space probabilistic calculus to describe non-quantum (trigonometric and hyperbolic) probabilistic transformations. We also analyse the superposition principle in this framework.

PACS numbers: 05.30.-d, 03.65.Ta, 03.65.Yz

1. Introduction

We analyse a well known expression for probability of an event in terms of the conditional probabilities based on another event. This expression often goes by the name Bayes' formula and is known not to apply in quantum mechanics when probabilities for incompatible, or non-commuting, observables are being evaluated, see e.g. [1–3].

In the classical case we have Bayes' formula:

$$p(A = a_i) = p(C = c_1)p(A = a_i/C = c_1) + p(C = c_2)p(A = a_i/C = c_2) \tag{1}$$

where $A = a_1, a_2$ and $C = c_1, c_2$ are two dichotomic random variables. In the quantum case we have the formula

$$p(A = a_i) = p(C = c_1)p(A = a_i/C = c_1) + p(C = c_2)p(A = a_i/C = c_2) \pm 2\sqrt{p(C = c_1)p(A = a_i/C = c_1)p(C = c_2)p(A = a_i/C = c_2)}\cos\theta \tag{2}$$

where θ is some phase.

The appearance of the interference term in quantum modification of Bayes' formula has led to the use of the term 'quantum probability' in contradiction to what could be called 'regular' or 'classical' probability (see e.g. [1–21] for extended discussions on this problem), but there is only one type of physical probability and it is one that is subject to measurement via counting and the generation of relative frequencies. It is the relative frequency probability (of von Mises) that is directly connected with data from experiment. We provide a frequency probabilistic analysis making a contribution to the understanding of probability and Bayes' formula within the context of quantum mechanics.

Our analysis begins with the relative frequency definition of the relevant probabilities. The probability for the eigenvalue of one observable is then expressed in terms of the conditional probabilities involving the eigenvalues of a second (in general incompatible) observable. In this way the 'non-classical' term in the quantum Bayes formula is shown as a perturbation due to the difference in preparation procedures of the different states. The perturbing term is then expressed in terms of a coefficient λ whose absolute value can be less than or equal to one, or it can be greater than one for each of the eigenvalues of the observable. This range of values for the coefficient λ then introduces three distinct types of perturbation which are called *trigonometric, hyperbolic and hyper-trigonometric*. Each case is then examined separately. Classical and quantum cases are then special cases of more general results. It is then shown that in the quantum case it is possible to reproduce a Hilbert space in which the probabilities are found in the usual way, but there is a case in which this is not possible; though the space is linear it is not a Hilbert space. In general it is not a complex linear space. In the case of hyperbolic probabilistic behaviour we have to use linear representation of probabilities over so-called hyperbolic numbers.

In fact, our approach to experimental probabilities is nothing other than the well known contextualist approach. In quantum theory such an approach was strongly supported by Bohr [11] and Heisenberg [3] (see also [5–21]). Heisenberg [3] already directly pointed out that quantum interference of probabilistic alternatives, (2), is a consequence of transition from one context (complex of physical conditions) to another. Bohr always pointed out that we have to take into account the experimental arrangement to determine 'quantum probabilities.' In this paper we took the important step to justify the contextualist approach: we derived the 'quantum probabilistic rule' in a purely classical frequency probabilistic framework (see also [22, 23]). In fact, we obtained much more than planned. We found that possible modifications of Bayes' formula (induced by context transitions) are not reduced to the 'quantum probabilistic rule' (see the above discussion).

As already mentioned, it is impossible to realize probabilistic transformations induced by all possible context transitions in a Hilbert space. We develop the non-Hilbert linear space probabilistic formalism. The main distinguishing feature of this formalism is *the violation of the superposition principle*. There is no more superposition transitivity: combination of two superpositions need not be again a superposition. The principle of superposition is the cornerstone of quantum formalism. There is still a large diversity of opinions on this principle. It may be that our models, in that the principle of superposition is violated, may be useful for analysis of this principle[1]. We note that there is a similarity with the quantum formalism based on the theory of POVM (positive operator-valued measures) (see, e.g., [6, 16–18, 24]) in that it is possible to consider non-orthogonal expansions of the unit operator.

[1] In our models the situation is not like that in models with superselection rules. 'Probability superselection' could not be represented by choosing a linear subspace in the Hilbert space of quantum states. In some sense our selections can be considered as nonlinear 'superselections'.

The results of this paper were presented in the authors' talks at the international conferences *Foundations of Probability and Physics* (Växjö, 2000), *Quantum Physics: Reconsideration of Foundations* (Växjö, 2001) and *Exploring Quantum Physics* (Venice, 2001).

2. Classification of transformation rules for probability distributions for three preparation procedures

Let $\mathcal{E}$ be some preparation procedure (see [5–10, 13]) that produce physical systems having two properties A and C. These properties are described by dichotomic variables $A = a_1$, a_2 and $C = c_1, c_2$. We fix the number of preparation acts, $N \equiv N_{\mathcal{E}}$, so $\mathcal{E}$ always produces ensembles $S = S_{\mathcal{E}}$ having $N = |S|$ elements.

Let $\mathcal{E}_1$ and $\mathcal{E}_2$ be two other preparation procedures. It is assumed that each of these preparation procedures can be applied to elements of S. By application of $\mathcal{E}_i$ to S we produce a new statistical ensemble[2] $S_i, i = 1, 2$. The main feature of the ensemble S_i is that $C = c_i$ for its elements ($i = 1, 2$). For example, $\mathcal{E}_i$ can be considered as filters with respect to the property C : $\mathcal{E}_i$ select elements of S such that $C = c_i$ ($i = 1, 2$). Such a filtration justifies the assumption that the number of elements in S_i could be chosen equal to the number of elements, N_i, in S having the property $C = c_i$ ($i = 1, 2$). So everywhere below

$$|S_i| = N_i \qquad i = 1, 2.$$

The crucial point of our considerations is that *in general we could not 'select', for example, elements with the property $C = c_1$ without disturbing the property A*. In general the sub-ensemble

$$S_{ij} = \{s \in S_i : A = a_j\}$$

of the ensemble S_i does not coincide with the sub-ensemble

$$S_{ij}^{(0)} = \{s \in S : C = c_i, A = a_j\}$$

of the original ensemble S. We set

$$n_{ij} = |S_{ij}^{(0)}| \qquad \text{and} \qquad m_{ij} = |S_{ij}|$$

(the numbers of elements in the sub-ensembles) and

$$N_i = |\{s \in S : C = c_i\}| \qquad n_j = |\{s \in S : A = a_j\}|$$

(the numbers of elements in S having, respectively, properties $C = c_i, i = 1, 2$ and $A = a_j, j = 1, 2$). We note that everywhere below the first number, i, in the index pair ij is related to the property C and the second one, j, to the property A. We shall use the frequency approach to probability (see, [25] and [19]): the probability is defined as the limit of relative frequencies when the number of trials $N \to \infty$.

Remark (foundations of probability and physics). As we have already discussed [19], the conventional probability theory based on Kolmogorov axiomatics [26] is not the best tool to work with 'quantum probabilities'. The formal use of an abstract, absolute, probability measure is the source of many misunderstandings. In particular, the Kolmogorov model is not the best one for operating with transitions from one context to another. In fact, all probabilities are conditional probabilities; there is no absolute probability (see [27] for the extended discussion). We prefer to work with frequency probabilities. Here contexts are described by collectives (random sequences) that are used to find relative frequencies. However, in this paper we shall not pay much attention to the mathematical details of the frequency framework (see [19] for

[2] In general we need two different ensembles $S_{\mathcal{E}}$ to produce two ensembles, S_1 and S_2.

the details). In fact, everybody who is familiar with von Mises frequency probability theory could recognize that in this paper we work with von Mises collectives. These collectives are produced by different preparation procedures (complexes of physical conditions). It may be even better to use the term 'collective' instead of the term 'ensemble' that is used in this paper. However, we are a little afraid to do this, because there is a rather strong prejudice against von Mises' approach (especially from the mathematical side).

We consider relative frequencies:

$$q_j^{(N)} \equiv p_j^a(N) = \frac{n_j}{N} \qquad p_i^{(N)} \equiv p_i^c(N) = \frac{N_i}{N}$$

(for the properties A and C in the ensembles prepared by $\mathcal{E}$);

$$p_{ij}^{a/c}(N_i) \equiv p_{ij}(N_i) = \frac{m_{ij}}{N_i}$$

(for the property $A = a_j$ in the ensemble prepared by $\mathcal{E}_i$) and the corresponding probabilities:

$$q_j \equiv p_j^a = \lim_{N\to\infty} p_j^a(N) \qquad p_i \equiv p_i^c = \lim_{N\to\infty} p_i^c(N) \qquad p_{ij} \equiv p_{ij}^{a/c} = \lim_{N_i\to\infty} p_{ij}(N).$$

As in general n_{ij} are not equal to m_{ij} (even asymptotically $N \to \infty$), we do not have the conventional formula of total probability. In general

$$\begin{aligned} q_j &= p_S(A = a_j) \neq p_1 p_{11} + p_2 p_{21} \\ &= p_S(C = c_1) p_{S_1}(A = a_1) + p_S(C = c_2) p_{S_2}(A = a_2). \end{aligned}$$

We want to investigate the various forms probabilities q_j can take, depending on perturbations induced by context transitions. In the general case we have

$$\begin{aligned} q_j(N) &= \frac{n_j}{N} = \frac{n_{1j}}{N} + \frac{n_{2j}}{N} = \frac{m_{1j}}{N} + \frac{m_{2j}}{N} + \delta_j^{(N)} \\ &= \frac{N_1}{N} \cdot \frac{m_{1j}}{N_1} + \frac{N_2}{N} \cdot \frac{m_{2j}}{N_2} + \delta_j^{(N)} \\ &= p_1(N) p_{1j}(N_1) + p_2(N) p_{2j}(N_2) + \delta_j^{(N)} \end{aligned}$$

where the perturbation term (which appears due to the transition from S to S_1 and S_2) has the form

$$\delta_j^{(N)} \equiv \delta_j(\mathcal{E}, \mathcal{E}_1, \mathcal{E}_2, N) = \frac{1}{N}[(m_{1j} - n_{1j}) + (m_{2j} - n_{2j})].$$

We remark that there exists the limit

$$\delta_j = \lim_{N\to\infty} \delta_j(N) = q_j - (p_1 p_{1j} + p_2 p_{2j}).$$

Thus in general we have $q_j = p_1 p_{1j} + p_2 p_{2j} + \delta_j$, where

$$\delta_j = \lim_{N\to\infty} \frac{1}{N}[(m_{1j} - n_{1j}) + (m_{2j} - n_{2j})].$$

It is useful to perform normalization by setting

$$\delta_j = 2\sqrt{p_1 p_{1j} p_2 p_{2j}}\lambda_j \qquad j = 1, 2.$$

The trivial (but important) remark is that there are three possibilities:

(T) $|\lambda_j| \leqslant 1$
(H) $|\lambda_j| > 1$
(HT) $|\lambda_1| \leqslant 1$ and $|\lambda_2| > 1$ or $|\lambda_1| > 1$ and $|\lambda_2| \leqslant 1$.

In case (T) we can always represent the coefficient as $\lambda_j = \cos\theta_j$, $j = 1, 2$; in case (H), as $\lambda_j = \pm\cosh\theta_j$, $j = 1, 2$; in case (HT), as $\lambda_1 = \cos\theta_1$ and $\lambda_2 = \pm\cosh\theta_2$ or vice versa. Probabilistic behaviours of the types (T), (H) and (HT) will be called *trigonometric, hyperbolic and hyper-trigonometric* behaviours, respectively.

We have studied the general case. There are three preparation procedures $\mathcal{E}$, $\mathcal{E}_1$ and $\mathcal{E}_2$ such that $\mathcal{E}_1$ and $\mathcal{E}_2$ are selections with respect to values $C = c_1$ and c_2. The general probabilistic transformation induced by transitions $\mathcal{E} \to \mathcal{E}_j$, $j = 1, 2$, has the form

$$p_j^a = p_1^c p_{1j}^{a/c} + p_2^c p_{2j}^{a/c} \pm 2\sqrt{p_1^c p_{1j}^{a/c} p_2^c p_{2j}^{a/c}}\lambda_j \tag{3}$$

where $\lambda_j = \cos\theta_j$ or $\lambda_j = \cosh\theta_j$, or $\lambda_1 = \cos\theta_1$ and $\lambda_2 = \pm\cosh\theta_2$ or vice versa. Here the coefficient λ_j gives the normalized statistical measure of the perturbations of A due to the transition $\mathcal{E} \to (\mathcal{E}_1, \mathcal{E}_2)$:

$$\lambda_j = \lim_{N\to\infty} \lambda_j^{(N)} \qquad \text{where} \quad \lambda_j^{(N)} = \frac{1}{2\sqrt{m_{1j}m_{2j}}}[n_{1j} - m_{1j}) + (n_{2j} - m_{2j})]. \tag{4}$$

If these perturbations are relatively small, namely $|\lambda_j| \leqslant 1$, $j = 1, 2$, then we observe T-behaviour; in particular, classical and quantum behaviours. If these perturbations are relatively large, namely $|\lambda_j| > 1$, $j = 1, 2$, then we observe H-behaviour. In fact, we can continuously transfer T-behaviour into H-behaviour, since λ_j, $|\lambda_j| = 1$, has both T- and H-representations: $\lambda_j = \pm\cos 0 = \pm\cosh 0$. If one of these perturbations, for instance λ_1, is relatively small, namely $|\lambda_1| \leqslant 1$, and another, λ_2, is relatively large, namely $|\lambda_2| > 1$, then we observe HT-behaviour.

Finally, we show that coefficients λ_1 and λ_2 are connected by a 'condition of orthogonality' (in the quantum formalism this is the real condition of orthogonality in the complex Hilbert space). We note that the matrix of probabilities $P = (p_{ij})$ is always a stochastic matrix:

$$p_{11} + p_{12} = 1 \qquad \text{and} \qquad p_{21} + p_{22} = 1 \tag{5}$$

(because $p_{i1} + p_{i2} = p_{S_i}(A = a_1) + p_{S_i}(A = a_2) = 1$). Thus we have

$$1 = p_1^a + p_2^a = p_1^c p_{11}^{a/c} + p_2^c p_{21}^{a/c} + p_1^c p_{12}^{a/c} + p_2^c p_{22}^{a/c} + 2\sqrt{p_1^c p_{11}^{a/c} p_2^c p_{22}^{a/c}}\lambda_1 + 2\sqrt{p_1^c p_{12}^{a/c} p_2^c p_{22}^{a/c}}\lambda_2.$$

To simplify considerations, we assume everywhere that all probabilities are strictly positive. This implies

$$\sqrt{p_{11}^{a/c} p_{21}^{a/c}}\lambda_1 + \sqrt{p_{12}^{a/c} p_{22}^{a/c}}\lambda_2 = 0. \tag{6}$$

We set

$$K = \sqrt{\frac{p_{12}^{a/c} p_{22}^{a/c}}{p_{11}^{a/c} p_{21}^{a/c}}}.$$

We obtain

$$\lambda_1 = -K\lambda_2.$$

We observe that probabilities p_j^c are not involved in the condition of orthogonality (6). In particular, in the T-case we always have

$$\cos\theta_1 = -K\cos\theta_2 \tag{7}$$

in the H-case we have

$$\cosh\theta_1 = K\cosh\theta_2 \tag{8}$$

(here $\lambda_1 = \pm\cosh\theta_1$ and $\lambda_2 = \mp\cosh\theta_2$).

In the HT-case we have

$$\cos\theta_1 = \pm K\cosh\theta_2 \qquad \text{or} \qquad \cosh\theta_1 = \pm K\cos\theta_2. \tag{9}$$

Finally, we remark that all the above considerations can be easily generalized to non-dichotomic variables: $A = a_1, \ldots, a_M$ and $C = c_1, \ldots, c_M$.

In this case the probability p_i^a can always be represented as

$$p_i^a = \sum_{j=1}^{M} p_j^c p_{ji}^{a/c} + 2\sum_{k<l}\sqrt{p_k^c p_l^b p_{ki}^{a/c} p_{li}^{a/c}}\,\lambda_{kl}^{(i)} \tag{10}$$

where the coefficients $\lambda_{kl}^{(i)} = \frac{\delta_{kl}^{(i)}}{2\sqrt{p_k^c p_l^c p_{ki}^{a/c} p_{li}^{a/c}}}$ and $\delta_{kl}^{(i)} = \frac{1}{M-1}\left[p_k^c\left(p_i^a - p_{ki}^{a/c}\right) + p_l^c\left(p_i^a - p_{li}^{a/c}\right)\right]$. Coefficients $\{\lambda_{kl}^{(i)}\}$ are normalized statistical deviations that arise due to the transition from the context determined by $\mathcal{E}$ to contexts $\mathcal{E}_j$.

To simplify analysis, we shall consider only dichotomic variables in the following sections.

3. Trigonometric probabilistic behaviour: classical, quantum and non-classical/quantum physics

In this section we consider probabilistic transformations for preparation procedures that produce relatively small statistical deviations:

$$|\lambda_j| \leqslant 1 \qquad j = 1, 2.$$

3.1. Classical probabilistic behaviour

Suppose that we can construct statistically 'perfect' preparation procedures $\mathcal{E}_1, \mathcal{E}_2$: selections of elements of the ensemble S with respect to values $C = c_1$ and c_2 produce statistically negligible changes of A. We set

$$\Delta_{ij}(N) = n_{ij} - m_{ij}.$$

Here n_{ij} is the number of elements of S having $C = c_i$ and $A = a_j$ and m_{ij} is the number of elements of S_i having $A = a_j$. The classical probabilistic behaviour is characterized by the condition

$$\lim_{N\to\infty}\frac{\Delta_{ij}(N)}{N} = 0 \qquad \text{for all} \quad i, j.$$

Here both $\lambda_j = 0$ and we have conventional rule (1).

3.2. Quantum probabilistic behaviour

Let us consider preparations which induce symmetric statistical deviations:

$$|\lambda_1| = |\lambda_2|. \tag{11}$$

Thus the coefficient K is equal to unity, so $p_{12}p_{22} = p_{11}p_{21}$. In the two-dimensional case this condition is equivalent to the well known condition of *double stochasticity*:

$$p_{11} + p_{21} = 1 \qquad p_{12} + p_{22} = 1. \tag{12}$$

Thus $p_{S_1}(A = a_1) + p_{S_2}(A = a_1) = 1$ and $p_{S_1}(A = a_2) + p_{S_2}(A = a_2) = 1$. These are 'conservation laws' for the A in the process of splitting of the ensemble S into ensembles S_1

and S_2. We also remark that (7) implies that $\cos\theta_1 = -\cos\theta_2$, so $\theta_2 = \theta_1 + \pi \pmod{2\pi}$. Thus we have the probabilistic transformations:

$$q_1 (\equiv p_1^a) = p_1 p_{11} + p_2 p_{21} + 2\sqrt{p_1 p_{11} p_2 p_{21}} \cos\theta \tag{13}$$

$$q_2 (\equiv p_2^a) = p_1 p_{12} + p_2 p_{22} - 2\sqrt{p_1 p_{12} p_2 p_{22}} \cos\theta. \tag{14}$$

This is the well known quantum probabilistic transformation. We now find complex representations of these probabilities that would linearize transformations (13) and (14). We use the well known formula:

$$A + B \pm 2\sqrt{AB}\cos\theta = |\sqrt{A} \pm \sqrt{B}\mathrm{e}^{\mathrm{i}\theta}|^2. \tag{15}$$

Thus

$$q_1 = |\sqrt{p_1}\sqrt{p_{11}} + \sqrt{p_2}\sqrt{p_{21}}\mathrm{e}^{\mathrm{i}\theta_1}|^2 \qquad q_2 = |\sqrt{p_1}\sqrt{p_{21}} + \sqrt{p_2}\sqrt{p_{22}}\mathrm{e}^{\mathrm{i}\theta_2}|^2$$

(in the quantum case $\theta_1 = \theta_2 + \pi$). These formulae can also be derived by $\boldsymbol{C}$-linear space computations. We represent the preparation procedure $\mathcal{E}$ by a vector φ in the two-dimensional complex Hilbert space:

$$\varphi = \sqrt{p_1}\varphi_1 + \sqrt{p_2}\mathrm{e}^{\mathrm{i}\theta}\varphi_2$$

where $\{\varphi_1, \varphi_2\}$ is an orthonormal basis corresponding to the physical observable C (the condition $p_1 + p_2 = 1$ implies that $\|\varphi\|^2 = 1$). Let ψ_1, ψ_2 be an orthonormal basis corresponding to the physical observable A. We have

$$\varphi_1 = \sqrt{p_{11}}\psi_1 + \mathrm{e}^{\mathrm{i}\gamma_1}\sqrt{p_{12}}\psi_2 \qquad \varphi_2 = \sqrt{p_{21}}\psi_1 + \mathrm{e}^{\mathrm{i}\gamma_2}\sqrt{p_{22}}\psi_2.$$

We remark that orthogonality of φ_1 and φ_2 is, in fact, equivalent to the condition of double stochasticity for $P = (p_{ij})$ and the relation $\gamma_2 = \gamma_1 + \pi \pmod{2\pi}$. By expanding φ with respect to the basis $\{\psi_1, \psi_1\}$ we obtain

$$\varphi = d_1\psi_1 + d_2\psi_2$$

where

$$d_1 = \sqrt{p_1}\sqrt{p_{11}} + \mathrm{e}^{\mathrm{i}\theta}\sqrt{p_2 p_{21}} \qquad d_2 = \mathrm{e}^{\mathrm{i}\gamma_1}\sqrt{p_1}\sqrt{p_{12}} + \mathrm{e}^{\mathrm{i}(\gamma_2+\theta)}\sqrt{p_2 p_{22}}. \tag{16}$$

By using the relation $\gamma_2 = \gamma_1 + \pi$ we reproduce the quantum probabilistic rule (13), (14).

We note that our considerations have demonstrated that the main distinguishing feature of quantum formalism is not the presence of the $\cos\theta$-factor in the 'quantum transformation of probabilities', but the double stochasticity of the matrix $P = (p_{ij}^{a/c})$ of transition probabilities and the relation

$$\gamma_2 = \gamma_1 + \pi \tag{17}$$

between phases in the expansions of φ_1 and φ_2 with respect to the basis $\{\psi_1, \psi_2\}$.

The 'double-stochasticity conservation laws', (12), and the 'phase conservation law', (17), imply the unitarity of the transformation U connecting $\{\varphi_1, \varphi_2\}$ and $\{\psi_1, \psi_2\}$. In fact, this is the root of the *superposition principle* (see the next subsection for the details).

Finally, we remark that there is a crucial difference between classical physical behaviour ($\lambda_1 = \lambda_2 = 0$) and quantum decoherence ($\lambda_1 = \lambda_2 = 0$). In the first case coefficients $\lambda_j = 0$, because statistical deviations are negligibly small. In the second case coefficients $\lambda_j = 0$, because statistical deviations compensate each other ($j = 1, 2$):

$$\frac{\Delta_{1j}}{N} \approx -\frac{\Delta_{2j}}{N} \qquad N \to \infty.$$

3.3. Non-classical/quantum trigonometric probability behaviour

Here the matrix $P = (p_{ij})$ of transition probabilities need not be double stochastic. We can find the probability distribution $q_j = p_j^a = p_S(A = a_j)$, $j = 1, 2$, by using the following transformation of probabilities:

$$q_j = p_1 p_{1j} + p_2 p_{2j} + 2\sqrt{p_1 p_{1j} p_2 p_{2j}} \cos\theta_j \tag{18}$$

where $\cos\theta_1 = -K\cos\theta_2$, $K = \sqrt{\frac{p_{12}p_{22}}{p_{11}p_{21}}}$. In general such a probabilistic transformation ('interference' between preparation procedures $\mathcal{E}_1$ and $\mathcal{E}_2$) could not be described by standard quantum formalism.

Example 3.1. Let $p_1 = p_2 = \frac{1}{2}$ (symmetric distribution of C in S; for example, the two-slit experiment with symmetric location of slits with respect to the source of particles) and let $p_{11} = p_{12} = \frac{1}{2}$ (symmetric distribution of A in S_1) and $p_{21} = \frac{1}{3}$, $p_{22} = \frac{2}{3}$ (asymmetric distribution of A in S_2). Thus the matrix P is not double stochastic.

The law of conservation of the A is violated in the process of the transition $S \to (S_1, S_2)$. The measure of this violation is given by the coefficient K. Here $K = \sqrt{2}$. Phases θ_1 and θ_2 must be chosen in such a way that $\cos\theta_1 = -\sqrt{2}\cos\theta_2$. For example, we can consider preparations such that $\theta_1 = \frac{3\pi}{4}$ and $\theta_2 = \frac{\pi}{3}$. In this case we have

$$p_1^a = \frac{5}{12} + \frac{\cos\frac{3\pi}{4}}{\sqrt{6}} \qquad p_2^a = \frac{7}{12} + \frac{\cos\frac{\pi}{3}}{\sqrt{3}}.$$

This probabilistic transformation could not be obtained in standard 'quantum linear calculus'. We shall see that it could be obtained by non-unitary generalization of 'quantum linear calculus'.

4. Hyperbolic probabilistic behaviour

In this section we consider examples of H- and HT-behaviour. We remark that H-behaviour can be exhibited by preparations having double-stochastic transition matrices.

Example 4.1. Let $p_1 = \alpha$ and $p_2 = 1 - \alpha$ $(0 < \alpha < 1)$ and let $p_{ij} = 1/2$, $i, j = 1, 2$. Here $K = 1$ (the transition matrix is double stochastic) and, hence, $\cosh\theta_2 = \cosh\theta_1$. We have

$$q_1 = \tfrac{1}{2} + \sqrt{\alpha(1-\alpha)}\cosh\theta$$
$$q_2 = \tfrac{1}{2} - \sqrt{\alpha(1-\alpha)}\cosh\theta.$$

In the opposite to the T-case the phase θ cannot take arbitrary values. There is a relation between θ and α that provides that q_1, q_2 have the meaning of probabilities. We set

$$e(\alpha) = \frac{1}{2\sqrt{\alpha(1-\alpha)}}.$$

We remark that $e(\alpha) \geqslant 1$ for all $0 < \alpha < 1$. The hyperbolic phase θ can be chosen as $\theta \in [0, \theta_{\max}]$, where $\theta_{\max} = \operatorname{arccosh} e(\alpha)$. For example, let $\alpha = \frac{1}{4}(1 - \alpha = 3/4)$. Thus $e(x) = \frac{2}{\sqrt{3}}$. Here we could observe hyperbolic interference for angles $0 \leqslant \theta \leqslant \operatorname{arccosh}\frac{2}{\sqrt{3}}$. We remark that if $p_1 = p_2 = \frac{1}{2}$, then $e(\alpha) = 1$ and the hyperbolic interference coincides with the ordinary interference $\cos 0 = \cosh 0 = 1$. In general the symmetric distribution $p_1 = p_2 = 1/2$ can produce non-trivial hyperbolic interference. We have for a general double-stochastic matrix p

$$q_j = \tfrac{1}{2}(p_{1j} + p_{2j}) + \sqrt{p_{1j}p_{2j}}\lambda_j = \tfrac{1}{2} + \sqrt{p_{1j}p_{2j}}\lambda_j = \tfrac{1}{2} + \sqrt{\alpha(1-\alpha)}\lambda_j$$

where we set $\alpha = p_{11} = p_{22}$ and $1 - \alpha = p_{12} = p_{21}$. If $\theta \in [0, \theta_{\max}]$, $\theta_{\max} = \operatorname{arccosh} e(\alpha)$, then $\lambda_j = \pm\cosh\theta$, $\theta \neq 0$.

We remark that the total symmetry (in S as well as S_1, S_2), namely $p_1 = p_2 = p_{ij} = 1/2$, produces the trivial H-interference (that coincides with the T-interference), so hyperbolic interference might be observed only for preparation procedures with asymmetric probability distributions for contexts.

Remark (negative probabilities). If we do not pay attention to the range of the H-phase parameter θ we could obtain *negative probabilities* and probabilities >1. It must be noted that such 'probabilities' appear with intriguing regularity in various extensions of quantum formalism ([28–30], see also [19] for the details). It may be that 'quantum negative probabilities' have the same origin as 'negative H-probabilities,' namely the use of non-physical values of some parameters (see [19] for the details).

Of course, our considerations induce the following natural question: 'is it possible to construct a linear space representation for the H-probabilistic transformations?' We shall study this question in section 6.

Finally, we consider an example of mixed HT-behaviour.

Example 4.2. Let $p_1 = p_2 = \frac{1}{2}$ and let $p_{11} = \frac{4}{5}$, $p_{12} = \frac{1}{5}$, $p_{21} = \frac{4}{5}$, $p_{22} = \frac{1}{5}$. We have $K = \frac{1}{4}$; so $\lambda_2 = -4\lambda_1$.

We have $q_1 = \frac{4}{5}(1 + \lambda_1)$, $q_2 = \frac{1}{5}(1 - 4\lambda_1)$. If $-1 \leqslant \lambda_1 \leqslant \frac{1}{4}$, then q_1 and q_2 have the meaning of probabilities. For example, let $\lambda_1 = \frac{-1}{2}$ and $\lambda_2 = 2$. Then $q_1 = \frac{2}{5}$, $q_2 = \frac{3}{5}$. Thus

$$q_1 = \tfrac{4}{5} + \tfrac{4}{5}\cos\tfrac{2}{3}\pi \qquad q_2 = \tfrac{1}{5} + \tfrac{1}{5}\cosh\left(\ln(2+\sqrt{3})\right).$$

We remark that mixed HT-behaviour cannot be produced on the basis of a double-stochastic matrix $P = (p_{ij})$.

Finally, we note that the H-phase has a symmetry, $\theta \to -\theta$, that is an analogue of the symmetry $\theta \to \theta + 2\pi$ for the T-phase. If $\lambda = \cosh\theta$, then θ can be chosen as

$$\theta = \ln\left(\lambda + \sqrt{\lambda^2 - 1}\right) \qquad \text{or} \qquad \theta = \ln\left(\lambda - \sqrt{\lambda^2 - 1}\right).$$

5. Complex linear space representation of the general trigonometric probabilistic rule

We shall study the possibility of representing a general probabilistic transformation (18) as a linear transformation in a complex linear space. As in general the transition probability matrix $P = (p_{ij})$ is not double stochastic, we could not expect that it would be possible to work with orthonormal bases in a complex Hilbert space. It seems that the inner product structure is not useful in the general case.

Let E be a two-dimensional linear space over the field of complex numbers $\boldsymbol{C}$. The choice of $\boldsymbol{C}$ as the basic number field has trivial explanation. Formula (15) gives the possibility of representing the T-probabilistic transformation in form (16), which is reduced to the transition from one basis to another. It is impossible to linearize a quantum probabilistic transformation by using real numbers, but it is possible to do this by using complex numbers. These arguments were already evident in our analysis of quantum theory. We now observe that they can be used in a more general situation.

Vectors of E are said to be quantum states. At the moment there is no Hilbert structure on E. There is nothing similar to the standard normalization condition for quantum states. We represent the ensemble S (the preparation procedure $\mathcal{E}$) by a vector φ in E and the ensembles S_1 and S_2 (the preparation procedures $\mathcal{E}_1$ and $\mathcal{E}_2$) by vectors φ_1 and φ_2.

It is supposed that the preparation procedures $\mathcal{E}_1$ and $\mathcal{E}_2$ determine some dichotomic physical variable, $C = c_1, c_2$. In the linear space calculus this assumption has the following counterpart: vectors $\{\varphi_1, \varphi_2\}$ are linearly independent in E.

Splitting S into S_1 and S_2 (due to the preparation procedures $\mathcal{E}_1$ and $\mathcal{E}_2$) is represented as expanding the vector φ with respect to a basis $\{\varphi_1, \varphi_2\}$ in E. We can always expand the vector ϕ with respect to the basis:

$$\varphi = \alpha_1\varphi_1 + \alpha_2\varphi_2$$

where α_1 and $\alpha_2 \in \boldsymbol{C}$. As in the ordinary quantum formalism the probabilities $p_i^c = P_S\ (C = c_i)$ are represented as $p_i^c = |\alpha_i|^2$ (generalization of Born's postulate). There is therefore a constraint for vectors φ and φ_1, φ_2:

$$|\alpha_1|^2 + |\alpha_2|^2 = 1. \tag{19}$$

In such a case the quantum state φ is said to be *C-decomposable*.

We now consider the measurement of A for ensembles S_i (prepared by $\mathcal{E}_i$). We consider such a measurement that a second measurement of A, performed immediately after the first one, will yield the same value of the observable. In quantum theory such measurements are often called *'measurements of the first kind'*. Thus such A-measurement can be interpreted as a preparation procedure.

To be more precise, we consider two preparation procedures $\mathcal{E}_1^a$ and $\mathcal{E}_2^a$ corresponding to selections of physical systems on the basis of values $A = a_1$ and $A = a_2$. The C-preparation procedures $\mathcal{E}_1$ and $\mathcal{E}_2$ we now denote by the symbols $\mathcal{E}_1^c$ and $\mathcal{E}_2^c$, respectively. $\mathcal{E}_j^a$ selects physical systems such that $A = a_j$, $j = 1, 2$. We remark that in general these selections may change the probability distribution of C. By applying $\mathcal{E}_j^a$ to the ensemble $S_i^c \equiv S_i$ (which was produced by the application of $\mathcal{E}_i^c$ to an ensemble S produced by $\mathcal{E}$) we obtain an ensemble S_{ij}^{ca}, $i, j = 1, 2$. In the same way we split the ensemble S (with the aid of $\mathcal{E}_1^a$ and $\mathcal{E}_2^a$) into ensembles S_j^a, $j = 1, 2$. Ensembles S_j^a, $j = 1, 2$, are represented by vectors ψ_j in the E. We assume that they also form a basis in E (this is a consequence of the fact that preparation procedures $\mathcal{E}_1^a$ and $\mathcal{E}_2^a$ determine the dichotomic physical variable A). Thus splitting $S \to (S_1^a, S_2^a)$ can be represented by the expansion

$$\varphi = \beta_1\psi_1 + \beta_2\psi_2$$

where $\beta_j \in \boldsymbol{C}$. Here probabilities $p_j^a = P_S(A = a_j) = |\beta_j|^2$, so

$$|\beta_1|^2 + |\beta_2|^2 = 1. \tag{20}$$

Thus φ is A-decomposable.

In the general case we have to represent ensembles S_{ij}^{ca}, $i, j = 1, 2$, by four different vectors ψ_{ij}. In general we cannot assume that these vectors belong to the same two-dimensional space E. The study of this general situation is too complicated. We restrict ourselves to the special case (which is the most interesting for applications). Let $\psi_{11} = \psi_1$ and $\psi_{21} = \psi_1, \psi_{21} = \psi_2$ and $\psi_{22} = \psi_2$. It was assumed that ψ_1 and ψ_2 are independent vectors.

We should like to predict the probabilities p_j^a on the basis of the transition from the basis $\{\varphi_1, \varphi_2\}$ to the basis $\{\psi_1, \psi_2\}$. Let $U = (\beta_{ij})$ be the transition matrix (the only restriction to U is its invertibility). Here each vector φ_i is A-decomposable[3]. Thus

$$|\beta_{i1}|^2 + |\beta_{i2}|^2 = 1 \qquad i = 1, 2. \tag{21}$$

[3] In general there is no composition (or it would be better to say decomposition) transitivity. For example, it may be that the state φ is C-decomposable and each state φ_i is A-decomposable, but φ is not A-decomposable. We suppose decomposability of all states under the consideration for physical reasons: the possibility to perform A and C measurements for elements of S. The violation of composition transitivity corresponds to the following situation: we can perform C-measurement on S and A-measurements on S_i^c, but we could not perform A-measurement on S.

We have

$$\beta_1 = \alpha_1\beta_{11} + \alpha_2\beta_{21} \qquad \beta_2 = \alpha_1\beta_{12} + \alpha_2\beta_{22}. \tag{22}$$

Coefficients α_j, β_{ij} are not independent. They satisfy constraint (20). Simple computations give us

$$\alpha_1\bar{\alpha}_2(\beta_{11}\bar{\beta}_{21} + \beta_{12}\bar{\beta}_{22}) + \bar{\alpha}_1\alpha_2(\bar{\beta}_{11}\beta_{21} + \bar{\beta}_{12}\beta_{22}) = 0. \tag{23}$$

One of the solutions of this equation is given by

$$\beta_{11}\bar{\beta}_{21} + \beta_{12}\bar{\beta}_{22} = 0. \tag{24}$$

This is the condition of unitarity of the transition matrix $U = (\beta_{ij})$. This solution gives the ordinary quantum formalism. In this formalism it is useful to introduce the inner product

$$\langle z, w\rangle = z_1\bar{w}_1 + z_2\bar{w}_2$$

and rewrite the above equation as the condition of orthogonality of vectors φ_1 and φ_2: $\langle\varphi_1, \varphi_2\rangle = 0$. However, equation (23) has other solutions which are not related to standard quantum formalism. These solutions give the complex linear space representation for the trigonometric probabilistic rule in the non-classical/quantum case. We set

$$\alpha_i = \sqrt{p_i}\mathrm{e}^{\mathrm{i}\xi_i} \qquad \beta_{ij} = \sqrt{p_{ij}}\mathrm{e}^{\mathrm{i}\gamma_{ij}}$$

where $p_1 + p_2 = 1$, $p_{11} + p_{12} = 1$, $p_{21} + p_{22} = 1$ and ξ_1, γ_{ij} are arbitrary phases. Thus the transition from one basis to another has the form

$$\varphi_1 = \sqrt{p_{11}}\mathrm{e}^{\gamma_{11}}\psi_1 + \sqrt{p_{12}}\mathrm{e}^{\gamma_{12}}\psi_2 \qquad \varphi_2 = \sqrt{p_{21}}\mathrm{e}^{\gamma_{21}}\psi_1 + \sqrt{p_{22}}\mathrm{e}^{\gamma_{22}}\psi_2. \tag{25}$$

In these notations equation (23) has the form

$$\cos(\eta + \gamma_1)\sqrt{p_{11}p_{21}} + \cos(\eta + \gamma_2)\sqrt{p_{12}p_{22}} = 0 \tag{26}$$

where $\eta = \xi_1 - \xi_2$, $\gamma_1 = \gamma_{11} - \gamma_{21}$, $\gamma_2 = \gamma_{12} - \gamma_{22}$.

We set $\theta_1 = \eta + \gamma_1$ and $\theta_2 = \eta + \gamma_2$. Equation (26) coincides with equation (6) in the T-case. Thus all possible probabilistic T-transformations can be represented in the complex linear space. A rather surprising fact is that equation (26) has a new (non-quantum solution) even for a double-stochastic matrix of transition probabilities.

Let P be a double-stochastic matrix. Equation (26) has the form

$$\cos(\eta + \gamma_1) + \cos(\eta + \gamma_2) = 0.$$

Thus

$$\cos\frac{(2\eta + \gamma_1 + \gamma_2)}{2} = 0 \qquad \text{or} \qquad \cos\frac{(\gamma_1 - \gamma_2)}{2} = 0.$$

There is a crucial difference between these equations. The first equation 'remembers' the state φ, splitting φ into $\{\varphi_1, \varphi_2\}$ (or S into S_1 and S_2). This memory is given by the phase shift η. The second equation does not contain any memory term. In fact, this is the standard quantum mechanical equation: $\gamma_1 - \gamma_2 = \pi \pmod{2\pi}$.

Thus we obtain a new (non-quantum) solution even for a double-stochastic matrix $P = (p_{ij})$:

$$2\eta + \gamma_1 + \gamma_2 = \pi \pmod{2\pi}.$$

In this case transformation (25) also reproduces the quantum probabilistic rule (13), (14): $q_j = p_1p_{1j} + p_2p_{2j} \pm 2\sqrt{p_1p_{1j}p_2p_{2j}}\cos\theta$. However, (25) is not unitary:

$$\beta_{11}\bar{\beta}_{21} + \beta_{12}\bar{\beta}_{22} = 1 - \mathrm{e}^{-2\mathrm{i}\eta} \neq 0 \qquad \eta \neq 0.$$

6. Linear space representation of the hyperbolic probabilistic rule

We want to find a kind of linear space calculus for the H-probabilistic transformation. It seems that it would be impossible to do this in a C-linear space. We propose to use a hyperbolic algebra G (see [31]). This is a two-dimensional real algebra with basis $e_0 = 1$ and $e_1 = \mathrm{j}$, where $\mathrm{j}^2 = 1$.

Elements of G have the form $z = x + \mathrm{j}y,\ x, y \in R$. We have $z_1 + z_2 = (x_1 + x_2) + \mathrm{j}(y_1 + y_2)$ and $z_1 z_2 = (x_1 x_2 + y_1 y_2) + \mathrm{j}(x_1 y_2 + x_2 y_1)$. This algebra is commutative. We introduce an involution in G by setting $\bar{z} = x - \mathrm{j}y$. We set

$$|z|^2 = z\bar{z} = x^2 - y^2.$$

We remark that $|z| = \sqrt{x^2 - y^2}$ is not well defined for an arbitrary $z \in G$. We set $G_+ = \{z \in G : |z|^2 \geqslant 0\}$. We remark that G_+ is the multiplicative semigroup $z_1, z_2 \in G^+ \to z = z_1 z_2 \in G_+$. This is a consequence of the equality

$$|z_1 z_2|^2 = |z_1|^2 |z_2|^2.$$

Thus, for $z_1, z_2 \in G_+$, we have $|z_1 z_2| = |z_1||z_2|$. We introduce

$$\mathrm{e}^{\mathrm{j}\theta} = \cosh\theta + \mathrm{j}\sinh\theta \qquad \theta \in R.$$

We remark that

$$\mathrm{e}^{\mathrm{j}\theta_1}\mathrm{e}^{\mathrm{j}\theta_2} = \mathrm{e}^{\mathrm{j}(\theta_1+\theta_2)} \qquad \overline{\mathrm{e}^{\mathrm{j}\theta}} = \mathrm{e}^{-\mathrm{j}\theta} \qquad |\mathrm{e}^{\mathrm{j}\theta}|^2 = \cosh^2\theta - \sinh^2\theta = 1.$$

Hence, $z = \pm\mathrm{e}^{\mathrm{j}\theta}$ always belongs to G_+. We also have

$$\cosh\theta = \frac{\mathrm{e}^{\mathrm{j}\theta} + \mathrm{e}^{-\mathrm{j}\theta}}{2} \qquad \sinh\theta = \frac{\mathrm{e}^{\mathrm{j}\theta} - \mathrm{e}^{-\mathrm{j}\theta}}{2\mathrm{j}}.$$

We set $G_+^* = \{z \in G_+ : |z|^2 > 0\}$. Let $z \in G_+^*$. We have

$$z = |z|\left(\frac{x}{|z|} + \mathrm{j}\frac{y}{|z|}\right) = \operatorname{sign} x\,|z|\left(\frac{x \operatorname{sign} x}{|z|} + \mathrm{j}\,\frac{y \operatorname{sign} x}{|z|}\right).$$

As $\frac{x^2}{|z|^2} - \frac{y^2}{|z|^2} = 1$, we can represent $x \operatorname{sign} x = \cosh\theta$ and $y \operatorname{sign} x = \sinh\theta$, where the phase θ is unequally defined. We can represent each $z \in G_+^*$ as

$$z = \operatorname{sign} x\ |z|\ \mathrm{e}^{\mathrm{j}\theta}.$$

By using this representation we can easily prove that G_+^* is the multiplicative group. Here $\frac{1}{z} = \frac{\operatorname{sign} x}{|z|}\mathrm{e}^{-\mathrm{j}\theta}$. The unit circle in G is defined as $S_1 = \{z \in G : |z|^2 = 1\} = \{z = \pm\mathrm{e}^{\mathrm{j}\theta}, \theta \in (-\infty, +\infty)\}$. It is a multiplicative subgroup of G_+^*.

Hyperbolic Hilbert space is G-linear space (module) E with a G-scalar product: a map $(\cdot,\cdot) : E \times E \to G$ that is:

(1) linear with respect to the first argument,

$$(az + bw, u) = a(z, u) + b(w, u) \qquad a, b \in G \quad z, w, u \in E;$$

(2) symmetric, $(z, u) = \overline{(u, z)}$;

(3) non-degenerated, $(z, u) = 0$ for all $u \in E$ iff $z = 0$.

We note that (1) and (2) imply that

$$(u, az + bw) = \bar{a}(u, z) + \bar{b}(u, w).$$

Remark. If we consider E as just a R-linear space, then $(\cdot,\cdot)$ is a (rather special) bilinear form which is not positively defined. In particular, in the two-dimensional case we have the signature $(+, -, +, -)$.

We shall represent the H-probabilistic transformation in the two-dimensional $\boldsymbol{G}$-linear space (module) $\boldsymbol{E}$. From the beginning we do not consider any $\boldsymbol{G}$-Hilbert structure on $\boldsymbol{E}$. Such a structure will appear automatically in the representation of one particular class of H-probabilistic transformations, H-quantum formalism. In the same way as in the previous section we introduce quantum states $\varphi, \{\varphi_1, \varphi_2\}, \{\psi_1, \psi_2\}$ corresponding to preparation procedures (statistical ensembles). By definition a quantum state is a vector belonging to a $\boldsymbol{G}$-linear space (no normalization!).

It is supposed that $\{\varphi_1, \varphi_2\}$ and $\{\psi_1, \psi_2\}$ are bases in the $\boldsymbol{G}$-linear space E.

It is supposed that the state φ is C and A-decomposable and the states φ_i are A-decomposable. Thus

$$\varphi = \alpha_1\varphi_1 + \alpha_2\varphi_2 \qquad |\alpha_1|^2 + |\alpha_2|^2 = 1 \qquad |\alpha_j|^2 \geqslant 0$$

and

$$\varphi_1 = \beta_{11}\psi_1 + \beta_{12}\psi_2 \qquad \varphi_2 = \beta_{21}\psi_1 + \beta_{22}\psi_2$$

where vectors of coefficients $\beta^{(1)} = (\beta_{11}, \beta_{12})$ and $\beta^{(2)} = (\beta_{21}, \beta_{22})$ are such that

$$|\beta_{11}|^2 + |\beta_{12}|^2 = 1 |\beta_{21}|^2 + |\beta_{22}|^2 = 1 \qquad \text{and} \qquad |\beta_{ij}|^2 \geqslant 0.$$

Thus

$$\varphi = \beta_1\psi_1 + \beta_2\psi_2$$

where the coefficients β_1, β_2 are given by (22). There is no formal difference between linear space transformations over $\boldsymbol{C}$ and $\boldsymbol{G}$. However, the assumption that the state φ is A-decomposable implies that the $\boldsymbol{G}$-linear space calculations have a physical meaning iff the vector $\beta = (\beta_1, \beta_2)$ is such that

$$|\beta_1|^2 = |\alpha_1\beta_{11} + \alpha_2\beta_{21}|^2 \geqslant 0 \qquad |\beta_2|^2 = |\alpha_1\beta_{12} + \alpha_2\beta_{22}|^2 \geqslant 0 \tag{27}$$

and

$$|\beta_1|^2 + |\beta_2|^2 = 1. \tag{28}$$

The latter equation coincides with equation (23) (with the only difference that all numbers belong to $\boldsymbol{G}$ instead of $\boldsymbol{C}$).

As we have already discussed in the T-case, in general there is no composition (in fact, decomposition) transitivity. In general the C-decomposability of φ and A-decomposability of φ_i need not imply that φ is also A-decomposable. Our assumptions on composition transitivity are based on the physical context of our considerations.

As in the $\boldsymbol{T}$-case, (23) has the solution given by equation (24) (the only difference is that now all coefficients belong to the hyperbolic algebra). This is the condition of orthogonality of vectors φ_1 and φ_2 with respect to the $\boldsymbol{G}$-linear product: $\langle z, w\rangle = z_1\bar{w}_1 + z_2\bar{w}_2$. So the matrix $U = (\beta_{ij})$ is a $\boldsymbol{G}$-unitary matrix, namely

$$\langle \beta^{(i)}, \beta^{(j)}\rangle = \delta_{ij}. \tag{29}$$

We now study the general case. Here the U need not be a unitary matrix. We consider only vectors with coefficients belonging to $\boldsymbol{G}_+^*$. We set $\alpha_i = \pm\sqrt{p_i}e^{j\xi_i}$, $\beta_{ij} = \pm\sqrt{p_{ij}}e^{j\gamma_{ij}}$, $i, j = 1, 2$. Condition (28) is equivalent to the condition

$$\sqrt{p_{12}p_{22}}\cosh\theta_2 + \sigma\sqrt{p_{11}p_{21}}\cosh\theta_2 = 0$$

where $\sigma = \sqcap_{ij}\operatorname{sign}\beta_{ij}$. This equation has a solution, namely phases θ_1 and θ_2, iff

$$\sigma = -1. \tag{30}$$

Thus the transition matrix $U = (\beta_{ij})$ must always satisfy (30).

Let us turn back to the case in which U is a G-unitary matrix. We shall call such a model *hyperbolic quantum formalism*. The orthogonality relation implies

$$0 = (\beta^{(1)}, \beta^{(2)}) = \operatorname{sign}\beta_{11}\operatorname{sign}\beta_{21}\sqrt{p_{11}p_{21}}e^{j(\gamma_{11}-\gamma_{21})} + \operatorname{sign}\beta_{12}\operatorname{sign}\beta_{22}\sqrt{p_{12}p_{22}}e^{j(\gamma_{12}-\gamma_{22})}$$

or

$$1 + \sigma K e^{j(\gamma_1-\gamma_2)} = 0$$

where $K = \sqrt{p_{12}p_{22}}/\sqrt{p_{11}p_{21}}$ and $\gamma_1 = \gamma_{12} - \gamma_{22}$, $\gamma_2 = \gamma_{11} - \gamma_{21}$. Thus $\sinh(\gamma_1 - \gamma_2) = 0$ and

$$\gamma_1 = \gamma_2 \tag{31}$$

(we recall that in the standard quantum formalism we have $\gamma_1 = \gamma_2 + \pi \pmod{2\pi}$). We also have

$$1 + \sigma K \cosh(\gamma_1 - \gamma_2) = 0.$$

Thus $\sigma = -1$ and $K = 1$. The sign condition (30) is therefore always satisfied for a unitary matrix $U = (\beta_{ij})$. The equality $K = 1$ is equivalent to double stochasticity of the transition matrix of probabilities $P = (p_{ij} = |\beta_{ij}|^2)$. Therefore the matrix $U = (\beta_{ij})$ is a G-unitary matrix iff the corresponding matrix of probabilities $P = (p_{ij})$ is a double-stochastic matrix, $\sigma = -1$, and hyperbolic phases satisfy (31).

The H-quantum formalism (special calculus in a G-linear space) represents probabilistic transformations

$$q_1 = p_1p_{11} + p_2p_{21} \pm 2\sqrt{p_1p_2p_{11}p_{21}}\cosh\theta$$
$$q_2 = p_1p_{12} + p_2p_{22} \mp 2\sqrt{p_1p_2p_{12}p_{22}}\cosh\theta$$

where $\theta = \gamma_{11} - \gamma_{21} = \gamma_{12} - \gamma_{22}$.

The situation is similar to the ordinary quantum formalism. However, there is an important difference between these formalisms. In the T-quantum formalism the condition of C-unitarity of $U = (\beta_{ij})$ was also sufficient to obtain physically meaningful transformation of probabilities: all possible phases θ give meaningful probabilistic transformation for the fixed C-unitary matrix $U = (\beta_{ij})$. This is not so in the H-quantum formalism. The G-unitarity of $U = \beta_{ij}$ is not sufficient to obtain physically meaningful probabilities for all H-phases θ. Besides condition (28), we also have condition (27), which provides non-negativity of probabilities $q_j = p_j^a = |\beta_j|^2$.

We set $t = p_{11} = p_{22}$ (so $p_{12} = p_{21} = 1 - t$), $0 < t < 1$ (we recall that $P = (p_{ij})$ is a double-stochastic matrix). We also set $p_1 = s$, so $p_2 = 1 - s$, $0 < s < 1$. Let us consider the case in which $\operatorname{sign}\beta_{11}\operatorname{sign}\beta_{21} = -1$. Hence $\operatorname{sign}\beta_{12}\operatorname{sign}\beta_{22} = 1$. Here

$$q_1 = st + (1-s)(1-t) - 2\sqrt{s(1-s)t(1-t)}\cosh\theta$$
$$q_2 = s(1-t) + (1-s)t + 2\sqrt{s(1-s)t(1-t)}\cosh\theta.$$

Thus

$$\cosh\theta \leqslant \frac{st + (1-s)(1-t)}{2\sqrt{s(1-s)t(1-t)}} = e(s,t).$$

Thus physical H-behaviour is possible only for probabilities s, t such that $e(s,t) \geqslant 1$ (in the case of the equality H- and T-behaviours coincide).

We note that there is no analogue of the superposition principle in the H-quantum formalism. G-unitary transformations preserve normalization condition (28), but they do not preserve positivity conditions (27).

We now turn back to the general case in which the P need not be double stochastic. We consider again equation (28), which is equivalent to (23) (with coefficients belonging to hyperbolic algebra). We have already studied the special class of solutions of equation (23) given by (24). These solutions are given by G-unitary matrices. We now consider the general equation:

$$\sigma K \cosh(\eta + \gamma_2) + \cosh(\eta + \gamma_1) = 0. \tag{32}$$

As $\sigma = -1$, we finally obtain the equation

$$K \cosh\theta_2 = -\cosh\theta_1$$

(compare to (7)).The presence of the H-phase $\eta = \xi_1 - \xi_2$ plays the role of memory in the preparation procedure $\mathcal{E}$ (which produced an ensemble S represented by the state φ).

We remark that equation (32) has following two solutions for $K = 1$ (double-stochastic matrix):

$$\cosh(\eta + \gamma_2) = \cosh(\eta + \gamma_1) \rightarrow \eta + \gamma_2 = \eta + \gamma_1 \qquad \text{or} \qquad \eta + \gamma_2 = -\eta - \gamma_1.$$

In the first case we have the H-quantum solution, $\gamma_1 = \gamma_2$, and in the second case we have a new solution, $2\eta + \gamma_2 + \gamma_1 = 0$, that corresponds to the non-unitary transition matrix U.

7. Conclusions

Our frequency analysis of probabilities related to transitions from one experimental arrangement (context, complex of physical conditions) to another showed the following.

(1) The 'quantum rule' for interference of probabilistic alternatives can be obtained in a purely contextualist approach; in particular, without applying wave arguments.
(2) Both 'quantum' and 'classical' probabilities can be interpreted as frequency probabilities. Specific 'quantum behaviour' of probabilities in experiments with quantum particles is related to the specific relation between elementary particles and experimental arrangement. There in the 'quantum world' transition from one context to another produces statistical perturbations that change the classical Bayes formula (by the additive interference term). In the 'classical world' such perturbations are negligibly small statistically.
(3) Transformations of probabilities corresponding to context transitions can be classified according to the magnitudes of statistical perturbations: trigonometric, hyperbolic or hyper-trigonometric. In particular, contextual modifications of the classical Bayes formula are not reduced to a 'quantum rule' for interference of probabilistic alternatives. There exists non-classical/quantum trigonometric interference of probabilistic alternatives as well as hyperbolic interference.
(4) The main distinguishing feature of 'quantum probabilistic transformations' is not the appearance of the $\cos\theta$-interference term, but the double stochasticity of the matrix of transition probabilities.
(5) Starting with the trigonometric transformation of probabilities, we obtain (with the aid of the cos-theorem) a complex amplitude representation of contextual probabilities. This gives the possibility of constructing a complex linear space representation of contextual probabilistic calculus. In general, we could not represent a trigonometric probabilistic transformation in a complex Hilbert space. This is possible only for double-stochastic matrices of transition probabilities (that corresponds to unitary transformations of a Hilbert space).
(6) One of the special features of general C-linear representation of contextual probabilities is the violation of the superposition principle. It seems that this fundamental principle is a consequence of double stochasticity.

(7) Hyperbolic probabilistic transformations can be represented as linear transformations in modules over the system (commutative algebra) $\boldsymbol{G}$ of hyperbolic numbers.
(8) One of the special features of general $\boldsymbol{G}$-linear representation of contextual probabilities is the violation of the superposition principle. In this case even double stochasticity of the matrix of transition probability ($\boldsymbol{G}$-unitarity of the corresponding $\boldsymbol{G}$-linear transformation) does not imply the superposition principle.
(9) Trigonometric transformations correspond to context transitions inducing relatively small statistical perturbations, hyperbolic relatively large.
(10) In principle, non-classical/quantum probabilistic behaviour (trigonometric as well as hyperbolic) could be simulated numerically (see [32]).

Finally, we make a remark on the contextualist viewpoint of superselection rules. Superselection rules are closely related to the superposition principle. With a superselection rule unitarity (double stochasticity of the matrix of transition probabilities) or linear combinations do not imply coherent superposition. Superselection rules also are important since they are relevant to macroscopic quantum systems [33, 34].

I think that superselection rules give restrictions on physical realization of some preparation procedures, namely filtration (selection) procedures that give a possibility of transforming an ensemble of physical systems prepared under one fixed complex of conditions $\mathcal{S}$ into an ensemble of physical systems prepared under some special complex of physical conditions $\mathcal{S}'$. There exist complexes $\mathcal{S}$ and $\mathcal{S}'$ such that it is impossible to create the corresponding transformation procedure. However, I think (and it may be that I am wrong) that superselection rules could not be analysed in a general probabilistic framework. Each rule is closely connected to some fixed class of physical systems under consideration. If we represent in the same linear (in particular, Hilbert) space contextual probabilities for distinct classes of physical systems, then we shall obtain distinct classes of contexts that could not be transformed into each other.

Acknowledgments

I should like to thank L Ballentine, S Albeverio, E Beltrametti, T Hida, D Greenberger, S Gudder, I Volovich, W De Muynck, J Summhammer, P Lahti, J-A Larsson, H Atmanspacher, B Coecke, S Aerts, A Peres, A Holevo, E Loubenets, L Polley, A Zeilinger, C Fuchs, R Gill, L Hardy, B Hiley, S Goldshtein, A Plotnitsky, A Shimony, R Jozsa, J Bub, C Caves, K Gustafsson and H Bernstein for fruitful (and rather critical) discussions.

References

[1] Dirac P A M 1995 *The Principles of Quantum Mechanics* (Oxford: Clarendon)
[2] von Neumann J 1955 *Mathematical Foundations of Quantum Mechanics* (Princeton, NJ: Princeton University Press)
[3] Heisenberg W 1930 *Physical Principles of Quantum Theory* (Chicago, IL: Chicago University Press)
[4] Feynman R and Hibbs A 1965 *Quantum Mechanics and Path Integrals* (New York: McGraw-Hill)
[5] Jauch J M 1968 *Foundations of Quantum Mechanics* (Reading, MA: Addison-Wesley)
[6] Busch P, Grabowski M and Lahti P 1995 *Operational Quantum Physics* (Berlin: Springer)
[7] d'Espagnat B 1995 *Veiled Reality. An Analysis of Present-Day Quantum Mechanical Concepts* (Reading, MA: Addison-Wesley)
[8] Peres A 1994 *Quantum Theory: Concepts and Methods* (Dordrecht: Kluwer)
[9] Beltrametti E and Cassinelli G 1981 *The Logic of Quantum Mechanics* (Reading, MA: Addison-Wesley)
[10] Ballentine L E 1989 *Quantum Mechanics* (Englewood Cliffs, NJ: Prentice-Hall)
[11] Bohr N 1935 *Phys. Rev.* **48** 696–702

[12] Accardi L 1984 The probabilistic roots of the quantum mechanical paradoxes *The Wave–Particle Dualism. A Tribute to Louis de Broglie on His 90th Birthday* ed S Diner, D Fargue, G Lochak and F Selleri (Dordrecht: Reidel) pp 297–330
Accardi L 1997 *Urne e Camaleoni: Dialogo sulla Realta, le Leggi del Caso e la Teoria Quantistica* (Rome: Saggiatore)
[13] Ballentine L E 1970 *Rev. Mod. Phys.* **42** 358–81
[14] Gudder S P 1984 *J. Math. Phys.* **25** 2397–401
[15] De Muynck W, De Baere W and Marten H 1994 *Found. Phys.* **24** 1589–663
De Muynck W and Stekelenborg J T 1988 *Ann. Phys., Lpz.* **45** 222–34
[16] Pitowsky I 1982 *Phys. Rev. Lett.* **48** 1299–302
Pitowsky I 1983 *Phys. Rev.* D **27** 2316–26
[17] Davies E B 1976 *Quantum Theory of Open Systems* (London: Academic)
[18] Ludwig G 1983 *Foundations of Quantum Mechanics* (Berlin: Springer)
[19] Khrennikov A Yu 1999 *Interpretations of Probability* (Utrecht: VSP)
[20] Khrennikov A Yu 2000 *J. Math. Phys.* **41** 5934–44
[21] Khrennikov A Yu 2001 *Phys. Lett.* A **278** 307–14
[22] Khrennikov A Yu 2000 Ensemble fluctuations and the origin of quantum probabilistic rule *Rep. MSI* 90, Växjö University
[23] Khrennikov A Yu 2000 Hyperbolic quantum mechanics *Preprint* quant-ph/0101002
[24] Holevo A S 1982 *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam: North-Holland)
[25] von Mises R 1964 *The Mathematical Theory of Probability and Statistics* (London: Academic)
Ville J 1939 *Etude Critique de la Notion de Collective* (Paris: Gauthier-Villars)
Tornier E 1936 *Wahrscheinlichkeitsrechnunug und Allgemeine Integrationstheorie* (Leipzig: University Press)
[26] Kolmogoroff A N 1933 *Grundbegriffe der Wahrscheinlichkeitsrech* (Berlin: Springer)
Kolmogoroff A N 1956 *Foundations of the Probability Theory* (New York: Chelsea) (reprinted)
[27] Ballentine L 2001 Interpretations of probability and quantum theory *Quantum Probability and Related Fields* at press
[28] Wigner E 1971 Quantum-mechanical distribution functions revisted *Perspectives in Quantum Theory* ed W Yourgrau and A van der Merwe (Cambridge, MA: MIT Press)
[29] Dirac P A M 1942 *Proc. R. Soc.* A **180** 1–39
[30] Feynman R P 1987 Negative probability *Quantum Implications, Essays in Honour of David Bohm* ed B J Hiley and F D Peat (London: Routledge and Kegan Paul) pp 235–46
[31] Khrennikov A Yu 1999 *Superanalysis* (Dordrecht: Kluwer)
[32] Khrennikov A 2001 Quantum statistics via perturbation effects of preparation procedures *Preprint* quant-ph/0103065
[33] Wan K K and Harrison F 1993 *Phys. Lett.* A **174** 1
[34] Wan K K, Bradshaw J, Trueman C and Harrison F 1998 *Found. Phys.* **28** 1739–83