Quantum probability law from 'environment-assisted invariance' in terms of pure-state twin unitaries

Quantum probability law from 'environment-assisted invariance' in terms of pure-state twin unitaries

F Herbut

Serbian Academy of Sciences and Arts, Knez Mihajlova 35, 11000 Belgrade, Serbia

E-mail: fedorh@infosky.net and fedorh@mi.sanu.ac.yu

Abstract
Zurek's derivation of the Born rule from envariance (environment-assisted invariance) is retraced in an attempt to present a more detailed derivation, which is extended to encompass the trace-rule form of the quantum probability law in its full generality. The investigation begins by a review presentation of Schmidt decomposition, and a detailed and complete theory of twin unitaries, which are the other face of envariance, and which stand in close connection with twin observables. The trace rule $\mathrm{tr}(E\rho)$, with E being an event (projector) and ρ a quantum state (density operator), is derived in five stages. The first three achieve the same as Zurek's derivation of Born's rule, but in a different way, using the full power of twin unitaries. Stage 4 extends the Zurek-like derivation to states (ray projectors) that are not eigenstates of the density operator. Finally, stage 5 utilizes the assumption of continuity of the probability law in the density operators to extend the trace rule to isolated (not correlated) systems.

PACS numbers: 03.65.Ta, 03.65.Ca

1. Introduction

Zurek has introduced [1] *envariance* (environment-assisted invariance) in the following way. He imagined a system $\mathcal{S}$ entangled with a dynamically decoupled environment $\mathcal{E}$, altogether described by a bipartite state vector $|\psi\rangle_{\mathcal{SE}}$. Further, he imagined two opposite-subsystem unitary operators $u_{\mathcal{S}}$ and $u_{\mathcal{E}}$ that 'counter-transformed' each other when elevated to the composite system $U_{\mathcal{S}} \equiv (u_{\mathcal{S}} \otimes 1_{\mathcal{E}})$, $U_{\mathcal{E}} \equiv (1_{\mathcal{S}} \otimes u_{\mathcal{E}})$, and applied to the bipartite state vector, e.g.,

$$U_{\mathcal{E}}U_{\mathcal{S}}|\psi\rangle_{\mathcal{SE}} = |\psi\rangle_{\mathcal{SE}}. \tag{1}$$

Zurek remarked that 'When the transformed property of the system can be so 'untransformed' by acting only on the environment, it is not the property of $\mathcal{S}$.' Zurek,

further, paraphrases Bohr's famous dictum: 'If the reader does not find envariance strange, he has not understood it.'

The *first aim* of this study is to acquire a full understanding of envariance, or rather of its other face, of twin unitaries. The wish to understand envariance as much as possible is not motivated only by its strangeness, but also by the fact that Zurek makes use of it to derive one of the basic laws of quantum mechanics: Born's rule.

Zurek's derivation gave rise to critical comments and inspired analogous attempts [2–5]. In a preliminary version of this study [6], quotations from each of these works and also from Zurek's most important paper on his derivation [7] were presented and commented upon from the point of view of this author's approach.

The *second aim* of this study is to utilize the full power of the complete theory of twin unitaries to rederive (in a different way) Zurek's result, and to go on to complete it to the full quantum probability law.

I favour what I call the quantum-logical approach to the quantum probability law, which is widely accepted since Gleason's paper [8]: if E is an event or property (mathematically a projector in the state space) of the system, and ρ is its state (mathematically a density operator), then the probability of the former in the latter is tr$(E\rho)$. (This form of the quantum probability law is called the 'trace rule').

An equivalent, and perhaps more practical, form of the probability law is the following: if $|\phi\rangle$ is an arbitrary state vector of the system, then $\langle\phi|\rho|\phi\rangle$ is the probability that in a suitable measurement on the system in the state ρ the *event* $|\phi\rangle\langle\phi|$ will occur. To follow Zurek's derivation as closely as possible, this form of the quantum probability law will be rederived, and, for brevity, $|\phi\rangle$ instead of the corresponding ray projector will be used throughout. A proof of the equivalence of the just stated probability formula of this paper and the trace rule is given below in relation (28).

All derivations of Born's rule from envariance in the literature are *restricted* to eigenstates $(\rho|\phi\rangle = r|\phi\rangle$, r a positive number) of the density operator ρ describing the state of the system. Four of the cited commentators of Zurek's argument (I have failed to get in touch with Fine) have pointed out to me that the restriction can be understood as natural in the context of (previous) system–environment interaction, which has led to decoherence (cf [9], section III.E.4), or if one takes the relative-state (or many-worlds) view, where the 'observer' is so entangled with the system in the measurement that the restriction covers the general case (cf [10]).

In the next section a precise and detailed presentation is given of Schmidt decomposition and of its special forms, canonical Schmidt decomposition and strong Schmidt decomposition. In this last, most specific form, the antiunitary correlation operator U_a (introduced in previous work [11]), the sole correlation entity inherent in a given bipartite state vector is made use of. It is the entity that turns Schmidt decomposition into strong Schmidt decomposition, making it complete and precise. This entity is lacking in almost all examples of the use of Schmidt decomposition in the literature. (For an alternative approach to the correlation operator via the antilinear operator representation of bipartite state vectors, see section 2 in [12].)

Twin unitaries, i.e., opposite-subsystem unitary operators that act equally on a given bipartite state vector, which are hence equivalent to envariance, are analysed in detail, and the group of all pairs of them is derived.

There is another derivation of the full set of envariance in the recent literature [13]. It is algebraic, i.e., in terms of matrices and suitable numbers, whereas the approach of this study is geometrical, i.e., it is in terms of state space decompositions and suitable maps.

In the next section also connection is established between twin unitaries and twin Hermitians. The latter are the so-called twin observables, studied in detail in pure bipartite

states in previous papers [11, 14]. Also a possibility of extending the notion of twin unitaries to mixed bipartite states is shortly discussed. Extension to twin Hermitians in mixed states was accomplished in previous work [15].

The next section actually consists of six subsections. Immediately before the first subsection, the structure of the six subsections is explained.

In section 3 the quantum probability law is derived. During the derivation some points come up that require more detailed discussion. In order to keep the derivation unbroken, these discussions are presented in the subsections of section 4.

2. Mathematical interlude: Strong Schmidt decomposition and twin unitaries

If one has two opposite factor-space unitaries u_1 and u_2 that, on defining $U_1 \equiv (u_1 \otimes 1_2)$ and $U_2 \equiv (1_1 \otimes u_2)$, *act equally* on a given composite state vector

$$U_1|\Psi\rangle_{12} = U_2|\Psi\rangle_{12}, \tag{2a}$$

then one speaks of *twin unitaries* (unitary twin operators) with respect to the bipartite state vector. They give another, equivalent, view of envariance (see the introduction), since, rewriting (2a) as

$$|\Psi\rangle_{12} = U_1^{-1}U_2|\Psi\rangle_{12}, \tag{2b}$$

one can see that U_1^{-1} 'untransforms' the action of U_2 (cf (1)) or symmetrically.

Incidentally, it is easy to see that $U_1|\Psi\rangle_{12}\langle\Psi|_{12}U_1^{-1} = U_2|\Psi\rangle_{12}\langle\Psi|_{12}U_2^{-1}$ is equivalent to

$$U_1|\Psi\rangle_{12} = e^{i\lambda}U_2|\Psi\rangle_{12}, \tag{2c}$$

where $\lambda \in \mathbf{R}_1$. This does not diminish the usefulness of definition (2a), because, if (2c) is valid for a pair (U_1, U_2), then one only has to replace these operators by $(U_1, e^{i\lambda}U_2)$, and the latter satisfy (2a).

Henceforth, we will write U_s both for u_s, $s = 1, 2$, and for $(1_1 \otimes u_2)$ or $(u_1 \otimes 1_2)$ (cf (1)).

Proposition. *If two opposite-subsystem unitaries U_1 and U_2 are twin unitaries, then so are U_1^{-1} and U_2^{-1}, as well as $U_1^{\dagger}$ and $U_2^{\dagger}$, where the dagger denotes adjoining. Besides, also the bra relation*

$$\langle\Psi|_{12}U_1 = \langle\Psi|_{12}U_2 \tag{2d}$$

is valid.

Proof. Utilizing relation (2b) in the last step, one has

$$U_1^{-1}|\Psi\rangle_{12} = \left(U_2^{-1}U_2\right)U_1^{-1}|\Psi\rangle_{12} = U_2^{-1}\left(U_2U_1^{-1}\right)|\Psi\rangle_{12} = U_2^{-1}|\Psi\rangle_{12}.$$

The second claim follows from $U_s^{-1} = U_s^{\dagger}$, $s = 1, 2$. The final claim (2d) is the dual (or bra relation) of the ket relation $U_1^{\dagger}|\Psi\rangle_{12} = U_2^{\dagger}|\Psi\rangle_{12}$. $\qquad\square$

It is the aim of this section to derive a complete theory of twin unitaries. To be precise, by this is meant answering the following three questions:

 (i) which unitaries U_s, $s = 1$ or 2, *have a twin?*
 (ii) if U_s, $s = 1$ or 2, does have a twin, how to utilize the given bipartite state vector $|\Psi\rangle_{12}$ *to evaluate a twin?*
(iii) is the twin unique?

We begin by giving a partial answer to the first question.

Theorem 1 (A). *If a subsystem unitary U_s, $s = 1$ or 2, has a twin, then it necessarily commutes with the corresponding reduced density operator:*

$$[U_s, \rho_s] = 0, \qquad s = 1 \ or \ 2, \tag{3a, b}$$

where

$$\rho_s \equiv \mathrm{tr}_t(|\Psi\rangle_{12}\langle\Psi|_{12}), \qquad s, t = 1 \ or \ 2, \qquad s \neq t, \tag{3c, d}$$

and 'tr$_t$' is the partial trace in the corresponding factor space. An (obvious) equivalent form of the necessary condition is invariance of the reduced density operator under the unitary:

$$U_s \rho_s U_s^{-1} = \rho_s, \qquad s = 1 \ or \ 2. \tag{3e, f}$$

Proof. Making use of (2d) in the next-to-last step, one has

$$U_1\rho_1 = U_1\mathrm{tr}_2(|\Psi\rangle_{12}\langle\Psi|_{12}) = \mathrm{tr}_2(U_1|\Psi\rangle_{12}\langle\Psi|_{12}) = \mathrm{tr}_2((U_2|\Psi\rangle_{12})\langle\Psi|_{12})$$
$$= \mathrm{tr}_2((|\Psi\rangle_{12}\langle\Psi|_{12})U_2) = \mathrm{tr}_2(|\Psi\rangle_{12}\langle\Psi|_{12}U_1) = \rho_1 U_1.$$

Symmetrically one derives the symmetrical relation for U_2. $\qquad\qquad\square$

In what follows it will be shown that $(3a, b)$ or $(3e, f)$ give also a sufficient condition, i.e., they give a complete answer to the first question. This will be done by answering also the second question. We will take an arbitrary U_s, $s = 1$ or 2, satisfying $(3e, f)$ with $(3c, d)$, and we will evaluate its twin. Finally, the third question will also be answered essentially in the affirmative.

For the evaluation of the twin a correlation entity inherent in $|\Psi\rangle_{12}$ will be utilized. To acquire this entity, a concise deviation is made in subsection 2.1. to give a review (without proofs) of Schmidt decompositions. In particular, canonical Schmidt decomposition is presented as a springboard for introducing the correlation entity in subsection 2.2.

In subsection 2.3. strong Schmidt decomposition, which involves the correlation entity, as well as the so-called correlated subsystem picture, which gives an alternative view of $|\Psi\rangle_{12}$, is presented. In subsection 2.4. the answer to question (i) is completed and questions (ii) and (iii) are also answered. In subsection 2.5. connection is established between twin unitaries and twin Hermitians. Finally, in subsection 2.6. twin unitaries in mixed bipartite states are shortly discussed.

2.1. Review of Schmidt decompositions

We assume that a completely arbitrary *bipartite state vector* $|\Psi\rangle_{12}$ is given. It is an arbitrary normalized vector in $\mathcal{H}_1 \otimes \mathcal{H}_2$, where the factor spaces are finite- or infinite-dimensional complex separable Hilbert spaces. The statements are, as a rule, asymmetric in the roles of the two factor spaces. But, as is well known, for every general asymmetric statement, also its symmetric one, obtained by exchanging the roles of 1 and 2, is valid.

We call an orthonormal complete basis simply 'basis'.

The natural framework for Schmidt decomposition is *general expansion in a factor-space basis*.

Let $\{|m\rangle_1 : \forall m\}$ be an arbitrary basis in $\mathcal{H}_1$. Then there exists a unique expansion

$$|\Psi\rangle_{12} = \sum_m |m\rangle_1 |m\rangle_2', \tag{4a}$$

where the generalized expansion coefficients $\{|m\rangle_2' : \forall m\}$ are elements of the opposite factor space $\mathcal{H}_2$, and they depend only on $|\Psi\rangle_{12}$ and the corresponding basis vectors $|m\rangle_1$, and not on the entire basis.

The generalized expansion coefficients are evaluated making use of the partial scalar product:

$$\forall m: \quad |m\rangle_2' = \langle m|_1 |\Psi\rangle_{12}. \tag{4b}$$

The partial scalar product is, in turn, evaluated expanding $|\Psi\rangle_{12}$ in arbitrary bases $\{|k\rangle_1 : \forall k\} \subset \mathcal{H}_1, \{|l\rangle_2 : \forall l\} \subset \mathcal{H}_2$, and by utilizing the ordinary scalar products in the composite and in the first factor spaces:

$$|\Psi\rangle_{12} = \sum_k \sum_l (\langle k|_1 \langle l|_2 |\Psi\rangle_{12}) |k\rangle_1 |l\rangle_2. \tag{4c}$$

Then (4b) reads

$$\forall m: \quad |m\rangle_2' = \sum_l \left\{ \sum_k [(\langle m|_1 |k\rangle_1)(\langle k|_1 \langle l|_2 |\Psi\rangle_{12})] \right\} |l\rangle_2, \tag{4d}$$

and the lhs is independent of the choice of the bases in the factor spaces.

The proof is straightforward.

Now we define Schmidt decomposition. It is well known and much used in the literature. It is only the first springboard for the main theory presented in subsection 2.4., for that of twin unitaries.

If, besides the basis vectors $|m\rangle_1$, the 'expansion coefficients' $|m\rangle_2'$ are also *orthogonal* in expansion (4a), then one speaks of a *Schmidt decomposition*. It is usually written in terms of normalized second-factor-space vectors $\{|m\rangle_2 : \forall m\}$:

$$|\Psi\rangle_{12} = \sum_m \alpha_m |m\rangle_1 |m\rangle_2, \tag{5}$$

where α_m are complex numbers and $\forall m : |m\rangle_1$ and $|m\rangle_2$ are referred to as *partners* in a pair of Schmidt states.

The term 'Schmidt decomposition' can be replaced by 'Schmidt expansion' or 'Schmidt form'. To be consistent and avoid confusion, we will stick to the first term throughout.

Expansion (4a) is *Schmidt decomposition if and only if* the first-factor-space basis $\{|m\rangle_1 : \forall m\}$ is an *eigenbasis* of ρ_1:

$$\forall m: \quad \rho_1 |m\rangle_1 = r_m |m\rangle_1, \qquad 0 \leqslant r_m. \tag{6}$$

We will make much use of this basic property of Schmidt decomposition.

As is obvious in relation (5), the two subsystems play symmetric roles in Schmidt decomposition. Hence, also the symmetric relation to the necessary and sufficient condition (6) is valid.

Next we define a special form of Schmidt decomposition that is sometimes more useful. It is called canonical Schmidt decomposition.

The non-trivial phase factors of the non-zero coefficients α_m in (5) can be absorbed either in the basis vectors in $\mathcal{H}_1$ or in those in $\mathcal{H}_2$ (or partly the former and partly the latter). If in a Schmidt decomposition (5) all non-zero α_m are non-negative real numbers, then we write, instead of (5), the following decomposition

$$|\Psi\rangle_{12} = \sum_i r_i^{1/2} |i\rangle_1 |i\rangle_2, \tag{7a}$$

and we confine the sum to non-zero terms (one is reminded of all this by the replacement of the index m by i in this notation). Relation ($7a$) is called *canonical Schmidt decomposition*. (The term 'canonical' reminds of the form of ($7a$), i.e., of $\forall i : r_i^{1/2} > 0$.)

Needless to say, every $|\Psi\rangle_{12}$ can be written as a canonical Schmidt decomposition.

Every canonical Schmidt decomposition ($7a$) is accompanied by the *spectral forms of the reduced density operators* (cf ($3c, d$)):

$$\rho_s = \sum_i r_i |i\rangle_s \langle i|_s, \qquad s = 1, 2. \qquad (7b, c)$$

(The same eigenvalues r_i appear both in ($7a$) and in the two spectral forms ($7b, c$).)

One should note that the topologically closed ranges $\bar{\mathcal{R}}(\rho_s)$, $s = 1, 2$ (subspaces) of the reduced density operators ρ_s, $s = 1, 2$ are *equally dimensional*. The range-projectors are

$$Q_s = \sum_i |i\rangle_s \langle i|_s, \qquad s = 1, 2. \qquad (7d, e)$$

The two reduced density operators have *equal eigenvalues* $\{r_i : \forall i\}$ (including equal possible degeneracies).

One has a canonical Schmidt decomposition ($7a$) *if and only if* the decomposition is bi-orthonormal and all expansion coefficients are positive.

The proof of these claims is straightforward. In the approach of this paper, ($7a$) is only the second springboard for the final, most special, form of the Schmidt decomposition (9) given below.

2.2. The antiunitary correlation operator

It is high time we introduced *the sole entanglement entity* inherent in any bipartite state vector, which is lacking from both forms of Schmidt decomposition discussed so far. It is an antiunitary map that takes the closed range $\bar{\mathcal{R}}(\rho_1)$ onto the symmetrical entity $\bar{\mathcal{R}}(\rho_2)$. (If the ranges are finite-dimensional, they are *ipso facto* closed, i.e., they are subspaces.) The map is called *the correlation operator*, and denoted by the symbol U_a [11, 14].

If a canonical Schmidt decomposition ($7a$) is given, then the two orthonormal bases of equal power $\{|i\rangle_1 : \forall i\}$ and $\{|i\rangle_2 : \forall i\}$ define an antiunitary, i.e., antilinear and unitary operator U_a, the correlation operator—the sole correlation entity inherent in the given state vector $|\Psi\rangle_{12}$:

$$\forall i : \quad |i\rangle_2 \equiv (U_a |i\rangle_1)_2. \qquad (8a)$$

The correlation operator U_a, mapping $\bar{\mathcal{R}}(\rho_1)$ onto $\bar{\mathcal{R}}(\rho_2)$, is well defined by ($8a$) and by the additional requirements of antilinearity (complex conjugation of numbers, coefficients in a linear combination) and by continuity (if the bases are infinite). (Both these requirements follow from that of antiunitarity.) Preservation of every scalar product up to complex conjugation, which, by definition, makes U_a antiunitary, is easily seen to follow from ($8a$) and the requirements of antilinearity and continuity because U_a takes an orthonormal basis into another orthonormal one.

Though canonical Schmidt decomposition ($7a$) is non-unique (even if ρ_s, $s = 1, 2$, are non-degenerate in all their positive eigenvalues, there is the non-uniqueness of the phase factors of $|i\rangle_1$), the correlation operator U_a is *uniquely* implied by a given bipartite state vector $|\Psi\rangle_{12}$.

This claim is proved in appendix A.

The uniqueness of U_a when $|\Psi\rangle_{12}$ is given is a slight compensation for the trouble one has treating an antilinear operator. (Though the difficulty is more psychological than practical,

because all that distinguishes an antiunitary operator from a unitary one is its antilinearity—it complex-conjugates the numbers in any linear combination—and its property that it preserves the absolute value, but complex-conjugates every scalar product.) The full compensation comes from the usefulness of U_a.

Once the orthonormal bases $\{|i\rangle_1 : \forall i\}$ and $\{|i\rangle_2 : \forall i\}$ of a canonical Schmidt decomposition (7a) are given, one can write

$$U_a = \sum_i |i\rangle_2 K \langle i|_1, \tag{8b}$$

where K is complex conjugation (denoted by an asterisk when numbers are concerned). This means, by definition,

$$\forall |\phi\rangle_1 : \quad U_a|\phi\rangle_1 = \sum_i (\langle i|_1|\phi\rangle_1)^*|i\rangle_2. \tag{8c}$$

2.3. Strong Schmidt decomposition and the correlated subsystem picture

We finally introduce the most specific form of Schmidt decomposition.

If one rewrites (7a) in terms of the correlation operator by substituting (8a) in (7a), then it takes the form

$$|\Psi\rangle_{12} = \sum_i r_i^{1/2}|i\rangle_1(U_a|i\rangle_1)_2. \tag{9}$$

This is called *strong Schmidt decomposition*.

If a strong Schmidt decomposition (9) is written down, then it can be viewed in two opposite ways:

(i) as a given bipartite state vector $|\Psi\rangle_{12}$ defining its two inherent entities, the reduced density operator ρ_1 in spectral form (cf (7b)) and the correlation operator U_a (cf (8a)), both relevant for the entanglement in the state vector;

(ii) as a given pair (ρ_1, U_a) (U_a mapping antiunitarily $\bar{\mathcal{R}}(\rho_1)$ onto some equally dimensional subspace of $\mathcal{H}_2$) defining a bipartite state vector $|\Psi\rangle_{12}$.

The second view of strong Schmidt decomposition allows a systematic generation and classification of all state vectors in $\mathcal{H}_1 \otimes \mathcal{H}_2$ (cf [16]).

One has

$$\rho_2 = U_a\rho_1 U_a^{-1} Q_2, \qquad \rho_1 = U_a^{-1}\rho_2 U_a Q_1 \tag{10a, b}$$

(cf (8a) and (7b, c)). Thus, the reduced density operators are, essentially, 'images' of each other via the correlation operator. (The term 'essentially' points to the fact that the dimensions of the null spaces are independent of each other.) Property (10a, b) is called *twin operators*.

When one takes into account the *eigensubspaces* $\mathcal{R}(Q_s^j)$ of ρ_s corresponding to (the common) distinct positive eigenvalues r_j of ρ_s, where Q_s^j projects onto the r_j-eigensubspace, $s = 1, 2$, then one obtains a *geometrical view* of the *entanglement* in a given state $|\Psi\rangle_{12}$ in terms of the so-called *correlated subsystem picture* [11]:

$$\bar{\mathcal{R}}(\rho_s) = \sum_j^{\oplus} \mathcal{R}(Q_s^j), \qquad s = 1, 2, \tag{10c, d}$$

where ' $\oplus$ ' denotes an orthogonal sum of subspaces,

$$\forall j : \quad \mathcal{R}(Q_2^j) = U_a\mathcal{R}(Q_1^j), \qquad \mathcal{R}(Q_1^j) = U_a^{-1}\mathcal{R}(Q_2^j), \tag{10e, f}$$

and, of course,

$$\bar{\mathcal{R}}(\rho_2) = U_a\bar{\mathcal{R}}(\rho_1), \qquad \bar{\mathcal{R}}(\rho_1) = U_a^{-1}\bar{\mathcal{R}}(\rho_2). \tag{10g, h}$$

In words, the correlation operator makes not only the ranges of the reduced density operators 'images' of each other, but also all positive-eigenvalue eigensubspaces. In other words, the correlation operator makes the eigendecompositions of the ranges 'images' of each other.

One should note that all positive-eigenvalue eigensubspaces $\mathcal{R}\left(Q_s^j\right)$ are necessarily finite dimensional because $\sum_i r_i = 1$ (a consequence of the normalization of $|\Psi\rangle_{12}$), and hence no positive-eigenvalue can have infinite degeneracy.

The correlated subsystem picture of a given bipartite state vector is very useful in investigating remote influences (as a way to understand physically the entanglement in the composite state) (see [12, 14]).

We will need the correlated subsystem picture of $|\Psi\rangle_{12}$ for the basic result of this section given below: the final theorem on twin unitaries.

2.4. Completing the answers

We are now able to complete our study of twin unitaries.

Theorem 1 (B)**.** *Every subsystem unitary satisfying (3a) or (3b) with (3c) or (3d) has a twin.*

Theorem 2. *If a unitary U_1 leaves ρ_1 invariant (cf (3e) with (3c)), then*

$$U_2 \equiv U_a U_1^{-1} U_a^{-1} Q_2 + U_2' Q_2^\perp \tag{11}$$

is its twin. (Here Q_2 projects onto the range of ρ_2, and U_2' in $U_2' Q_2^\perp$ is an arbitrary second-subsystem unitary commuting with Q_2.)

Proof of theorems 1(B) and 2. Let U_1 satisfy (3e). Then so does U_1^{-1}, and hence $\left\{U_1^{-1}|i\rangle_1 : \forall i\right\}$ is an eigenbasis of ρ_1 in its range just like $\{|i\rangle_1 : \forall i\}$. Therefore, we can write down a strong Schmidt decomposition as follows (cf (9)):

$$|\Psi\rangle_{12} = \sum_i r_i^{1/2}\left(U_1^{-1}|i\rangle_1\right)\left(U_a\left(U_1^{-1}|i\rangle_1\right)\right)_2.$$

Applying U_1 to this, we obtain

$$U_1|\Psi\rangle_{12} = \sum_i r_i^{1/2}|i\rangle_1\left(U_a U_1^{-1}|i\rangle_1\right)_2. \tag{12a}$$

On the other hand, application of $U_2 \equiv U_a U_1^{-1} U_a^{-1} Q_2 + U_2' Q_2^\perp$ to

$$|\Psi\rangle_{12} = \sum_i r_i^{1/2}|i\rangle_1(U_a|i\rangle_1)_2$$

results in

$$U_2|\Psi\rangle_{12} = \sum_i r_i^{1/2}|i\rangle_1\left(\left(U_a U_1^{-1} U_a^{-1}\right)U_a|i\rangle_1\right)_2 = \sum_i r_i^{1/2}|i\rangle_1\left(U_a U_1^{-1}|i\rangle_1\right)_2. \tag{12b}$$

The rhs of (12a) and that of (12b) are equal. $\square$

Theorem 3. *If U_1 is a subsystem unitary that leaves ρ_1 invariant (cf (3e)), then it has the essentially unique twin unitary $U_2 \equiv U_a U_1^{-1} U_a^{-1} Q_2 + U_2' Q_2^\perp$, where U_2' satisfies $[U_2', Q_2] = 0$, and otherwise it is an arbitrary unitary.*

Proof. Let (U_1, U_2) be twin unitaries. Applying them to the two strong Schmidt decompositions, $|\Psi\rangle_{12} = \sum_i r_i^{1/2}\left(U_1^{-1}|i\rangle_1\right)\left(U_a\left(U_1^{-1}|i\rangle_1\right)\right)_2 = \sum_i r_i^{1/2}|i\rangle_1\left(U_a|i\rangle_1\right)_2$, definition (2a) gives

$$\sum_i r_i^{1/2}|i\rangle_1\left(U_a U_1^{-1}|i\rangle_1\right)_2 = \sum_i r_i^{1/2}|i\rangle_1(U_2 U_a|i\rangle_1)_2.$$

Since the second tensor factor in each term of canonical Schmidt decomposition is unique, one has

$$\forall i: \quad \left(\left(U_a U_1^{-1}\right)|i\rangle_1\right)_2 = \left((U_2 U_a)|i\rangle_1\right)_2,$$

which can be rewritten as

$$U_2 = U_a U_1^{-1} U_a^{-1} Q_2 + U_2' Q_2^{\perp}. \qquad \qquad \square$$

It is straightforward to show (along the lines of the proof just presented) that the twin unitaries are also responsible for the non-uniqueness of strong (or of canonical) Schmidt decomposition. To put this more precisely, besides (9) (besides (7a)) all other strong Schmidt decompositions (canonical Schmidt decompositions) are obtained by replacing $\{|i\rangle_1 : \forall i\}$ in (9) by $\{U_1|i\rangle_1 : \forall i\}$, where $[U_1, \rho_1] = 0$ (by replacing $\{|i\rangle_1|i\rangle_2 : \forall i\}$ in (7a) by $\left\{(U_1|i\rangle_1)\left(U_2^{-1}|i\rangle_2\right) : \forall i\right\}$, where $[U_s, \rho_s] = 0, s = 1, 2,$ and (11) is satisfied).

The set of all pairs of twin unitaries (U_1, U_2) is a *group*, if one defines the composition law by $(U_1', U_2') \times (U_1, U_2) \equiv (U_1' U_1, U_2 U_2')$ (note the inverted order in $\mathcal{H}_2$). Taking the inverse turns out to be $(U_1, U_2)^{-1} = \left(U_1^{-1}, U_2^{-1}\right)$. The claim is proved in appendix B.

Having in mind the correlated subsystem picture (10a)–(10h) of $|\Psi\rangle_{12}$, it is immediately seen that the three theorems on twin unitaries can be cast in the following equivalent unified form.

The final theorem on twin unitaries. The group of all twin unitaries (U_1, U_2) consists of all pairs of opposite factor-space unitaries

(i) that reduce in every positive-eigenvalue eigensubspace $\mathcal{R}\left(Q_s^j\right), s = 1, 2$ (cf (10c, d)), and

(ii) the reducees are connected by relation (11) and its symmetrical one mutatis mutandis, or, to be more precise, in (11) and its symmetrical relation Q_s is replaced by $Q_s^j, s = 1, 2,$ and the obtained system of relations enumerated by j are valid simultaneously for all j values.

In the more transparent language of formulae, the group consists of all pairs of unitaries (U_1, U_2)

(i) that can be written in the form

$$U_s = \sum_j U_s^j Q_s^j + U_s' Q_s^{\perp}, \qquad s = 1, 2, \qquad (13a, b)$$

where $\forall j : U_s^j$ is a unitary in $\mathcal{R}\left(Q_s^j\right)$, and $[U_s', Q_s] = 0$, otherwise U_s' is an arbitrary unitary, $s = 1, 2$; and

(ii)

$$\forall j: \quad U_2^j Q_2^j = U_a \left(U_1^j\right)^{-1} U_a^{-1} Q_2^j, \qquad (13c)$$

$$U_1^j Q_1^j = U_a^{-1} \left(U_2^j\right)^{-1} U_a Q_1^j. \qquad (13d)$$

Note that within each positive-eigenvalue eigensubspace $\mathcal{R}\left(Q_s^j\right)$ of $\rho_s, s = 1, 2,$ *all* unitaries U_s^j are encompassed (but not independently, cf (13c, d)). This will be important in the application in the next section.

The next two (short) subsections round out the study of twin unitaries. The reader who is primarily interested in the argument leading to the quantum probability law is advised to skip them.

2.5. Connection with twin Hermitians

There is a notion closely connected with twin unitaries in a pure bipartite state: it is that of twin Hermitians (in that state). If a pair (H_1, H_2) of opposite factor-space Hermitian operators commute with the corresponding reduced density operators, and

$$H_2 = U_a H_1 U_a^{-1} Q_2 + H_2' Q_2^\perp, \qquad H_1 = U_a^{-1} H_2 U_a Q_1 + H_1' Q_1^\perp \qquad (14a, b)$$

is valid ($[H_s', Q_s] = 0$, otherwise H_s' is an arbitrary Hermitian, $s = 1, 2$), then one speaks of twin Hermitian operators. (Relations $(14a, b)$, in analogy with (11) and its symmetrical one, state that the reducees in the ranges of the reduced density operators are 'U_a-images' of each other, and the reducees in the null spaces are completely arbitrary.)

One should note that twin unitaries and twin Hermitians are, actually, defined in the same way. To see this, one has to replace U_s^{-1} by $U_s^\dagger$ in (11) and in its symmetrical relation, and H_s by $H_s^\dagger$, $s = 1, 2$, in $(14a, b)$.

Twin Hermitians have important physical meaning [12, 14]. But here we are only concerned with their connection with twin unitaries.

If $U_s, s = 1$ or $s = 2$, are symmetry operators of the corresponding reduced density operators (cf $(3e, f)$), then there exist Hermitian operators H_s that also commute with the latter and

$$U_s = e^{iH_s} Q_s + U_s' Q_s^\perp, \qquad s = 1 \quad \text{or} \quad s = 2 \qquad (15a, b)$$

is valid. And *vice versa*, if $H_s, s = 1$ or $s = 2$ are Hermitians that commute with the corresponding reduced density operators, then there exist analogous unitaries given by $(15a, b)$. (The unitary and Hermitian reducees in the ranges determine each other in $(15a, b)$, and the reducees in the null spaces are arbitrary.)

The latter claim is obvious. But to see that also the former is valid, one should take into account that commutation with the corresponding reduced density operator implies reduction in each (finite dimensional) eigensubspace $\mathcal{R}(Q_s^j)$ of ρ_s (cf $(10c, d)$). Then one can take the spectral form of each reducee of U_s, and $(15a, b)$ becomes obvious (and the corresponding reducees of H_s are unique if their eigenvalues are required to be, e.g., in the intervals $[0, 2\pi)$.)

The connection $(15a, b)$, which goes in both directions, can be *extended to twin operators*.

If (U_1, U_2) are twin unitaries, then $(15a, b)$ (with 'or' replaced by 'and') determine corresponding twin Hermitians, and *vice versa*, if (H_1, H_2) are twin Hermitians, then the same relations determine corresponding twin unitaries.

2.6. Mixed states

If ρ_{12} is a *mixed bipartite density operator*, then we no longer have the correlation operator U_a and the correlated subsystem picture $(10a)$–$(10h)$. Nevertheless, in some cases, twin Hermitians, defined by

$$H_1 \rho_{12} = H_2 \rho_{12}, \qquad (16a)$$

have been found [15]. (Their physical meaning was analogous to that in the pure-state case.) It was shown that $(16a)$ implied

$$[H_s, \rho_s] = 0, \qquad s = 1, 2, \qquad (16b, c)$$

where ρ_s are again the reduced density operators. (Unlike in the case where ρ_{12} is a pure state, in the mixed-state case the commutations $(16b, c)$ are not sufficient for the existence of a twin observable.)

Relations $(16b, c)$, in turn, again imply reduction of H_s in every positive-eigenvalue eigensubspace $\mathcal{R}(Q_s^j)$ of $\rho_s, s = 1, 2$, but now the dimensions of the corresponding,

i.e., equal-j, eigensubspaces are, unlike in ($10c, d$), completely independent of each other (but finite dimensional). In each of them, relations ($15a, b$) (with 'and' instead of 'or') hold true, and define *twin unitaries* satisfying ($2a$) with ρ_{12} instead of $|\Psi\rangle_{12}\langle\Psi|_{12}$.

Thus, in some cases, the concept of envariance can be extended to mixed states.

3. Quantum probability law from twin unitaries

The forthcoming argument is given in five stages; the first three stages are an attempt to essentially reproduce Zurek's argument [1, 7, 17, 18] in a different way by utilizing the group of all pairs of twin unitaries (presented in subsection 2.4). This derivation is intended to be more explicit.

Zurek's 'environment', which, after the standard interaction with the system under consideration, establishes special, measurement-like correlations with it, is explicitly treated as any entangled bipartite pure state $|\Psi\rangle_{12}$, where subsystem 1 is the system under consideration, and 2 is some opposite subsystem with an *infinite dimensional* state space $\mathcal{H}_2$. We shall try to see to what extent and how the quantum probability law follows from the quantum correlations, i.e., the entanglement in $|\Psi\rangle_{12}$.

The forth stage is new. It is meant to extend the argument to states $|\phi\rangle_1$ which are not eigenvectors of the reduced density operator $\rho_1 \equiv \mathrm{tr}_2(|\Psi\rangle_{12}\langle\Psi|_{12})$. The fifth stage is also new. It extends the argument to isolated (not correlated) systems.

Let $|\Psi\rangle_{12}$ be an arbitrary entangled bipartite state vector. We assume that subsystems 1 and 2 are not interacting. (They may have interacted in the past and thus have created the entanglement. But it also may have been created in some other way, e.g., by an external field as the spatial-spin entanglement in a Stern–Gerlach apparatus.)

We want to obtain the probability law in subsystem 1. By this we assume that there exist probabilities, and we do not investigate why this is so; we only want to obtain their form.

On p 4, left column, in [7] (the archive copy), Zurek lists three 'facts', which he considers basic to his approach.

'*Fact 1*. Unitary transformations must act on the system to alter its state. (That is, when the evolution operator does not operate on the Hilbert space $\mathcal{H}_S$ of the system, i.e., when it has a form $\cdots \otimes \mathbf{1}_S \otimes \cdots$ the state of S remains the same.)

Fact 2. The state of the system S is all that is needed (and all that is available) to predict measurement outcomes, including their probabilities.

Fact 3. The state of a larger composite system that includes S as a subsystem is all that is needed (and all that is available) to determine the state of the system S.'

Zurek adds '. . . the above *facts* are interpretation-neutral and the states (e.g., 'the state of S') they refer to need not be pure.'

I find Zurek's 'facts' fully convincing, and I am going to build them into the present Zurek-like approach. Mentioning them explicitly will, I hope, also facilitate a comparison between Zurek's original derivation and my first three stages below.

The *first stipulation* is as follows.

(a) Though the given pure state $|\Psi\rangle_{12}$ determines all properties in the composite system, therefore also all those of subsystem 1, the latter must be *determined actually by the subsystem alone*. This is, by (vague) definition, what is meant by *local* or subsystem properties.

(b) There exist local or subsystem probabilities of all elementary events $|\phi\rangle_1\langle\phi|_1$, $|\phi\rangle_1 \in \mathcal{H}_1$. (As it has been stated in the introduction, we will write the event shortly as the state vector that determines it.)

Since $|\Psi\rangle_{12} \in (\mathcal{H}_1 \otimes \mathcal{H}_2)$, subsystem 1 is somehow connected with the state space $\mathcal{H}_1$, but it is not immediately clear precisely how. Namely, since we start out *without the probability rule*, the reduced density operator $\rho_1 \equiv \mathrm{tr}_2(|\Psi\rangle_{12}\langle\Psi|_{12})$, though mathematically at our disposal, is yet devoid of physical meaning. We need a precise definition of what is local or what is the subsystem state. We will achieve this gradually, and thus ρ_1 *will be gradually endowed with the standard physical meaning*.

The *second stipulation* is that subsystem or *local properties must not be changeable by remote action*, i.e., by applying a second-subsystem unitary U_2 to $|\Psi\rangle_{12}$ or any unitary U_{23} applied to the opposite subsystem with an ancilla (subsystem 3).

If this were not so, then there would be no sense in calling the properties at issue 'local' and not 'global' in the composite state. We are dealing with a *definition of local* or subsystem properties. By assumption (b) in the first stipulation, the probability law that we are endeavouring to obtain should be local.

The most important part of the precise mathematical formulation of the second stipulation is in terms of twin unitaries (cf (2a)).

Stage 1. No local unitary U_1 that has a twin U_2 must be able to change any local property. We know from theorem 1 that such local unitaries U_1 are all those that commute with ρ_1, and no others. In this way the mathematical entity ρ_1 is already beginning to obtain some physical relevance for local properties.

We know from the final theorem that we are dealing with U_1 that are orthogonal sums of *arbitrary* unitaries acting within the positive-eigenvalue eigensubspaces $\mathcal{R}(Q_1^j)$ of ρ_1 (cf (13a)).

Let $|\phi\rangle_1$ and $|\phi\rangle_1'$ be any two distinct state vectors from one and the same positive-eigenvalue eigensubspace $\mathcal{R}(Q_1^j)$ of ρ_1. Evidently, there exists a unitary U_1^j in this subspace that maps $|\phi\rangle_1$ into $|\phi\rangle_1'$, and, adding to it orthogonally any other eigensubspace unitaries (cf (13a)), one obtains a unitary U_1 in $\mathcal{H}_1$ that has a twin, i.e., the action of which can be given rise to from the remote second subsystem. ('Remote' here refers in a figurative way to lack of interaction. Or, to use Zurek's terms, 1 and 2 are assumed to be 'dynamically decoupled' and 'causally disconnected'.) Thus, we conclude that the two first-subsystem states at issue must have the *same probability*.

In other words, arguing *ab contrario*, if the probabilities of the two distinct states $|\phi\rangle_1, |\phi\rangle_1' \in \mathcal{R}(Q_1^j)$ were distinct, then, by remote action (by applying the twin unitary U_2 of the above unitary U_1 to $|\Psi\rangle_{12}$), one could transform one of the states into the other, which would locally mean changing the probability value without any local cause.

Putting our conclusion differently, all eigenvectors of ρ_1 that correspond to one and the same eigenvalue $r_j > 0$ *have one and the same probability* in $|\Psi\rangle_{12}$. Let us denote by $p(Q_1^j)$ the probability of the, in general, composite event that is mathematically represented by the eigenprojector Q_1^j corresponding to the eigenvalue r_j of ρ_1 (cf (10c)), and let the multiplicity of r_j (the dimension of $\mathcal{R}(Q_1^j)$) be d_j. Then the probability of $|\phi\rangle_1$ in ρ_1 is $p(Q_1^j)/d_j$. To see this, one takes a basis $\{|\phi_k\rangle_1 : k = 1, 2, \ldots, d_j\}$ spanning $\mathcal{R}(Q_1^j)$, or, equivalently, $Q_1^j = \sum_{k=1}^{d_j} |\phi_k\rangle_1\langle\phi_k|_1$, with, e.g., $|\phi_{k=1}\rangle_1 \equiv |\phi\rangle_1$. Further, one makes use of the *additivity rule of probability*: probability of the sum of *mutually exclusive* (orthogonal) events (projectors) equals the same sum of the probabilities of the event terms in it. Finally, one takes into account the fact that all d_j basis vectors have one and the same probability in ρ_1.

Actually, the σ-additivity rule of probability is the *third stipulation*. It requires that the probability of every finite or infinite sum of exclusive events be equal to the same sum of the

probabilities of the event terms. We could not proceed without it. (The need for infinite sums will appear four passages below.)

In the *special case*, when ρ_1 has only one positive eigenvalue of multiplicity d (the dimension of the range of ρ_1), the probability of $|\phi\rangle_1 \in \mathcal{R}(\rho_1)$ is $p(Q_1)/d$ (where Q_1 is the range projector of ρ_1.) To proceed, we need to evaluate $p(Q_1)$.

To this purpose, we make the *fourth stipulation*: every state vector $|\phi\rangle_1$ that belongs to the *null space* of ρ_1 (or, equivalently, when $|\phi\rangle_1\langle\phi|_1$, acting on $|\Psi\rangle_{12}$, gives zero) has *probability zero*. (The twin unitaries do not influence each other in the respective null spaces, cf (13a, b). Hence, this assumption is independent of the second stipulation.)

Justification for the fourth stipulation lies in Zurek's original framework. Namely, if the opposite subsystem is the environment, which establishes measurement-like entanglement, then the Schmidt states, e.g., the above eigen-sub-basis, obtain partners in a Schmidt decomposition (cf (5)), and this leads to measurement. States from the null space do not appear in this, and cannot give a positive measurement result.

One has $1_1 = Q_1 + \sum_l |l\rangle_1\langle l|_1$, where $\{|l\rangle_1 : \forall l\}$ is a basis spanning the null space of ρ_1, which may be infinite dimensional. Then, $p(Q_1) = p(1_1) = 1$ follows from the third stipulation (σ-additivity) and the fourth one. Finally, in the above special case of only one positive eigenvalue of ρ_1, the probability of $|\phi\rangle_1 \in \mathcal{R}(\rho_1)$ is $1/d$, which equals the only positive eigenvalue of ρ_1 in this case.

Our next aim is to derive $p(Q_1^j)$ in a more general case.

Stage 2. In this stage we confine ourselves to composite state vectors $|\Psi\rangle_{12}$ (i) that have finite entanglement, i.e., the first-subsystem reduced density operator of which has a finite-dimensional range, (ii) such that each eigenvalue r_j of ρ_1 is a rational number.

We rewrite the eigenvalues with an equal denominator: $\forall j : r_j = m_j/M$. Since $\sum_j d_j r_j = 1$, one has $\sum_j d_j m_j = M$ (d_j being the multiplicity of r_j).

Now we assume that $|\Psi\rangle_{12}$ has a special structure:

(i) The opposite subsystem 2 is bipartite in turn, hence we replace the notation 2 by (2+3), and $|\Psi\rangle_{12}$ by $|\Phi\rangle_{123}$.

(ii) (a) We introduce a two-indices eigen-sub-basis of ρ_1 spanning the closed range $\bar{\mathcal{R}}(\rho_1) : \{|j, k_j\rangle_1 : k_j = 1, 2, \ldots, d_j; \forall j\}$ so that the sub-basis is, as one says, adapted to the spectral decomposition $\rho_1 = \sum_j r_j Q_1^j$ of the reduced density operator, i.e., $\forall j : Q_1^j = \sum_{k_j=1}^{d_j} |j, k_j\rangle_1\langle j, k_j|_1$.

(b) We assume that $\mathcal{H}_2$ is at least M dimensional, and we introduce a basis $\{|j, k_j, l_j\rangle_2 : l_j = 1, 2, \ldots, m_j; k_j = 1, 2, \ldots, d_j; \forall j\}$ spanning a subspace of $\mathcal{H}_2$.

(c) We assume that also $\mathcal{H}_3$ is at least M dimensional, and we introduce a basis $\{|j, k_j, l_j\rangle_3 : l_j = 1, 2, \ldots, m_j; k_j = 1, 2, \ldots, d_j; \forall j\}$ spanning a subspace of $\mathcal{H}_3$.

(d) Finally, we define via a canonical Schmidt decomposition $1 + (2 + 3)$ (cf (7a)):

$$|\Phi\rangle_{123} \equiv \sum_j \sum_{k_j=1}^{d_j} (m_j/M)^{1/2} \left[|j, k_j\rangle_1 \otimes \left(\sum_{l_j=1}^{m_j} (1/m_j)^{1/2}|j, k_j, l_j\rangle_2 |j, k_j, l_j\rangle_3 \right) \right]. \quad (17a)$$

Equivalently,

$$|\Phi\rangle_{123} \equiv \sum_j \sum_{k_j=1}^{d_j} \sum_{l_j=1}^{m_j} (1/M)^{1/2}|j, k_j\rangle_1 |j, k_j, l_j\rangle_2 |j, k_j, l_j\rangle_3. \quad (17b)$$

Viewing (17b) as a state vector of a bipartite $(1+2)+3$ system, we see that it is a canonical Schmidt decomposition (cf (7a)). Having in mind (7b), and utilizing the final conclusion of stage one, we can state that the probability of each state vector $|j, k_j\rangle_1 |j, k_j, l_j\rangle_2$ is $1/M$.

On the other hand, we can view ($17a$) as a state vector of the bipartite system $1 + (2 + 3)$ in the form of a canonical Schmidt decomposition. One can see that $\forall j$, $\left(Q_1^j \otimes 1_2\right)$ and $\sum_{k_j=1}^{d_j} \sum_{l_j=1}^{m_j} |j, k_j\rangle_1 \langle j, k_j|_1 \otimes |j, k_j, l_j\rangle_2 \langle j, k_j, l_j|_2$ act equally on $|\Phi\rangle_{123}$. On the other hand, it is easily seen that the former projector can be written as a sum of the latter sum of projectors and of an orthogonal projector that acts as zero on $|\Phi\rangle_{123}$, and therefore has zero probability on account of stipulation four. Thus, $\left(Q_1^j \otimes 1_2\right)$ and the above sum have equal probabilities, which is

$$p\left(Q_1^j \otimes 1_2\right) = d_j m_j / M. \tag{18}$$

As was concluded in stage 1, the probability of any state vector $|\phi\rangle_1$ in $\mathcal{R}\left(Q_1^j\right)$ is $p\left(Q_1^j\right)/d_j$. The projectors Q_1^j and $\left(Q_1^j \otimes 1_2\right)$ stand for the same event (viewed locally and more globally respectively); hence they have the same probability in $|\Phi\rangle_{123}$. Thus, $p(|\phi\rangle_1) = m_j/M = r_j$, i.e., it equals the corresponding eigenvalue of ρ_1.

We see that also the eigenvalues, not just the eigensubspaces, i.e., the entire operator ρ_1 is relevant for the local probability. At this stage we do not yet know if we are still lacking some entity or entities. We will write X for the possible unknown.

How do we justify replacing $|\Psi\rangle_{12}$ by $|\Phi\rangle_{123}$? In the state space $(\mathcal{H}_2 \otimes \mathcal{H}_3)$ there is a pair of orthonormal sub-bases of $d = \sum_j d_j$ vectors that appear in ($17a$). Evidently, there exists a unitary operator U_{23} that maps the Schmidt-state partners $|j, k_j\rangle_2$ of $|j, k_j\rangle_1$ in $|\Psi\rangle_{12}$ tensorially multiplied with an initial state $|\phi_0\rangle_3$ into the vectors:

$$\forall k_j, \forall j: \quad U_{23}: \quad |j, k_j\rangle_2 |\phi_0\rangle_3 \longrightarrow \sum_{l_j=1}^{m_j} (1/m_j)^{1/2} |j, k_j, l_j\rangle_2 |j, k_j, l_j\rangle_3. \tag{19}$$

On account of the second stipulation, any such U_{23}, which transforms by interaction an ancilla (subsystem 3) in state $|\phi_0\rangle_3$ and subsystem 2 as it is in $|\Psi\rangle_{12}$ into the $(2 + 3)$-subsystem state as it is $|\Phi\rangle_{123}$, does not change any local property of subsystem 1. Hence, it does not change the probabilities either.

Stage 3. We make the *fifth stipulation*: the probability formula that we are trying to derive is *continuous* in ρ_1, i.e., if $\rho_1 = \lim_{n \to \infty} \rho_1^n$, then $p(E_1, \rho_1, X) = \lim_{n \to \infty} p\left(E_1, \rho_1^n, X\right)$, for every event (projector) E_1. (We assume that X, if it exists, does not change in the convergence process.)

Let $\rho_1 = \sum_{j=1}^J r_j Q_1^j$, J a natural number, be the spectral form of an arbitrary density operator with finite-dimensional range. One can write $\rho_1 = \lim_{n \to \infty} \rho_1^n$, where $\rho_1^n = \sum_{j=1}^J r_j^n Q_1^j$, with $r_j = \lim_{n \to \infty} r_j^n$, $j = 1, 2, \ldots, J$, and all r_j^n are rational numbers. (Note that the eigenprojectors Q_1^j are assumed to be the same all over the convergence.) Then the required continuity gives for an eigenvector $|r_{j_0}\rangle$ of ρ_1 corresponding to the eigenvalue r_{j_0}: $p\left(|r_{j_0}\rangle, \rho_1, X\right) = \lim_{n \to \infty} p\left(|r_{j_0}\rangle, \rho_1^n, X\right) = r_{j_0}$. This extends the conclusion of stage two to *all ρ_1 with finite-dimensional ranges*, and their eigenvectors.

Let $\rho_1 = \sum_{j=1}^{\infty} r_j Q_1^j$ have an infinite-dimensional range. We define $\rho_1^n \equiv \sum_{j=1}^n \left(r_j / \left(\sum_{k=1}^n r_k\right)\right) Q_1^j$. (Note that we are taking the same eigenprojectors Q_1^j.) Then $\rho_1 = \lim_{n \to \infty} \rho_1^n$, and for any eigenvector $|r_{j_0}\rangle$ one has $p(|r_{j_0}\rangle, \rho_1, X) = \lim_{n \to \infty} p(|r_{j_0}\rangle, \rho_1^n, X) = \lim_{n \to \infty} r_{j_0} / \left(\sum_{k=1}^n r_k\right) = r_{j_0}$. This extends the conclusion of the preceding stage to *all reduced density operators*.

As a final remark about stage three, we point out that the continuity postulated is meant with respect to the so-called strong operator topology in Hilbert space [19]. Thus, if $A = \lim_{n \to \infty} A_n$ in this topology (A and A_n any bound linear operators), then, and only then, for every vector $|\psi\rangle$ one has $A|\psi\rangle = \lim_{n \to \infty} A_n |\psi\rangle$. This, in turn, means, as well

known, that $\lim_{n\to\infty}\|A|\psi\rangle - A_n|\psi\rangle\| = 0$ (where the 'distance' in the Hilbert space is made use of).

Stage 4. The result of the preceding stages can be put as follows: if $\rho_1|\phi\rangle_1 = r|\phi\rangle_1$, then the probability is

$$p(|\phi\rangle_1, \rho_1) = r = \langle\phi|_1\rho_1|\phi\rangle_1. \tag{20}$$

(We have dropped X because we already know that, as far as eigenvectors of ρ_1 are concerned, nothing is missing.) Now we wonder what about state vectors $|\phi\rangle_1$ that are *not eigenvectors of* ρ_1. Entanglement in a composite state vector $|\Psi\rangle_{12}$ that implies ρ_1 through $\rho_1 \equiv \mathrm{tr}_2(|\Psi\rangle_{12}\langle\Psi|_{12})$ is of no further use. We must change $|\Psi\rangle_{12}$ and ρ_1 in a suitable manner.

Our *sixth stipulation* consists of the assumption that the least invasive non-demolition measurement, i.e., the so-called *ideal measurement*, of the observable

$$A_1 \equiv |\phi\rangle_1\langle\phi|_1 \tag{21}$$

is performable in the state ρ_1. (Earlier used synonyms for 'non-demolition' are 'repeatable', 'predictive' and 'of the first kind'.) By this measurement one can disregard the fact that ρ_1 is the state of a subsystem of $|\Psi\rangle_{12}$ (as proved in appendix D).

As well known [20], in ideal measurement the Lüders change-of-state formula is applicable, which in our case reads

$$\begin{aligned}\rho_1 \to \rho_1' &\equiv (|\phi\rangle_1\langle\phi|_1)\rho_1(|\phi\rangle_1\langle\phi|_1) + (|\phi\rangle_1\langle\phi|_1)^\perp\rho_1(|\phi\rangle_1\langle\phi|_1)^\perp \\ &= (\langle\phi|_1\rho_1|\phi\rangle_1)|\phi\rangle_1\langle\phi|_1 + (|\phi\rangle_1\langle\phi|_1)^\perp\rho_1(|\phi\rangle_1\langle\phi|_1)^\perp.\end{aligned} \tag{22}$$

Since ρ_1' describes both individual systems and ensembles, the very concept of measurement requires that the statistical weights in the mixture (22), which is the final state of the measurement interaction, be equal to the corresponding probabilities; in particular,

$$p(|\phi\rangle_1, \rho_1) = \langle\phi|_1\rho_1|\phi\rangle_1. \tag{23}$$

Thus, (20) is valid also for states $|\phi\rangle_1$ that are not eigenstates of ρ_1.

It should be, perhaps, noted that the Lüders change-of-state formula (22), which is characteristic for ideal measurement, has been shown to be equivalent to minimal change of state [21]. By this the closest state among states to which non-demolition measurement can lead is taken using the distance in the metric space of Hilbert–Schmidt operators, to which density operators belong. (From the mathematical point of view, however, this is not the natural metric space for density operators, because also the more restricted ideal of trace-class operators, of which the density operators also form a subset, is a metric space as every vector space with a norm. The larger ideal of Hilbert–Schmidt operators is the completion of the ideal of trace class operators in the Hilbert–Schmidt norm (cf appendix C)[1].)

Stage 5. Finally, we have to find out what the probability law should be when ρ is not an improper, but a proper mixture [22], i.e., *when there are no correlations with another system*.

Though these two kinds of mixtures are physically very different on the global level, it is known that locally there is no difference. We could therefore just stipulate this fact and thus extend the local probability law unchanged to proper mixtures. But there is an alternative that is based on a weaker assumption.

We make the *seventh stipulation* assuming that the continuity assumption (the fifth stipulation) is valid without the need to distinguish proper and improper mixtures locally, i.e., in subsystem 1.

We take first an isolated pure state $|\psi\rangle$.

---

[1] I owe this insight to an unknown board member, who was kindly adjudicating the references for this paper.

We start with an infinite sequence of correlated bipartite state vectors $\{|\Psi_{12}\rangle^n : n = 1, 2, \ldots, \infty\}$ such that, as far as the reduced density operator is concerned, one has

$$\forall n : \quad \rho_1^n = (1 - 1/n)|\psi\rangle_1\langle\psi|_1 + (|\psi\rangle_1\langle\psi|_1)^\perp \rho_1^n (|\psi\rangle_1\langle\psi|_1)^\perp, \tag{24}$$

where $|\psi\rangle_1$ actually equals $|\psi\rangle$. (It is well known that for every density operator ρ_1 there exists a state vector $|\Psi\rangle_{12}$ such that $\rho_1 = \text{tr}_2(|\Psi\rangle_{12}\langle\Psi|_{12})$. This claim is easily proved using the spectral form (7b) of ρ_1 and the canonical Schmidt decomposition (7a).) We now write index 1 because we now do have correlations with subsystem 2.

Obviously

$$|\psi\rangle_1\langle\psi|_1 = \lim_{n\to\infty} \rho_1^n. \tag{25}$$

According to our fifth stipulation, the probability formula is continuous in the density operator. Hence,

$$\forall|\phi\rangle : \quad p(|\phi\rangle, |\psi\rangle\langle\psi|) = \lim_{n\to\infty} p(|\phi\rangle_1, \rho_1^n) = \lim_{n\to\infty} \langle\phi|_1\rho_1^n|\phi\rangle_1 = \langle\phi|_1 \lim_{n\to\infty} \rho_1^n|\phi\rangle_1.$$

This finally gives

$$\forall|\phi\rangle : \quad p(|\phi\rangle, |\psi\rangle\langle\psi|) = \langle\phi|(|\psi\rangle\langle\psi|)|\phi\rangle = |\langle\phi||\psi\rangle|^2. \tag{26}$$

In this way, the same probability formula is extended to isolated pure states.

If ρ is *an isolated mixed* state, i.e., a proper mixture, one can take any of its (infinitely many) decompositions into pure states, say,

$$\rho = \sum_k w_k|\psi_k\rangle\langle\psi_k|,$$

where w_k are the statistical weights ($\forall k : w_k > 0; \sum_k w_k = 1$). Then

$$p(|\phi\rangle, \rho) = \sum_k w_k\langle\phi|(|\psi_k\rangle\langle\psi_k|)|\phi\rangle.$$

This finally gives

$$p(|\phi\rangle, \rho) = \langle\phi|\rho|\phi\rangle, \tag{27}$$

extending the same probability formula to mixed isolated states. (It is obvious that the choice of the above decomposition into pure states is immaterial. One can take the spectral decomposition, e.g.)

The form $\langle\phi|\rho|\phi\rangle$ of the probability law obtained, following Zurek, in the present version is *equivalent to the* (much more generally looking) *trace rule* on account of σ-additivity. Namely, if we take any event E and write it as a sum of orthogonal ray projectors $E = \sum_{i=1} |i\rangle\langle i|$, additivity or σ-additivity (in the case of a finite or an infinite sum respectively) enables us to transform the present form into the trace rule:

$$p(E, \rho) = \sum_{i=1}\langle i|\rho|i\rangle = \sum_{i=1}\text{tr}(\rho|i\rangle\langle i|) = \text{tr}\left(\rho\left(\sum_{i=1}|i\rangle\langle i|\right)\right) = \text{tr}(\rho E). \tag{28}$$

On the other hand, it is obvious that the trace rule reduces to $\langle\phi|\rho|\phi\rangle$ in the case of elementary events (ray projectors) $E \equiv |\phi\rangle\langle\phi|$.

4. Final remarks

In the long derivation of the preceding section some points require additional comments. They are displaced into this final section in an attempt to keep the derivation with as few deviations as possible.

4.1. Non-contextuality of probability in terms of events

This remark should shed light on an aspect of probabilities of events that is usually, and also in the preceding section, taken for granted.

(a) *The event non-contextuality.* On account of the implication relation in the structure of all events (the projector E implies the projector F, i.e., $E \leqslant F$ if and only if $EF = E$), every composite event can occur as a consequence of the occurrence of different elementary events that imply it. Nevertheless, *the probability does not depend on this*.

As a matter of fact, the probabilities of the composite events are in section 3 of this paper, following Zurek, defined in terms of mutually exclusive elementary events (orthogonal ray-projectors, each defined by a state vector) using σ-additivity. This implies event non-contextuality of probability.

(b) *Non-contextuality with respect to observables.* A given elementary (or composite) event can, in general, be the eigenevent (eigenprojector) of different observables. (This, essentially, amounts to the so-called eigenvalue–eigenstate link.) Correspondingly, the event can occur in measurement of different observables. The probability of the event does not depend on this. This non-contextuality has not been utilized in the present quantum logical approach, in which, by definition, observables more general than events (projectors) are avoided.

4.2. How much probability has been put into the derivation?

In their concluding remarks Schlosshauer and Fine [2] say:

> '…a fundamental statement about any probabilistic theory: we cannot derive probabilities from a theory that does not already contain some probabilistic concept; at some stage, we need to 'put probabilities in to get probabilities out'.

We have 'put in' seven stipulations. More will be said on the 'price' that has been 'paid' for the probability law in the next subsection.

4.3. Zurek-like derivation and Gleason's theorem

In an effort to make Zurek's argument more transparent, his 'small natural' and some tacit assumptions have been avoided as much as possible. The most disquieting consequence was raising σ-additivity to the status of a stipulation. This was no different than in Gleason's well-known theorem [8], which goes as follows.

One assumes that one has a map associating a number p from the doubly closed interval $[0, 1]$ with every subspace, or, equivalently, with every projector E (projecting onto a subspace) observing σ-additivity, i.e.

$$p\left(\sum_i E_i\right) = \sum_i p(E_i) \tag{29a}$$

for every orthogonal decomposition (finite or infinite) of every projector. Then, for every such map, there exists a unique density operator ρ such that

$$p(E, \rho) = \mathrm{tr}(E\rho) \tag{29b}$$

for every projector (the trace rule). Thus, the set of all density operators and that of all quantum probabilities stand in a natural and simple one-to-one relation.

In spite of the generality and elegance of Gleason's theorem, any Zurek-like derivation is very valuable. Perhaps a famous dictum of Wigner can help to make this clear. When

faced with the challenge of computer simulations to replace analytical solutions of intricate equations of important physical meaning, Wigner has allegedly said 'I am glad that your computer understands the solutions; but I also would like to understand them.'

Schlosshauer and Fine say (in the introduction to their paper):

'...Gleason's theorem is usually considered as giving rather little physical insight into the emergence of quantum probabilities and the Born rule.'

Gleason gives the complete answer at once in the form of the trace rule. One can then derive from it the other six postulates of the present version and more. To use Wigner's words, the mathematics in the proof of Gleason's theorem 'understands' the uniqueness and the other wonders of the quantum probability law, but we do not.

Now, the extra 6 stipulations in the present version (besides σ-additivity), though logically unnecessary in view of Gleason's theorem, nevertheless, thanks to Zurek's ingenuity, help to unfold before our eyes the simplicity and full generality of the quantum probability law.

4.4. The role of entanglement

In this remark a critical view is taken of the role of entanglement in the present derivation.

Entanglement enters through the sole entanglement entity of the approach—the correlation operator U_a (see the correlated subsystem picture in subsection 2.3.). In terms of this entity a complete answer was obtained to the questions which unitaries have a twin, and which opposite-subsystem unitary is the (unique) twin.

However, in section 3, in unfolding the present version, the correlation operator (and hence entanglement) was not made use of at all. All that was utilized was the general form of a first-subsystem unitary that has a twin: $U_1 = \sum_j U_1^j Q_1^j + U_1' Q_1^\perp$, where $1_1 = \sum_j Q_1^j + Q_1^\perp$ is the eigenresolution of the unity with respect to (distinct eigenvalues) of the reduced density operator $\rho_1 (\equiv \mathrm{tr}_2(|\Psi\rangle_{12}\langle\Psi|_{12}))$, and $\forall j : U_1^j$ is an *arbitrary* unitary in the eigensubspace $\mathcal{R}(Q_1^j)$ corresponding to the positive eigenvalue r_j of ρ_1 (cf subsection 2.4.).

These unitaries (Zurek's envariance unitaries) are utilized to establish what are local or first-subsystem properties, in particular, local probabilities. It immediately follows that any two distinct eigenvectors corresponding to the same eigenvalue of ρ_1 determine equal-probability events (cf Stage one in section 3). Thus, envariance is made use of in the first and most important step of Zurek's argument without 'putting in probability' (cf subsection 4.2).

Nevertheless, twin unitaries (envariance) is due to entanglement, and Zureks argument is based on the latter. Entanglement is, as well known, the basic stuff of which quantum communication and quantum computation are made of. No wonder that entanglement is increasingly considered to be a fundamental physical entity. As an illustration for this, one may mention that preservation of entanglement has been proposed as an equivalent second law of thermodynamics for composite systems (cf [23] and the references therein).

4.5. σ-additivity

It was argued in the preceding section that the presented derivation cannot work without the assumption of σ-additivity. Zurek, on the other hand, has a somewhat different claim: on pp 18 and 19 [7] he says:

'To demonstrate Lemma 5 (a key step in his endeavour, FH) we need one more property—the fact that when a certain event $\mathcal{U}(p(\mathcal{U}) = 1)$ can be decomposed into two mutually exclusive events, $\mathcal{U} = k \vee k^\perp$, their probability must add up to unity:

$$p(\mathcal{U}) = p(k \vee k^\perp) = p(k) + p(k^\perp) = 1.$$

This assumption introduces (in a very limited setting) additivity. It is equivalent to the statement that 'something will certainly happen'.

To get an idea how full σ-additivity differs from Zureks 'very limited' assumption, we put the former in the form of a 'staircase' of gradually strengthened partial assumptions.

The starting point is the fact that if any event E occurs, the opposite event $E^\perp(\equiv (1 - E))$ does not occur (in suitable measurement, of course).

(1) It is plausible to assume that $E + E^\perp = 1$ has $p(E) + p(E^\perp) = 1$ as its consequence in any quantum state.

(2) If $E + F = G$ (all being events, i.e., projectors, and $EF = 0$), then, in view of the fact that, e.g., F is the opposite event of E in G, i.e., $F = E^\perp G$, and in view of assumption (1), it is plausible to assume that $E + F = G$ implies $p(E) + p(F) = p(G)$ in any quantum state. Obviously, assumption (2) is a strengthening of assumption (1).

Lemma. *Assumption (2) implies additivity for every finite orthogonal sum of events:* $\sum_i E_i = G \Rightarrow \sum_i p(E_i) = p(G)$ *in any quantum state.*

Proof. If the lemma is valid for n terms, then

$$p\left(\sum_{i=1}^{(n+1)} E_i\right) = p\left(\left(\sum_{i=1}^{n} E_i\right) + E_{(n+1)}\right) = p\left(\sum_{i=1}^{n} E_i\right) + p(E_{(n+1)}) = \sum_{i=1}^{(n+1)} p(E_i),$$

i.e., it is valid also for $(n + 1)$ terms. By assumption, it is valid for two terms. By total induction, it is then valid for every finite sum. $\square$

(3) If one has the projector relation $G = \lim_{n\to\infty} F_n$ and the projector sequence $\{F_n : n = 1, 2, \ldots, \infty\}$ is non-descending ($\forall n : F_{(n+1)} \geqslant F_n \Leftrightarrow F_{(n+1)} F_n = F_n$), then the assumption of *continuity* in the probability $p(G) = \lim_{n\to\infty} p(F_n)$ is plausible (otherwise one could have jumps in probability and no event responsible for it). Assuming the validity of assumption (2), it implies

$$p\left(\sum_{i=1}^{\infty} E_i\right) \equiv p\left(\lim_{n\to\infty} \sum_{i=1}^{n} E_i\right) = \lim_{n\to\infty} \sum_{i=1}^{n} p(E_i) \equiv \sum_{i=1}^{\infty} p(E_i),$$

i.e., σ-additivity ensues.

If one wants to estimate how 'steep' each of these 'stairs' is, one is on intuitive ground burdened with feeling and arbitrariness. Assumption (1) seems to be the largest 'step' (with respect to the stated fact that is its premise). Once (1) is given, assumption (2) (equivalent to additivity of probability) seems very natural, hence less 'steep'. The final assumption (3) seems even more natural, and hence least 'steep'.

Let us revisit the Zurek-like argument with respect to σ-additivity. Let $|\Psi\rangle_{12}$ be infinitely entangled, or, equivalently, let $\rho_1 \equiv \text{tr}_2(|\Psi\rangle_{12}\langle\Psi|_{12})$ have an infinite-dimensional range. Further, let $\{|i\rangle_1 : i = 1, 2, \ldots, \infty\}$ be a set of eigenvectors of ρ_1 corresponding to different eigenvalues, but let them not span the whole range $\bar{\mathcal{R}}(\rho_1)$. Without the validity of σ-additivity we cannot infer the probability $p(E_1, \rho_1)$, where $E_1 \equiv \sum_{i=1}^{\infty} |i\rangle_1\langle i|_1$. Thus, if one wants the general form of the probability law, and in the present version nothing less is wanted, then one must make the above assumption (2) and the continuity assumption in (3) (assumptions of which the σ-additivity concept consists).

4.6. Why unitary operators?

Both envariance and its other face, unitary twins, are expressed in terms of unitary operators. One can raise the question in the title of the subsection.

The answer lies in the notion of *distant influence* and in the notion of Zurek's 'Fact 1' (see the sixth paragraph in section 3). One assumes that the nearby subsystem 1 is dynamically decoupled from another subsystem 2, but not statistically. Entanglement is assumed to exist between the two subsystems. On account of the quantum correlations that are established, one can manipulate subsystem 2 in order to make changes in subsystem 1 (without interaction with it). By definition, local are those properties of the nearby subsystem that cannot be changed by the described distant influence. Probabilities of events on subsystem 1 were stipulated to be local.

One is thinking in terms of so-called bare quantum mechanics, i.e., quantum mechanics without collapse. Then all conceivable manipulations of the distant subsystem are unitary evolutions (suitable interactions of suitably chosen subsystems—all without any interaction with subsystem 1). As Zurek puts it in his 'fact 1': 'Unitary transformations must act on the system to alter its state.' (This goes for the distant subsystem which should exert the distant influence.)

Unitary evolution preserves the total probability of events. The suspicion has been voiced that restriction to unitary operators might just be a case of 'putting in probability in order to get out probability'[2]. Even if this is so, it appears to be even milder than Zurek's 'putting in' probability-one assumptions (cf last passage in subsection B.1 in [7]).

One may try to argue that the unitarity of the evolution operator (of the dynamical law) does not contain any probability assumption. Namely, one may start with the Schrödinger equation, of which the unitary evolution operator is the integrated form (from instantaneous tendency of change in a finite interval). At first glance, the Schrödinger equation has nothing to do with probabilities. But this is not quite so. The dynamical law, instantaneous or for a finite interval, gives the change of the quantum state, which is, in turn, equivalent to the totality of probability predictions (cf Gleason's theorem in subsection 4.3.).

Perhaps the 'fundamental statement' quoted in subsection 4.2. is true, and one should not expect to derive probabilities exclusively from notions completely different from them.

Acknowledgments

Appendix A

We prove now that the correlation operator U_a is independent of the choice of the eigen-sub-basis of ρ_1 (cf (7b)) that spans $\bar{\mathcal{R}}(\rho_1)$ in which the strong Schmidt decomposition of $|\Psi\rangle_{12}$ (cf (9)) is written.

---

[2] The question in the title of the subsection was raised by Schlosshauer (in private communication). He voiced the suspicion that restriction to unitary operators might be a way of 'putting in probabilities to get out probabilities'.

Let $\{|j,k_j\rangle_1 : \forall k_j, \forall j\}$ and $\{|j,l_j\rangle_1 : \forall l_j, \forall j\}$ be two arbitrary eigen-sub-bases of ρ_1 spanning $\mathcal{R}(\rho_1)$. The vectors are written with two indices, j denoting the eigensubspace $\mathcal{R}(Q_1^j)$ to which the vector belongs, and the other index k_j (l_j) enumerates the vectors within the subspace.

The proof goes as follows. Let

$$\forall j : \quad |j,k_j\rangle_1 = \sum_{l_j} U_{k_j,l_j}^{(j)} |j,l_j\rangle_1,$$

where $\left(U_{k_j,l_j}^{(j)}\right)$ are unitary sub-matrices. Then, keeping U_a one and the same, we can start out with the strong Schmidt decomposition in the k_j-eigen-sub-basis, and after a few simple steps (utilizing the antilinearity of U_a and the unitarity of the transition sub-matrices), we end up with the strong Schmidt decomposition (of the same $|\Psi\rangle_{12}$) in the l_j-eigen-sub-basis:

$$|\Psi\rangle_{12} = \sum_j \sum_{k_j} r_j^{1/2} |j,k_j\rangle_1 (U_a|j,k_j\rangle_1)_2$$

$$= \sum_j \sum_{k_j} \left\{ r_j^{1/2} \left(\sum_{l_j} U_{k_j,l_j}^{(j)} |j,l_j\rangle_1 \right) \otimes \left[U_a \left(\sum_{l_j'} U_{k_j,l_j'}^{(j)} |j,l_j'\rangle_1 \right) \right]_2 \right\}$$

$$= \sum_j \sum_{l_j} \sum_{l_j'} \left\{ r_j^{1/2} \left(\sum_{k_j} U_{k_j,l_j}^{(j)} U_{k_j,l_j'}^{(j)*} \right) |j,l_j\rangle_1 \otimes (U_a|j,l_j'\rangle_1)_2 \right\}$$

$$= \sum_j \sum_{l_j} \sum_{l_j'} \left\{ r_j^{1/2} \delta_{l_j,l_j'} |j,l_j\rangle_1 \otimes (U_a|j,l_j'\rangle_1)_2 \right\}$$

$$= \sum_j \sum_{l_j} r_j^{1/2} |j,l_j\rangle_1 (U_a|j,l_j\rangle_1)_2.$$

Appendix B

We elaborate now the *group of pairs of unitary twins*.

Let (U_1', U_2') and (U_1, U_2) be two pairs of twin unitaries for a given bipartite state vector $|\Psi\rangle_{12}$, i.e., let $U_1'|\Psi\rangle_{12} = U_2'|\Psi\rangle_{12}$, and $U_1|\Psi\rangle_{12} = U_2|\Psi\rangle_{12}$, be valid. Then, applying U_2 to both sides of the former relation, exchanging the rhs and the lhs, and utilizing the latter relation, one has

$$U_2 U_2'|\Psi\rangle_{12} = U_2 U_1'|\Psi\rangle_{12} = U_1' U_2|\Psi\rangle_{12} = U_1' U_1|\Psi\rangle_{12}.$$

Hence, $(U_1' U_1, U_2 U_2')$ are twin unitaries, and one can define a composition law as $(U_1', U_2') \times (U_1, U_2) \equiv (U_1' U_1, U_2 U_2')$. (Note the inverse order in the second-subsystem operators.) Naturally, the trivial twin unitaries $(1_1, 1_2)$ are the unit element. The inverse of (U_1, U_2) is $\left(U_1^{-1}, U_2^{-1}\right)$ (cf the proposition in subsection 2.).

Appendix C

Bounded linear operators A in a complex separable Hilbert space for which $\|A\|_1 \equiv \mathrm{tr}((A^\dagger A)^{1/2}) < \infty$ ($A^\dagger$ being the adjoint of A) are trace-class ones. The set of all trace-class operators is an ideal (and hence a vector space) with a norm. It implies a distance $d_1(A, A') \equiv \|A - A'\|_1$.

Those linear operators A are Hilbert–Schmidt ones for which $\mathrm{tr}(A^\dagger A) < \infty$. The scalar product in the Hilbert space $\mathcal{H}_{HS}$ of all linear Hilbert–Schmidt operators is $(A, B) \equiv \mathrm{tr}(A^\dagger B)$

(cf the definition after theorem VI.21 and problem VI.48(a) in [19]). The scalar product implies the norm $\|A\|_2 \equiv \operatorname{tr}(A^\dagger A)$, and this in turn implies a distance $d_2(A, A') \equiv \|A - A'\|_2$.

One has always $\|A\|_2 \leqslant \|A\|_1$, and hence $d_2(A, A') \leqslant d_1(A, A')$.

Appendix D

We assume that ideal measurement of the observable $(|\phi\rangle_1\langle\phi|_1 \otimes I_2)$ (I_2 being the identity operator in the state space of the second subsystem) in the state $|\Psi\rangle_{12}$ is performed, and we are interested in the change of the state of subsystem 1.

Then $\rho_1 (\equiv \operatorname{tr}_2(|\Psi\rangle_{12}\langle\Psi|_{12}))$ goes over into the state

$$\operatorname{tr}_2((|\phi\rangle_1\langle\phi|_1 \otimes I_2)(|\Psi\rangle_{12}\langle\Psi|_{12})(|\phi\rangle_1\langle\phi|_1 \otimes I_2)$$
$$+ (|\phi\rangle_1\langle\phi|_1 \otimes I_2)^\perp(|\Psi\rangle_{12}\langle\Psi|_{12})(|\phi\rangle_1\langle\phi|_1 \otimes I_2)^\perp)$$
$$= |\phi\rangle_1\langle\phi|_1\rho_1|\phi\rangle_1\langle\phi|_1 + (|\phi\rangle_1\langle\phi|_1)^\perp\rho_1(|\phi\rangle_1\langle\phi|_1)^\perp.$$

This justifies disregarding the entanglement and the second subsystem in $|\Psi\rangle_{12}$.

References

[1] Zurek W H 2003 *Phys. Rev. Lett.* **90** 120404 (*Preprint* quant-ph/0211037)
[2] Schlosshauer M and Fine A 2005 *Found. Phys.* **35** 197 (*Preprint* quant-ph/0312058)
[3] Barnum H 2003 No-signalling-based version of Zurek's derivation of quantum probabilities: a note on 'Environment-assisted invariance, entanglement and probabilities in quantum physics' *Preprint* quant-ph/0312150
[4] Mohrhoff U 2004 *Int. J. Quant. Inf.* **2** 221 (*Preprint* quant-ph/0401180)
[5] Caves C M Notes on Zurek's derivation of the quantum probability rule (http://info.phys.unm.edu/˜caves/reports/ZurekBornderivation.pdf)
[6] Herbut F 2006 'Complete 'Born's rule' from 'environment-assisted invariance' in terms of pure-state twin unitaries' *Preprint* quant-ph/0611220
[7] Zurek W H 2005 *Phys. Rev.* A **71** 052105 (*Preprint* quant-ph/0405161)
[8] Gleason A M 1957 *J. Math. Mech.* **6** 885
[9] Schlosshauer M 2004 *Rev. Mod. Phys.* **76** 1267 (*Preprint* quant-ph/0312059)
[10] Barnum H The many-worlds interpretation of quantum mechanics: psychological versus physical bases for the multiplicity of 'worlds', unpublished (http://info.phys.unm.edu/papers/ papers.html)
[11] Herbut F and Vujičić M 1971 A new development in the description of correlations between two quantum systems *Foundations of Quantum Mechanics, Proc. Int. School of Physics 'Enrico Fermi' course IL,* ed B D'Espagnat (New York: Academic) p 316
[12] Herbut F 2006 *J. Math. Phys.* **47** 122103 (*Preprint* quant-ph/0609073) (reprinted in *Virtual J. Quant. Inf.* Dec. 2006 issue)
[13] Paris M G A 2005 *Int. J. Quant. Inf.* **3** 655 (*Preprint* quant-ph/0502025)
[14] Herbut F and Vujičić M 1976 *Ann. Phys., NY* **96** 382
     Vujičić M and Herbut F 1984 *J. Math. Phys.* **25** 2253
     Herbut F and Vujičić M 1987 *J. Phys. A: Math. Gen.* **20** 5555
     Herbut F 2002 *Phys. Rev.* A **66** 052321 (*Preprint* quant-ph/0305187)
[15] Herbut F and Damnjanović M 2000 *J. Phys. A: Math. Gen.* **33** 6023 (*Preprint as* quant-ph/0004085)
     Herbut F 2002 *J. Phys. A: Math. Gen.* **35** 1691 (*Preprint* quant-ph/0305181)
     Herbut F 2003 *J. Phys. A: Math. Gen.* **36** 8479 (*Preprint* quant-ph/0309181)
[16] Herbut F 1990 *J. Phys. A: Math. Gen.* **23** 367
[17] Zurek W H 2003 *Rev. Mod. Phys.* **75** 715 (*Preprint* quant-ph/0105127)
[18] Zurek W H 2004 Quantum darwinism and envariance *Science and Ultimate Reality: From Quantum to Cosmos* ed J D Barrow, P C W Davies and C H Harper (Cambridge: Cambridge University Press) (*Preprint* quant-ph/0308163)
[19] Reed M and Simon B 1972 *Methods of Modern Mathematical Physics. Functional Analysis* vol 1 (New York: Academic) chapter VI, section 1
[20] Lüders G 1951 *Ann. Phys., Lpz.* **8** 322

Messiah A 1961 *Quantum Mechanics* vol I (Amsterdam: North-Holland) p 333

Cohen-Tannoudji C, Diu B and Laloe F 1977 *Quantum Mechanics* vol I (New York: Wiley) p 221 (fifth postulate)

[21] Herbut F 1969 *Ann. Phys., NY* **55** 271

[22] d'Espagnat B 1976 *Conceptual Foundations of Quantum Mechanics* 2nd edn (Reading, MA: Benjamin) subsection 7.2

[23] Horodecki M and Horodecki R 1998 *Phys. Lett.* A **244** 473