A quantum measure of coherence and incompatibility

2005 J. Phys. A: Math. Gen. 38 2959

(http://iopscience.iop.org/0305-4470/38/13/010)

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

J. Phys. A: Math. Gen. **38** (2005) 2959–2974 doi:10.1088/0305-4470/38/13/010

A quantum measure of coherence and incompatibility

F Herbut

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

E-mail: fedorh@infosky.net

Received 28 September 2004, in final form 10 January 2005
Published 14 March 2005
Online at stacks.iop.org/JPhysA/38/2959

Abstract
The well-known two-slit interference is understood as a special relation between an observable (localization at the slits) and a state (being on both slits). The relation between an observable and a quantum state is investigated in the general case. It is assumed that the amount of coherence equals that of the incompatibility between an observable and a state. On these grounds, an argument is presented that leads to a natural quantum measure of coherence, called 'coherence or incompatibility information'. Its properties are studied in detail, making use of 'the mixing property of relative entropy' derived in this paper. A precise relation between the measure of coherence of an observable and that of its coarsening is obtained and discussed from the intuitive point of view. Convexity of the measure is proved, and thus the fact that it is an information entity is established. A few more detailed properties of the coherence information are derived with a view to investigating the final-state entanglement in general repeatable measurement, and, more importantly, general bipartite entanglement in follow-ups of this study.

PACS numbers: 03.65.Ta, 03.67.Mn

1. Introduction

In a preceding paper [1] *coherence in a relative sense*, i.e., understood as a relation between a given observable and a given quantum state, was postulated to be *identical with the incompatibility* between the observable and the state as far as its *quantity* I_C is concerned. (For notation see the passage immediately following the proof of proposition 5 below.) Then it was shown that the bipartite pure state entanglement is expressible as I_C (with a suitable observable).

Pure states cannot be obtained as mixtures. Therefore, the question if I_C is concave, i.e., a genuine entropy quantity, or convex, i.e., a genuine information one, or something third, could not be put in this context. The first aim of this study is to clarify this point. (This is done in

0305-4470/05/132959+16$30.00

proposition 5.) To enable this, the mixing property of relative entropy (paralleling the mixing property of entropy and Donald's identity for relative entropy, see the remark) is derived.

In a follow-up of the mentioned paper [2] the special case of the final bipartite pure state $|\psi\rangle_{12}$ in repeatable measurement, when the initial state is pure, was studied. It was shown that the initial quantity of incompatibility between the measured observable and the initial state reappears as the amount of entanglement in $|\psi\rangle_{12}$, and is further preserved when it is shifted in reading the measurement result. This completes Vedral's result [3] that the information transfer from object (subsystem 1) to measuring apparatus (subsystem 2) does not exhaust the mutual information I_{12} in the final state.

I think that it is of interest to find out if the mentioned preservation of the quantity of incompatibility between the measured observable and the initial pure state is restricted to a pure state, or it can be generalized to a mixed initial state. This is not a straightforward generalization. It requires more knowledge on I_C. The second aim of this study is to provide such knowledge, which will be possible due to the mentioned auxiliary relative entropy relations (see section 3).

In a further preceding paper [4] an arbitrary discrete incomplete observable A and its completion A^c to a complete observable were investigated and it was shown that $I_C(A, \rho) \leqslant I_C(A^c, \rho)$ for any state ρ. This inequality is expected if the assumption on the identity of the amount of coherence and that of incompatibility is correct. But it is desirable to evaluate $I_C(A^c, \rho) - I_C(A, \rho)$ and thus to try to acquire more insight into the nature of I_C. This is the third aim of this paper. (See the discussion after the proof of the theorem below.)

The fourth aim of this paper is to present an argument that starts with the mentioned identity assumption and leads to an expression for the quantity of coherence in a natural way. Will this expression be the same as the *ad hoc* introduced one? This is done in section 2 and an affirmative answer is obtained. It is summed up in the conclusion (subsection 5.2).

The fifth and last aim of this investigation is perhaps the most important one. Namely, in [4] it was established that I_C also plays an important role in some mixed bipartite states. This line of research should be continued in a follow up because it may contribute to our understanding of how mutual information in general bipartite states breaks up into a quasi-classical part and entanglement, which is the object of study of a wide circle of researchers, e.g. [5, 6]. To this purpose, one may need more detailed knowledge of the properties of I_C. To acquire such knowledge is the fifth aim of this paper (see section 4).

1.1. Background in classical statistical physics

To obtain a background for our quantum study of coherence, we assume that a classical discrete variable $A(q) = \sum_l a_l \chi_l(q)$ is given (all $a_l \in \mathbb{R}$ being distinct). The symbol q denotes the continuous state variables (as a rule, it consists of twice as many variables as there are degrees of freedom in the system); χ_l are the characteristic functions $\forall l : \chi_l(q) \equiv 1$ if $q \in \mathcal{A}_l$, and zero otherwise. Naturally, $\mathcal{A}_l$ are (Lebesgue measurable) sets such that $A(q) = a_l$ if and only if $q \in \mathcal{A}_l$, and $\sum_l \mathcal{A}_l = \mathcal{Q}$, where $\mathcal{Q}$ is the entire state space (or phase space) and the sum is the union of disjoint sets.

Let $\rho(q)$ be a continuous probability distribution in $\mathcal{Q}$ with the physical meaning of a statistical 'state' of the system. One can think of $\rho(q)$ as a mixture

$$\rho(q) = \sum_l p_l \rho_l(q), \tag{1}$$

where $\forall l: \; p_l \equiv \int_Q \rho(q)\chi_l(q)\,\mathrm{d}q$ are the statistical weights (probabilities of the results a_l if $A(q)$ is measured in $\rho(q)$), and $\forall l,\, p_l > 0: \rho_l(q) \equiv \rho(q)\chi_l(q)/p_l$ are the 'states' with definite (or sharp) values of $A(q)$.

Let $B(q)$ be any other continuous or discrete variable. Then, utilizing (1), its average can be written

$$\langle B\rangle_\rho \equiv \int_Q \rho(q)B(q)\,\mathrm{d}q = \sum_l p_l \langle B\rangle_{\rho_l}. \tag{2}$$

One distinguishes the contributions of the individual eigenvalues a_l of $A(q)$ through the terms on the rhs. They each contribute to $\langle B\rangle_\rho$ separately.

All this serves only as a classical background to help us understand the non-classical, i.e., purely quantum relations between the analogous quantum entities.

1.2. Transition to the quantum mechanical case

The quantum mechanical analogues of the mentioned classical entities are the following.

Discrete observables (Hermitian operators) $A = \sum_l a_l P_l$ (spectral form in terms of distinct eigenvalues), ρ quantum state (density operator), and B an arbitrary observable (Hermitian operator). The quantum average is $\langle B\rangle_\rho \equiv \mathrm{tr}(\rho B)$.

In the transition from classical to quantum one runs into a surprise, that is known but, perhaps, not sufficiently well known. Before we formulate it in the form of a lemma, let us introduce the Lüders state ρ_L [7] in order to obtain the quantum analogues of relations (1) and (2). It is that mixture of states, each with a definite value of A, which has a *minimal* Hilbert–Schmidt distance from the given state ρ [8]. It is defined as

$$\rho_L \equiv \sum_l p_l \rho_L^l, \tag{3a}$$

where

$$\forall l: \quad p_l \equiv \mathrm{tr}(\rho P_l) \tag{3b}$$

are again the statistical weights in (3*a*) (or the probabilities of the results a_l when A is measured in ρ), and

$$\forall l,\, p_l > 0: \qquad \rho_L^l \equiv P_l \rho P_l / p_l \tag{3c}$$

are the states with definite values a_l of A. Finally,

$$\langle B\rangle_{\rho_L} = \sum_l p_l \langle B\rangle_{\rho_L^l}. \tag{3d}$$

Decomposition (3*a*) is the analogue of (1), and (3*d*) is that of (2).

Lemma 1. *The following four statements are equivalent.*

(i) The state ρ cannot be written as a mixture of states in each of which the observable A has a definite value.

(ii) The observable A and the state ρ are incompatible, i.e., the operators do not commute $[A, \rho] \neq 0$.

(iii) The Lüders state ρ_L given by (3a)–(3c) is distinct from the original state ρ.

(iv) There exists an observable B such that

$$\langle B\rangle_\rho \neq \langle B\rangle_{\rho_L}, \tag{4}$$

where the rhs is given by (3d).

Proof is given in appendix 1.

The physical meaning of lemma 1 is that it defines a kind of *quantum coherence* as a special relation between an observable and a state. Experimentally it is exhibited in *interference*. In this relative sense (relation between a variable and a state) it is lacking in classical physics because there a state can always be written as a mixture of states in each of which the variable in question has a definite value (negation of (i), cf (1)). Though classical waves do exhibit a kind of coherence and show interference, but this is in a different sense (cf section 5).

One should note that the Lüders state needs no other characterization than its role in lemma 1 (in particular (iii)). The fact that it is 'closest' to ρ in Hilbert–Schmidt metrics, though actually not important for this study, raises the thought-provoking questions if 'closest' is also true in other metrics; if not, why are the Hilbert–Schmidt metrics more suitable?

We take *two-slit interference* [9] to serve as an illustration for lemma 1.

Let A be a dichotomic position observable with two eigenvalues: localization at the left slit, and localization at the right slit on the first screen. Let ρ be a wave packet that has just arrived at this two-slit screen. Next, one has to find a suitable observable B such that inequality (4) be satisfied at the mentioned moment. Moreover, one wants to observe experimentally the lhs of (4), or rather the individual probabilities of the eigenvalues of B (that go into the lhs).

To this purpose, one actually replaces B by another localization observable A' on a second screen, to which the photon will arrive some time later. This observable is suitable for observation (of its localization probabilities). Hence, one can define $B \equiv U^{-1}A'U$, U being the evolution operator expressing the movement of the particle from the two-slit screen to the second one. One should note that B is not a position observable though A' is because the Hamiltonian that generates U contains the kinetic energy (square of linear momentum).

Claim (i) of lemma 1 says that the particle is not moving through either the left or the right slit. Claim (ii) expresses the same fact algebraically. Namely, ρ, being a pure state $|\psi\rangle\langle\psi|$, would commute with A only if $|\psi\rangle$ lay in an eigensubspace of A. In our case this would mean that the particle traverses one of the slits.

The Lüders state ρ_L is, in some sense, the best approximation to ρ of a state traversing one or the other of the slits. Naturally, $\rho \neq \rho_L$ as claimed by (iii). Claim (iv), i.e., relation (4), amounts to the same as the fact that the interference pattern on the second screen is not equal to the sum of those that would be obtained when only one of the slits was open (for some time) and then the other (for another, disjoint, equally long time).

In the two-slit experiment one actually observes the time-delayed equivalent of (4):

$$\langle A'\rangle_{U\rho U^{-1}} \neq \langle A'\rangle_{U\rho_L U^{-1}}. \tag{5}$$

Since the lhs of (5) is *distinct* from the rhs, one speaks of the former as *interference*. In the described two-slit case the lhs of (5) gives fringes, whereas the rhs does not. Nevertheless, it is not always true that the lhs of (5) itself means interference. This is the case only with a suitable pair of A and ρ (cf (ii) in lemma 1). Let me give a counterexample.

Let us take another two-slit experiment in which the slits have polarizers that give opposite linear polarization to the light passing the slits [10]. The state ρ in the slits is then such that we have equality in (5) (though A' is the same), and there is no interference because $[A, \rho] = 0$. (The state $\rho = |\psi\rangle\langle\psi|$ is now in the composite spatial-polarization state space, and the spatial subsystem state—the reduced statistical operator—is a Lüders state.)

One should note that when interference is displayed, one has three ingredients: the state ρ, the observable A the two eigenvalues of which *play a cooperative role*, and the second observable A' the probabilities of eigenvalues of which are observed. Since in theory there can be many observables such as A', or B in (4), one likes to omit them. Then one speaks of *coherence* of the observable A in the state ρ. We make use of the same concepts in the general theory.

Definition 1. *The lhs of relation (4), in case inequality (4) is valid, is called interference. If an observable A and a state ρ stand in such a mutual relation that any of the four claims of lemma 1 is known to be valid, then one speaks of coherence.*

One should note that the concepts of interference and of coherence stand in a peculiar relation to each other: There is no coherence (between A and ρ) unless an observable B that exhibits interference can be, in principle, found; if the latter is the case, and only then, one may forget about B, and concentrate on the relation between A and ρ, i.e., on coherence. The kind of quantum coherence investigated in this paper can be more fully called 'eigenvalue coherence of an observable in relation to a state' in view of the cooperative role of some eigenvalues (or, more precisely, their quantum numbers, because the values of the eigenvalues play no role) as seen in (4).

Thus, any of the four (equivalent) claims in lemma 1 defines coherence. But for the investigation in this paper the important claim is (ii): coherence exists if and only if A and ρ do not commute. This remark is the cornerstone of the expounded approach to investigating coherence (as in the preceding studies [1, 4]).

2. How to obtain a quantum measure of coherence?

We start with the assumption that coherence of an observable A with respect to a state ρ is *essentially the same thing* as incompatibility of A and ρ: $[A, \rho] \neq 0$. The quantum measure will be called *coherence* or incompatibility *information*, and it will be denoted by $I_C(A, \rho)$ or shortly I_C (cf (10) below).

One wonders what the meaning of a larger value of I_C for coherence is. It is more of what? The only answer I can think of is in accordance with the above assumption: more of incompatibility of A and ρ.

The next question is: do we know what is a 'larger amount of incompatibility'?

The seminal review on entropy of Wehrl [11] (section III.C there) explains that each member of the Wigner–Yanase–Dyson family of skew information

$$I_p(\rho, A) \equiv -S_p(\rho, A) \equiv (1/2)\,\mathrm{tr}([\rho^p, A][\rho^{1-p}, A]), \qquad 0 < p < 1, \tag{6}$$

is a good measure of incompatibility of ρ and A. Namely, $I_p(\rho, A)$ is positive unless ρ and A commute, when it is zero. It is also convex as an information quantity should be.

Substituting the spectral form of A in (6), one obtains

$$I_p = (1/2)\,\mathrm{tr}\left(\sum_l \sum_{l'} a_l[\rho^p, P_l]a_{l'}[\rho^{1-p}, P_{l'}]\right).$$

One can see that I_p depends on the eigenvalues of A.

As is well known, A and ρ are compatible if and only if all eigenprojectors P_l of the former are compatible with the latter. The eigenvalues of A do not enter this relation. Hence, $I_p(\rho, A)$ given by (6) is not the kind of incompatibility measure that we are looking for. One wonders if there is any other kind.

To obtain an answer, we turn to a neighbouring quantity: the quantum amount of *uncertainty* of A in ρ. It is the entropy $S(A, \rho)$:

$$S(A, \rho) \equiv H(p_l), \tag{7a}$$

where $H(p_l)$ is the Shannon entropy

$$H(p_l) \equiv -\sum_l p_l \log p_l, \tag{7b}$$

and

$$\forall l: \quad p_l \equiv \mathrm{tr}(P_l\rho). \tag{7c}$$

It is known that whenever A and ρ are incompatible, and A is a complete observable, i.e., if all its eigenvalues are nondegenerate (we will write it as A^c), then always $S(A^c,\rho) > S(\rho)$. When A^c is compatible with ρ, the two quantities are equal. The interpretation that the larger the difference $S(A^c,\rho) - S(\rho)$, the more incompatible A^c and ρ are seems plausible. Hence, we require for complete observables A^c, that $I_C(A^c,\rho)$ should equal this quantity: $I_C(A^c,\rho) \equiv S(A^c,\rho) - S(\rho)$. Equivalently, one can require that the following peculiar decomposition of the entropy in the case of a complete observable should hold:

$$S(\rho) = S(A^c,\rho) - I_C(A^c,\rho). \tag{8}$$

On the other hand, if A is a discrete observable that is complete or incomplete but *compatible* with ρ, then the following decomposition parallels (8):

$$S(\rho) = S(A,\rho) + \sum_l p_l S(P_l\rho P_l/p_l) \tag{9}$$

(cf (7*a*), (7*b*) and (7*c*)). If $p_l = 0$, the corresponding term in the sum is by definition zero.

Decomposition (9) is obtained by application of *the mixing property of entropy* [11] (see sections II.F and II.B there). It applies to an *orthogonal state decomposition*, in this case to $\rho = \sum_l p_l(P_l\rho P_l/p_l)$, and it reads $S(\rho) = H(p_l) + \sum_l p_l S(P_l\rho P_l/p_l)$ (cf (7*b*)).

The coherence information I_C does not appear in (9). This is as it should be because it is zero due to the assumed compatibility of A and ρ.

In the case of a general discrete A, which is complete or incomplete, compatible with ρ or not, we must interpolate between (8) and (9). This can be done by observing that both decompositions can be rewritten in a unified way as

$$I_C(A,\rho) = S\left(\sum_l P_l\rho P_l\right) - S(\rho) \tag{10}$$

(valid for either $A = A^c$ or for $[A,\rho] = 0$). The sought interpolated formula should thus be the same relation (10), but valid this time for all discrete A. Thus, $I_C(A,\rho)$ is obtained by the presented argument.

Making use of the mixing property of entropy, we can rewrite (10) equivalently as the following general decomposition of entropy:

$$S(\rho) = S(A,\rho) + \sum_l p_l S(P_l\rho P_l/p_l) - I_C(A,\rho). \tag{11}$$

(Note that A is any discrete observable in (11).)

In order to derive a number of properties of coherence information, we make a deviation into relative entropy theory.

3. Useful relative-entropy relations

The *relative entropy* $S(\rho\|\sigma)$ of a state (density operator) ρ with respect to a state σ is by definition

$$S(\rho\|\sigma) \equiv \mathrm{tr}[\rho\log(\rho)] - \mathrm{tr}[\rho\log(\sigma)] \tag{12a}$$

$$\text{if} \quad \mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma); \tag{12b}$$

or else $S(\rho\|\sigma) = +\infty$ (see page 16 in [12]). By 'support', denoted by 'supp', is meant the subspace that is the topological closure of the range.

If σ is singular and condition (12b) is valid, then the orthocomplement of the support (i.e., the null space) of ρ, contains the null space of σ, and both operators reduce in supp(σ). Relation (12b) is valid in this subspace. Both density operators also reduce in the null space of σ. Here the log is not defined, but it comes after zero, and it is generally understood that zero times an undefined quantity is zero. We will refer to this as *the zero convention*.

The more familiar concept of (von Neumann) quantum entropy, $S(\rho) \equiv -\mathrm{tr}[\rho\log(\rho)]$, also requires the zero convention. If the state space is infinite dimensional, then, in a sense, entropy is almost always infinite (cf page 241 in [11]). In finite-dimensional spaces, entropy is always finite.

There is an *equality for entropy* that is much used, and we have utilized it, *the mixing property* concerning *orthogonal state decomposition* (cf page 242 in [11]):

$$\sigma = \sum_k w_k\sigma_k, \tag{13}$$

$\forall k: w_k \geqslant 0$; for $w_k > 0$, $\sigma_k > 0, \mathrm{tr}\,\sigma_k = 1$; $\forall k \neq k'$: $\sigma_k\sigma_{k'} = 0$; $\sum_k w_k = 1$. Then $S(\sigma) = H(w_k) + \sum_k w_k S(\sigma_k)$, $H(w_k) \equiv -\sum_k [w_k\log(w_k)]$ being the Shannon entropy of the probability distribution $\{w_k : \forall k\}$.

The first aim of this section is to derive an analogue of the mixing property of entropy. The second aim is to derive two corollaries that we shall need in this paper.

We will find it convenient to make use of an *extension* $\log^{\mathrm{e}}$ of the logarithmic function to the entire real axis: if $0 < x:\ \log^{\mathrm{e}}(x) \equiv \log(x)$, if $x \leqslant 0:\ \log^{\mathrm{e}}(x) \equiv 0$.

The following elementary property of the extended logarithm will be utilized.

Lemma 2. *If an orthogonal state decomposition (13) is given, then*

$$\log^{\mathrm{e}}(\sigma) = {\sum_k}'[\log(w_k)]Q_k + {\sum_k}'\log^{\mathrm{e}}(\sigma_k), \tag{14}$$

where Q_k is the projector onto the support of σ_k, and the prime on the sum means that the terms corresponding to $w_k = 0$ are omitted.

Proof. Spectral forms $\forall k$, $w_k > 0$: $\sigma_k = \sum_{l_k} s_{l_k}|l_k\rangle\langle l_k|$ (all s_{l_k} positive) give a spectral form $\sigma = \sum_k\sum_{l_k} w_k s_{l_k}|l_k\rangle\langle l_k|$ of σ on account of the orthogonality assumed in (13) and the zero convention. Since numerical functions define the corresponding operator functions via spectral forms, one further obtains

$$\log^{\mathrm{e}}(\sigma) \equiv \sum_k\sum_{l_k}\left[\log^{\mathrm{e}}\left(w_k s_{l_k}\right)\right]|l_k\rangle\langle l_k| = {\sum_k}'\sum_{l_k}\left[\log(w_k) + \log\left(s_{l_k}\right)\right]|l_k\rangle\langle l_k|$$
$$= {\sum_k}'[\log(w_k)]Q_k + {\sum_k}'\sum_{l_k}\left[\log\left(s_{l_k}\right)\right]|l_k\rangle\langle l_k|.$$

(In the last step $Q_k = \sum_{l_k}|l_k\rangle\langle l_k|$ for $w_k > 0$ was made use of.) The same is obtained from the rhs when the spectral forms of σ_k are substituted in it. □

Proposition 1. *Let condition (12b) be valid for the states ρ and σ, and let an orthogonal state decomposition (13) be given. Then one has*

$$S(\rho\|\sigma) = S\left(\sum_k Q_k\rho Q_k\right) - S(\rho) + H(p_k\|w_k) + \sum_k p_k S(Q_k\rho Q_k/p_k\|\sigma_k), \tag{15}$$

where, for $w_k > 0$, Q_k projects onto the support of σ_k, and $Q_k \equiv 0$ if $w_k = 0$, $p_k \equiv \mathrm{tr}(\rho Q_k)$, and

$$H(p_k\|w_k) \equiv \sum_k[p_k\log(p_k)] - \sum_k[p_k\log(w_k)] \tag{16}$$

is the classical discrete counterpart of the quantum relative entropy, valid because $(p_k > 0) \Rightarrow (w_k > 0)$.

One should note that the claimed validity of the classical analogue of (12*b*) is due to the definitions of p_k and Q_k. Besides, (13) implies that $\left(\sum_k Q_k\right)$ projects onto supp(σ). Further, as a consequence of (12*b*), $\left(\sum_k Q_k\right)\rho = \rho$. Hence, $\mathrm{tr}\left(\sum_k Q_k \rho Q_k\right) = \mathrm{tr}\left(\sum_k Q_k \rho\right) = 1$.

We call decomposition (15) *the mixing property of relative entropy*.

Proof of proposition 1**.** We define

$$\forall k,\ p_k > 0: \qquad \rho_k \equiv Q_k \rho Q_k / p_k. \tag{17}$$

First we prove that (12*b*) implies

$$\forall k,\ p_k > 0: \qquad \mathrm{supp}(\rho_k) \subseteq \mathrm{supp}(\sigma_k). \tag{18}$$

Let k, $p_k > 0$, be an arbitrary fixed value. We take a pure-state decomposition

$$\rho = \sum_n \lambda_n |\psi_n\rangle\langle\psi_n|, \tag{19a}$$

$\forall n:\ \lambda_n > 0$. Applying $Q_k, \ldots, Q_k$ to (19*a*), one obtains another pure-state decomposition

$$Q_k \rho Q_k = p_k \rho_k = \sum_n \lambda_n Q_k |\psi_n\rangle\langle\psi_n| Q_k \tag{19b}$$

(cf (17)). Let $Q_k|\psi_n\rangle$ be a nonzero vector appearing in (19*b*). Since (19*a*) implies that $|\psi_n\rangle \in \mathrm{supp}(\rho)$ (cf appendix 2(ii)), condition (12*b*) further implies $|\psi_n\rangle \in \mathrm{supp}(\sigma)$. Let us write down a pure-state decomposition

$$\sigma = \sum_m \lambda'_m |\phi_m\rangle\langle\phi_m| \tag{20}$$

with $|\phi_1\rangle \equiv |\psi_n\rangle$. (This can be done with $\lambda'_1 > 0$ cf [13].) Then, applying $Q_k, \ldots, Q_k$ to (20) and taking into account (13), we obtain the pure-state decomposition

$$Q_k \sigma Q_k = w_k \sigma_k = \sum_m \lambda'_m Q_k |\phi_m\rangle\langle\phi_m| Q_k.$$

(Note that $w_k > 0$ because $p_k > 0$ by assumption.) Thus, $Q_k|\psi_n\rangle = Q_k|\phi_1\rangle \in \mathrm{supp}(\sigma_k)$. This is valid for any nonzero vector appearing in (19*b*), and these span $\mathrm{supp}(\rho_k)$ (cf appendix 2(ii)). Therefore, (18) is valid.

On account of (12*b*), the standard logarithm can be replaced by the extended one in definition (12*a*) of relative entropy: $S(\rho\|\sigma) = -S(\rho) - \mathrm{tr}[\rho \log^{\mathrm{e}}(\sigma)]$. Substituting (13) on the rhs, and utilizing (14), the relative entropy $S(\rho\|\sigma)$ becomes

$$-S(\rho) - \mathrm{tr}\left\{\rho\left[\sum_k{}' [\log(w_k)] Q_k + \sum_k{}' [\log^{\mathrm{e}}(\sigma_k)]\right]\right\}$$
$$= -S(\rho) - \sum_k{}' [p_k \log(w_k)] - \sum_k{}' \mathrm{tr}[\rho \log^{\mathrm{e}}(\sigma_k)].$$

Adding and subtracting $H(p_k)$, replacing $\log^{\mathrm{e}}(\sigma_k)$ by $Q_k[\log^{\mathrm{e}}(\sigma_k)]Q_k$, and taking into account (16) and (17), one further obtains

$$S(\rho\|\sigma) = -S(\rho) + H(p_k) + H(p_k\|w_k) - \sum_k{}' p_k\, \mathrm{tr}[\rho_k \log^{\mathrm{e}}(\sigma_k)].$$

(The zero convention is valid for the last term because the density operator $Q_k \rho Q_k / p_k$ may not be defined. Note that replacing $\sum_k$ by $\sum'_k$ in (16) does not change the lhs because only $p_k = 0$ terms are omitted.)

Adding and subtracting the entropies $S(\rho_k)$ in the sum, one further has

$$S(\rho\|\sigma) = -S(\rho) + H(p_k) + H(p_k\|w_k) + {\sum_k}' p_k S(\rho_k) + {\sum_k}' p_k\{-S(\rho_k) - \mathrm{tr}[\rho_k \log^{\mathrm{e}}(\sigma_k)]\}.$$

Utilizing the mixing property of entropy, one can put $S\big(\sum_k p_k\rho_k\big)$ instead of $[H(p_k) + \sum_k' p_k S(\rho_k)]$. Owing to (18), we can replace $\log^{\mathrm{e}}$ by the standard logarithm and thus obtain the rhs (15). □

Remark. In a sense, (15) runs parallel to Donald's identity

$$S(\rho\|\sigma) = \sum_k p_k S(\rho_k\|\sigma) - H(p_k),$$

when an orthogonal decomposition $\rho = \sum_k p_k\rho_k$ of the first state ρ in relative entropy is given.

For a general decomposition $\rho = \sum_k p_k\rho_k$ of the first state Donald's identity reads

$$S(\rho\|\sigma) = \sum_k p_k S(\rho_k\|\sigma) - \sum_k p_k S(\rho_k\|\rho)$$

[14, 15] (relation (5) in the latter). The more special relation in the remark follows from this on account of the relation that generalizes the mixing property of entropy: if $\rho = \sum_k p_k\rho_k$ is any state decomposition, then

$$S(\rho) = \sum_k p_k S(\rho_k\|\rho) + \sum_k p_k S(\rho_k)$$

is valid (cf lemma 4 and remark 1 in [16]).

Now we turn to the derivation of some consequences of proposition 1.

Let ρ be a state and $A = \sum_i a_i P_i + \sum_j a_j P_j$ be a spectral form of a discrete observable (Hermitian operator) A, where the eigenvalues a_i and a_j are all distinct. The index i enumerates all the detectable eigenvalues, i.e., $\forall i\colon\ \mathrm{tr}(\rho P_i) > 0$, and $\mathrm{tr}\big[\rho\big(\sum_i P_i\big)\big] = 1$.

The simplest quantum measurement of A in ρ changes this state into the Lüders state

$$\rho_L(A) \equiv \sum_i P_i\rho P_i \tag{21}$$

(cf (3*a*) and (3*c*)). Such a measurement is often called 'ideal'.

Corollary 1. *The relative-entropic 'distance' from any quantum state to its Lüders state is the difference between the corresponding quantum entropies:*

$$S\left(\rho \,\middle\|\, \sum_i P_i\rho P_i\right) = S\left(\sum_i P_i\rho P_i\right) - S(\rho).$$

Proof. First we prove that

$$\mathrm{supp}(\rho) \subseteq \mathrm{supp}\left(\sum_i P_i\rho P_i\right). \tag{22}$$

To this purpose, we write down a decomposition (19*a*) of ρ into pure states. One has $\mathrm{supp}\big(\sum_i P_i\big) \supseteq \mathrm{supp}(\rho)$ (equivalent to the certainty of $\big(\sum_i P_i\big)$ in ρ, cf [4]), and the decomposition (19*a*) implies that each $|\psi_n\rangle$ belongs to $\mathrm{supp}(\rho)$ (cf appendix 2(ii)). Hence, $|\psi_n\rangle \in \mathrm{supp}\big(\sum_i P_i\big)$; equivalently, $|\psi_n\rangle = \big(\sum_i P_i\big)|\psi_n\rangle$. Therefore, one can write

$$\forall n\colon\quad |\psi_n\rangle = \sum_i (P_i|\psi_n\rangle). \tag{23a}$$

On the other hand, (19a) implies

$$\sum_i P_i \rho P_i = \sum_i \sum_n \lambda_n P_i |\psi_n\rangle\langle\psi_n| P_i. \tag{23b}$$

As seen from (23b), all vectors $(P_i|\psi_n\rangle)$ belong to $\mathrm{supp}\left(\sum_i P_i \rho P_i\right)$. Hence, so do all $|\psi_n\rangle$ (due to (23a)). Since ρ is the mixture (19a) of the $|\psi_n\rangle$, the latter span $\mathrm{supp}(\rho)$ (cf appendix 2(ii)). Thus, finally, (22) also follows.

In our case $\sigma \equiv \sum_i P_i \rho P_i$ in (15). We replace k by i. Next, we establish

$$\forall i: \quad Q_i \rho Q_i = P_i \rho P_i. \tag{24}$$

Since Q_i is, by definition, the support projector of $(P_i \rho P_i)$, and $P_i(P_i \rho P_i) = (P_i \rho P_i)$, one has $P_i Q_i = Q_i$ (see appendix 2(i)). One can write $P_i \rho P_i = Q_i (P_i \rho P_i) Q_i$, from which then (24) follows.

Realizing that $w_i \equiv \mathrm{tr}(Q_i \rho Q_i) = \mathrm{tr}(P_i \rho P_i) \equiv p_i$ due to (24), one obtains $H(p_i \| w_i) = 0$ and $\forall i$: $S(Q_i \rho Q_i / p_i \| P_i \rho P_i / w_i) = 0$ in (15) for the case at issue. This completes the proof. □

Now we turn to a peculiar further implication of corollary 1.

Let $B = \sum_k \sum_{l_k} b_{kl_k} P_{kl_k}$ be a spectral form of a discrete observable (Hermitian operator) B such that all eigenvalues b_{kl_k} are distinct. Besides, let B be more complete than A or, synonymously, a refinement of the latter. This, by definition, means that

$$\forall k: \quad P_k = \sum_{l_k} P_{kl_k} \tag{25}$$

is valid. Here k enumerates both the i and the j index values in the spectral form of A.

Let $\rho_L(A)$ and $\rho_L(B)$ be the Lüders states (21) of ρ with respect to A and B respectively.

Corollary 2. *The states ρ, $\rho_L(A)$ and $\rho_L(B)$ lie on a straight line with respect to the relative entropy, i.e., $S(\rho \| \rho_L(B)) = S(\rho \| \rho_L(A)) + S(\rho_L(A)) \| \rho_L(B))$, or explicitly*

$$S\left(\rho \middle\| \sum_i \sum_{l_i} \left(P_{il_i} \rho P_{il_i}\right)\right) = S\left(\rho \middle\| \sum_i (P_i \rho P_i)\right) + S\left(\sum_i (P_i \rho P_i) \middle\| \sum_i \sum_{l_i} \left(P_{il_i} \rho P_{il_i}\right)\right).$$

Note that all eigenvalues b_{kl_k} of B with indices others than il_i are undetectable in ρ.

Proof. Corollary 1 immediately implies

$$S(\rho \| \rho_L(B)) = [S(\rho_L(B)) - S(\rho_L(A))] + [S(\rho_L(A)) - S(\rho)],$$

and, as easily seen from (21), $\rho_L(B) = (\rho_L(A))_L(B)$ due to $P_{il_i} P_{i'} = \delta_{i,i'} P_{il_i}$ (cf (25)). □

4. Properties of coherence information

To begin with, we notice in (10) that I_C depends on ρ and A, actually only on the eigenprojectors of the latter.

As a consequence of (10), one can also write the definition of I_C in the form of a relative entropy

$$I_C = S\left(\rho \middle\| \sum_l P_l \rho P_l\right) \tag{26}$$

as follows from corollary 1.

It was proved long ago [17] that $S\big(\sum_l P_l\rho P_l\big) > S(\rho)$ if and only if A and ρ are incompatible, and the two entropies are equal otherwise. Thus, in the case of compatibility $[A,\rho]=0$, I_C is zero, otherwise it is positive. This is what we would intuitively expect.

It was proved in [4] (theorem 2 there) that

$$I_C = w_{\mathrm{inc}} I_C\left(\sum_l^{\mathrm{inc}} a_l P_l, \left(\sum_l^{\mathrm{inc}} P_l\right)\rho\left(\sum_l^{\mathrm{inc}} P_l\right)\Big/ w_{\mathrm{inc}}\right), \tag{27}$$

where 'inc' on the sum denotes summing only over all those values of l the corresponding P_l of which are incompatible with ρ, and $w_{\mathrm{inc}} \equiv \mathrm{tr}\left(\rho\sum_l^{\mathrm{inc}} P_l\right)$.

This corresponds to an intuitive expectation that the quantity I_C should depend only on those eigenprojectors P_l of A that do not commute with ρ, and not at all on those that do.

We obtain (27) as a special case of a much more general result below (cf the theorem and propositions 2 and 3).

We shall need another known concept. For the sake of precision and clarity, we define it.

Definition 2. *One says that a discrete observable $\bar{A}=\sum_m \bar{a}_m\bar{P}_m$ (spectral form in terms of distinct eigenvalues $\bar{a}_m$) is coarser than or a coarsening of $A=\sum_l a_l P_l$ if there is a partitioning Π in the set $\{l:\forall l\}$ of all index values of the latter*

$$\Pi: \qquad \{l:\forall l\} = \sum_m C_m,$$

such that

$$\forall m: \quad \bar{P}_m = \sum_{l\in C_m} P_l$$

(C_m are classes of values of the index l, and the sum is the union of the disjoint classes). One also says that A is finer than or a refinement of $\bar{A}$.

Theorem. *Let $\bar{A}$ be any coarsening of A (cf definition 2). Then*

$$I_C(A,\rho) = I_C(\bar{A},\rho) + \sum_m [p_m I_C(\bar{P}_m A, \bar{P}_m\rho\bar{P}_m/p_m)], \tag{28}$$

and $\forall m$: $p_m \equiv \mathrm{tr}(\rho\bar{P}_m)$. (If $p_m = 0$, then, by the zero convention, the corresponding I_C in (28) need not be defined. The product is by definition zero.)

Before we prove the theorem, we apply corollary 2 to our case.

Under the assumptions of the theorem, one has

$$S\left(\rho\,\middle\|\,\sum_l (P_l\rho P_l)\right) = S\left(\rho\,\middle\|\,\sum_m (\bar{P}_m\rho\bar{P}_m)\right) + S\left(\sum_m (\bar{P}_m\rho\bar{P}_m)\,\middle\|\,\sum_l (P_l\rho P_l)\right). \tag{29}$$

Proof of the Theorem**.** On account of (26), (29) takes the form

$$I_C(A,\rho) = I_C(\bar{A},\rho) + I_C\left(A, \sum_m (\bar{P}_m\rho\bar{P}_m)\right). \tag{30}$$

Utilizing (10) for the second term on the rhs, the latter becomes $S\big(\sum_l (P_l\rho P_l)\big) - S\big(\sum_m (\bar{P}_m\rho\bar{P}_m)\big)$. Making use of the mixing property of entropy in both these terms, and cancelling out $H(p_m)$ (cf (7b) *mutatis mutandis*), this difference, further, becomes $\sum_m p_m S\big(\big(\sum_{l\in C_m} P_l\rho P_l\big)\big/ p_m\big) - \sum_m p_m S(\bar{P}_m\rho\bar{P}_m/p_m))$. Its substitution in (30) with the help of (10) (and definition 2) then gives the claimed relation (28). (Naturally, one must be aware of the fact that $\bar{A}$ is a coarsening of A, hence $\forall m$: $[\bar{P}_m, A] = 0$, implying $A \equiv \sum_m \sum_{m'} \bar{P}_m A \bar{P}_{m'} = \sum_m \bar{P}_m A$.) □

If $\bar{A}$ is any coarsening of A, then the index values m of the former replace classes C_m of index values l of the latter. Hence, coherence in $\bar{A}$ — as a cooperative role of index values — must be poorer than in A. Therefore, one would intuitively expect that $I_C(\bar{A}, \rho)$ must not be larger than $I_C(A, \rho)$. The theorem confirms this, and tells more: it gives the expression by which $I_C(A, \rho)$ exceed $I_C(\bar{A}, \rho)$. One wonders what the intuitive meaning of this is.

Discussion of the theorem. Let us think of ρ as describing a laboratory ensemble, and let us imagine that an ideal measurement of $\bar{A}$ is performed on each quantum system in the ensemble. The ensemble ρ is then replaced by the mixture $\sum_m p_m(\bar{P}_m \rho \bar{P}_m / p_m)$ of subensembles $(\bar{P}_m \rho \bar{P}_m / p_m)$. One can think of the measurement of the more refined observable A as taking place in two steps: the first is the mentioned measurement of the coarser observable $\bar{A}$, and the second is a continuation of measurement of A in each subensemble $(\bar{P}_m \rho \bar{P}_m / p_m)$. Let us assume *additivity* of I_C in two-step measurement.

Further, let us bear in mind that, though I_C is meant to be a property of each individual member of the ensemble ρ, it is *statistical*, i.e., it is given in terms of the ensemble. Finally, in the second step we have an ensemble of subensembles (a superensemble). Since our system is anywhere in the entire ensemble $\sum_m p_m(\bar{P}_m \rho \bar{P}_m / p_m)$ of the second step, one must average over the superensemble with the statistical weights p_m of its subensemble members $(\bar{P}_m \rho \bar{P}_m / p_m)$.

If $m' \neq m$, then the part $\bar{P}_{m'} A$ of $A = \sum_{m''} \bar{P}_{m''} A$ is evidently undetectable in the subensemble ρ_m. Hence, only $\bar{P}_m A$ is relevant from the entire A, i.e., $I_C(A, \rho)$ reduces to $I_C(\bar{P}_m A, \rho_m)$ there.

In this way one can understand relation (28). What have we learnt from this? It is that I_C is additive and statistical. This conclusion is in keeping with the neighbouring quantity $S(A, \rho)$. Namely, one can easily derive a relation similar to (28) for it:

$$S(A, \rho) = S(\bar{A}, \rho) + \sum_m p_m S(\bar{P}_m A, \bar{P}_m \rho \bar{P}_m / p_m).$$

That I_C and $S(A, \rho)$ behave equally in an additive and statistical way is no surprise since they are terms in the same general decomposition (11) of the entropy $S(\rho)$ of the state ρ.

The theorem is a substantially stronger form of a previous result (theorem 3 in [4]), in which $I_C(A, \rho) \geqslant I_C(\bar{A}, \rho)$ was established with necessary and sufficient conditions for equality, which are obvious in the theorem. (I_C was denoted by E_C in previous work, cf my comment following proposition 5 below.)

The theorem has the following immediate consequences.

Proposition 2. *If the coarsening $\bar{A}$ defined in definition 2 is compatible with ρ, then (28) reduces to*

$$I_C(A, \rho) = \sum_m [p_m I_C(\bar{P}_m A, \bar{P}_m \rho \bar{P}_m / p_m)]. \tag{31}$$

Proposition 3. *Let us define a coarsening Π (cf definition 2) that partitions $\{l : \forall l\}$ into at most three classes: C_{inc} comprising all index values l for which a_l is detectable (i.e., of positive probability) and P_l is incompatible with ρ, C_{comp} consisting of all l for which a_l is detectable and P_l is compatible with ρ, and, finally, C_{und} which is made up of all l for which a_l is undetectable. The coarsening thus defined is compatible with ρ, and (31) reduces to (27).*

Proof. In the coarsening Π of proposition 3 the index m takes on three 'values': 'inc', 'comp', and 'und'. It is easily seen that the coarser observable $\bar{A}$ thus defined is compatible with ρ. Hence, (31) applies. Further, the second and third terms are zero. In this way, (27) ensues. □

Proposition 4. *Coherence information I_C is unitary invariant, i.e., $I_C(A, \rho) = I_C(UAU^\dagger, U\rho U^\dagger)$, where U is an arbitrary unitary operator.*

Proof. Relative entropy is known to be unitary invariant. On account of (26), so is I_C. □

This is as it should be because I_C should not depend on the basis in the state space: UAU^{-1} and $U\rho U^{-1}$ can be understood as A and ρ respectively viewed in another basis.

Proposition 5. *Coherence information I_C is convex.*

Proof. This is an immediate consequence of the known convexity of relative entropy (cf (26)) under joint mixing of the two states in it. □

On account of convexity we know that I_C is an *information entity*, and not an entropy one (or else it would be concave). In previous work [1, 4, 2] the same quantity (the rhs of (10)) was erroneously denoted by $E_C(A, \rho)$ and treated as an entropy quantity. But this does not imply that any of the applications of $E_C(A, \rho)$ was erroneous. All one has to do is to replace this symbol by $I_C(A, \rho)$ and keep in mind that one is dealing with an information quantity.

5. Conclusion

Perhaps it is of interest to comment upon the more standard uses of the term 'coherence' in the literature.

One encounters the basic use of the word 'coherence' in the properties of light waves. One distinguishes two types of coherence there: (i) temporal coherence, which is a measure of the correlation between the phases of a light wave at different points along the direction of propagation, and (ii) spatial coherence, which is a measure of the correlation between the phases of a light wave at different points transverse to the direction of propagation. (The fascinating phenomenon of holography requires a large measure of both temporal and spatial coherence of light.)

Quantum 'coherence' also refers to large numbers of particles that cooperate collectively in a single quantum state. The best known examples are superfluidity, superconductivity and laser light, all macroscopic phenomena. In the last example different parts of the laser beam are related to each other in phase, which can lead to interference effects. 'Coherence' is often related to different kinds of correlations, see, e.g., [18].

In all the mentioned examples 'coherence' refers to an *absolute* property of the quantum state of the system; in contrast with the use of the term in this paper, which expresses a *relative* property: relation between an observable and a state. As it was mentioned, the kind of quantum coherence studied in this paper can be more fully called 'eigenvalue coherence of an observable in relation to a state' in view of the cooperative role of the eigenvalues (or rather their quantum numbers, because the values of the eigenvalues play no role) as seen in (4).

In the literature one often finds the claim that quantum pure states are coherent. From the analytical point of view of this paper one can say that a pure state $|\psi\rangle$ is *not coherent* with respect to any observable for which $|\psi\rangle\langle\psi|$ is an eigenprojector. But it is coherent with respect to all other observables.

5.1. On generality of the results

A question may linger on to the end of this study: what if the observable is not a discrete one? Can one still speak of eigenvalue coherence in relation to a given state ρ?

It seems to me that the answer is that one should write down the following partial spectral form of a general observable A':

$$A' = \sum_l a_l P_l + P^\perp A' P^\perp,$$

where the summation goes over *all eigenvalues* of A', and $P \equiv \sum_l P_l$. One should take the discrete coarsening A of A':

$$A \equiv \sum_l a_l P_l + a P^\perp,$$

where the eigenvalue a is arbitrary but distinct from all $\{a_l : \forall l\}$. Then the expounded eigenvalue coherence theory should by applied to A, and it should be valid for A' (as the best we can do for the latter). In a preceding paper [4] the case when $P^\perp \neq 0$ with the eigenvalue a undetectable was studied.

One has eigenvalue coherence of a general observable A' in relation to a state ρ if either A' has at least two eigenvalues or if A' has at least one eigenvalue and $P^\perp \neq 0$.

Another question that may linger on is whether the state ρ that was used in this paper is really general. If ρ has an infinite-dimensional range and A has infinitely many eigenvalues, it may happen that there are infinitely many detectable ones. The expounded theory covers also this case.

5.2. Summing up

In an attempt to understand the essential features of two-slit interference (see lemma 1 followed by its application to two-slit interference in subsection 1.2) a general coherence theory was developed based on the assumption that 'coherence' equals 'incompatibility' $[A, \rho] \neq 0$ between an observable and a state. Since this relation means that ρ is incompatible with at least one eigenevent (eigenprojector) P_l of A, and this property is independent of the eigenvalues, it was argued that the entire family of observables with one and the same decomposition of the identity $\sum_l P_l = I$ (the latter is called 'closure relation' if A is complete) should have the same amount of incompatibility. This discarded the Wigner–Yanase–Dyson family of skew information (6). Further, it was argued that the necessarily nonnegative quantity $S(A^c, \rho) - S(\rho)$ was a natural measure of incompatibility between a complete observable A^c and the state ρ satisfying the stated claim. Finally, interpolating between the case of a complete and that of a compatible observable (see (8)–(10)), the general expression (10) was obtained.

Thus, a natural quantum measure of how much of coherence, and, equivalently, incompatibility, there is if a discrete observable $A = \sum_l a_l P_l$ and a state ρ are given was derived along the expounded argument. It was called coherence or incompatibility information (denoted by $I_C(A, \rho)$ or shortly I_C) in section 2.

A deviation into a general relative-entropy investigation was made in section 3. What was called 'the mixing property of relative entropy' (paralleling that of entropy) was derived, and so were two corollaries.

The relative-entropy results were utilized to express the coherence information $I_C(A, \rho)$ in the form of a relative entropy (cf (26)) in section 4. Connection between the coherence information $I_C(\bar{A}, \rho)$ of any coarsening $\bar{A}$ (cf definition 2) of an observable A and $I_C(A, \rho)$

was obtained in the theorem. Its intuitive meaning was discussed. It was concluded that I_C is additive in two-step measurement and statistical.

The corresponding relation took a much simpler form in the case $\bar{A}$ was compatible with ρ (cf proposition 2). In a special case of this a result from previous work was recognized (cf proposition 3 and (27)). The coherence information was shown to be unitary invariant (proposition 4) and convex (proposition 5).

In previous work [1, 2, 4] the coherence information I_C was successfully utilized in analysing bipartite quantum correlations. The last one of them filled in an information–theoretical gap noted in a preceding investigation of the measurement process [3].

Since a number of new properties of I_C have now been obtained, even more fruitful applications can be expected.

Appendix 1.

We prove the equivalence of the negations of the four claims in lemma 1. ('$\neg$ (i)' is the negation of (i) etc, and '($\Leftrightarrow$)' is the claim of '$\Leftrightarrow$') The logical scheme of the proof is: $\neg$ (ii) $\Leftrightarrow \neg$ (iii) $\Leftrightarrow \neg$ (iv); $\neg$ (ii) $\Rightarrow \neg$ (i) $\Rightarrow \neg$ (iii).

$\neg$ (ii) ($\Leftrightarrow$)$\neg$ (iii): one can always write $\rho = \sum_l \sum_{l'} P_l \rho P_{l'}$. Since A and ρ commute if and only if each eigenprojector P_l of A commutes with ρ, the claimed equivalence is obvious.

($\neg$ (iii) $\Rightarrow \neg$ (iv)) is obvious. To prove ($\neg$ (iv) $\Rightarrow \neg$ (iii)), we restrict the operators B to ray projectors $|a\rangle\langle a|$. Then $\neg$ (iv) implies $\mathrm{tr}(\rho|a\rangle\langle a|) = \langle a|\rho|a\rangle = \langle a|\rho_L|a\rangle$ for every state vector $|a\rangle$. But then, as is well known, one must have $\rho = \rho_L$, which is $\neg$ (iii).

$\neg$ (ii) ($\Rightarrow$)$\neg$ (i): in view of $\rho = \sum_l \sum_{l'} P_l \rho P_{l'}$, commutation of ρ with each P_l implies $\neg$ (i).

$\neg$ (i) ($\Rightarrow$)$\neg$ (iii): let us assume that $\rho = \sum_l p_l \rho_l$, and that each state ρ_l has the sharp value of the corresponding eigenvalue a_l of A. Then $\rho_l = P_l \rho_l P_l$ (cf lemma A.4. in [19]). Substituting this in the state decomposition, and subsequently evaluating ρ_L according to (3a)–(3c), one can see that $\neg$ (iii) follows.

Appendix 2.

Let $\rho = \sum_n \lambda_n |n\rangle\langle n|$ be an arbitrary decomposition of a density operator into ray projectors, and let E be any projector. Then

$$E\rho = \rho \quad \Leftrightarrow \quad \forall n\colon E|n\rangle = |n\rangle \tag{A.1}$$

(cf lemmas A.1. and A.2. in [20]).

(i) If the above decomposition is an eigendecomposition with positive weights, then $\sum_n |n\rangle\langle n| = Q$, Q being now the support projector of ρ, and, on account of (A.1),

$$E\rho = \rho \quad \Rightarrow \quad EQ = Q. \tag{A.2}$$

(ii) Since one can always write $Q\rho = \rho$, (A.1) implies that all $|n\rangle$ in the arbitrary decomposition belong to $\mathrm{supp}(\rho)$. Further, defining a projector F so that $\mathrm{supp}(F) \equiv \mathrm{span}(\{|n\rangle : \forall n\})$, one has $FQ = F$. Equivalence (A.1) implies $F\rho = \rho$. Hence, (A.2) gives $QF = Q$. Altogether, $F = Q$, i.e., the unit vectors $\{|n\rangle : \forall n\}$ span $\mathrm{supp}(\rho)$.

References

[1] Herbut F 2002 *Phys. Rev.* A **66** 052321 (*Preprint* quant-ph/0305187)

[2] Herbut F 2003 On the meaning of entanglement in quantum measurement *Preprint* quant-ph/0311192

[3] Vedral V 2003 *Phys. Rev. Lett.* **90** 050401 (*Preprint* quant-ph/0207116)
[4] Herbut F 2003 *J. Phys. A: Math. Gen.* **36** 8479 (*Preprint* quant-ph/0309181)
[5] Henderson L and Vedral V 2001 *J. Phys. A: Math. Gen.* **34** 6899 (*Preprint* quant-ph/0105028)
[6] Zurek W H 2000 *Ann. Phys. (Lpz)* **9** 853 (*Preprint* quant-ph/0011039)
Ollivier H and Zurek W H 2002 *Phys. Rev. Lett.* **88** 17901 (*Preprint* quant-ph/0105072)
[7] Lüders G 1951 *Ann. Phys. (Leipzig)* **8** 322 (in German)
Messiah A 1961 *Quantum Mechanics* vol 1 (Amsterdam: North-Holland) p 333
Cohen-Tannoudji C, Diu B and Laloe F 1977 *Quantum Mechanics* vol 1 (New York: Wiley) p 221 (fifth postulate)
[8] Herbut F 1969 *Ann. Phys. (New York)* **55** 271 section 2
[9] Young T 1802 *Trans. R. Soc.* **12** 387
Herbut F 1992 *Am. J. Phys.* **60** 146
[10] Henneberger W C and Zitter R N 1983 *Am. J. Phys.* **51** 464
Holladay W 1993 *Phys. Lett.* A **183** 280
[11] Wehrl A 1978 *Rev. Mod. Phys.* **50** 221
[12] Ohya M and Petz D 1993 *Quantum Entropy and Its Use* (Berlin: Springer)
[13] Hadjisavvas N 1981 *Lett. Math. Phys.* **5** 327
[14] Donald M J 1987 *Math. Proc. Camb. Phil. Soc.* **101** 363
[15] Schumacher B and Westmoreland M D 2000 *Relative Entropy in Quantum Information Theory: Proc. AMS spec. session on Quantum Information and Computation* (*Preprint* quant-ph/0004045)
[16] Herbut F 2004 *J. Phys. A: Math. Gen.* **37** 3535 (*Preprint* quant-ph/0311193)
[17] Lindblad G 1972 *Commun. Math. Phys.* **28** 245
[18] Jaeger G and Sarkar S 2003 *Revisiting the Foundations of Relativistic Physics* ed A Ashtekar *et al* (Dordrecht: Kluwer) pp 523–42
[19] Herbut F 1996 *Found. Phys. Lett.* **9** 437
[20] Herbut F 1994 *J. Phys. A: Math. Gen.* **27** 7503