ELSEVIER

Optics Communications 6389 (2000) xxx

OPTICS COMMUNICATIONS

www.elsevier.com/locate/optcom

Quantitative quantum erasure [1]

Berthold-Georg Englert [a,b,*], János A. Bergou [c]

[a] Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, 85748 Garching, Germany
[b] Institute for Quantum Studies and Department of Physics, Texas A & M University, College Station, TX 77843-4242, USA
[c] Department of Physics and Astronomy, Hunter College of the City University of New York, 695 Park Avenue, New York, NY 10021, USA

Received 3 August 1999; received in revised form 21 October 1999; accepted 18 November 1999

Abstract

Quantum erasure aims at large fringe visibility in sorted subensembles. Since the achieved visibility depends on the specific sorting scheme, we optimize the sorting procedure under general circumstances. The possible compromises between fringe visibility (wave-typical) and which-alternative knowledge (particle-typical) are governed by a number of inequalities that are quantitative statements about wave-particle duality. Among them is the 'erasure relation', which states an absolute upper bound on the visibility that a quantum eraser can yield. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Wave-particle duality; Quantum erasure

1. Introduction

The quantum eraser was a sleeping beauty [2] until the early 1980's when Marlan Scully [4–6] noticed its significance for our understanding of measurements on quantum systems and, in particular, of the central role that entanglement plays in them. Since then, quantum erasure was a repeating theme in Marlan's published work and a powerful *deus ex machina* in numerous discussions. Not surprisingly, his recent textbook with Zubairy [7] contains a section on the subject. It is, therefore, an expression of our gratitude, and a great pleasure indeed, to dedicate these remarks on quantitative aspects of quantum erasure to Marlan Scully, to whom we owe our own interest in the matter.

After introducing the notational conventions used throughout (Sections 2.1 and 2.2), we discuss in Section 2.3 how labeled subensembles are identified – the decisive initial stage of quantum erasure. Two particular ways of forming subensembles, the which-alternative sorting and the quantum-erasure sorting, are then dealt with in Section 2.4. This enables us to report the various inequalities that make quantitative statements about wave-particle duality and quantum erasure (Section 2.5). Before prov-

* Corresponding author. E-mail: bge@physik.uni-ulm.de

[1] Dedicated to Marlan Scully on the occasion of his 60th birthday.

[2] The earliest mentioning of the essence of quantum erasure that we are aware of is an incidental remark by Groenewold in 1957 [1], concerning Einstein's recoiling-slit version of Young's double-slit interferometer (see Bohr's account in Ref. [2]). It is amusing that Groenewold's insight became an issue in a research paper some 35 years later [3].

0030-4018/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S0030-4018(99)00718-X

ing these inequalities in Section 2.7, we illustrate the richness of the field with the aid of the five simple examples of Section 2.6. Credit for contributions to the discovery of the inequalities of Section 2.5 is given (fairly, we hope) in Section 2.8, where the history of the subject is reviewed. Section 2 closes with a discussion of some mathematical aspects (Section 2.9). The general, and rather abstract, considerations of Section 2 are supplemented by two realistic examples in Section 3 – both with close relation to Marlan's work.

2. Interference of two quantum alternatives

2.1. Q-bits by themselves

Succumbing to the fashion of the day we'll call a two-state quantum degree of freedom a *q-bit*, and denote the kets that symbolize its basic alternatives by $|\mathrm{Q}_1\rangle$ and $|\mathrm{Q}_2\rangle$ (with $\langle\mathrm{Q}_j|\mathrm{Q}_k\rangle = \delta_{jk}$). Q-bits can have many different physical forms. Familiar examples include the binary alternatives of a Stern–Gerlach experiment ('spin up' or 'spin down'); of a photon's helicity ('left handed' or 'right handed'); of Young's double-slit set-up ('through this slit' or 'through that slit'); of Mach–Zehnder interferometers ('reflected at the entry beam splitter' or 'transmitted at it'); and of Ramsey interferometers ('transition in the first zone' or 'in the second zone'). The first and the last are discussed in some detail in Section 3, but in the present context, which concerns the physical concepts and the general formalism, we need not be specific about the actual physical nature of the q-bits because all two-alternative degrees of freedom are kinematically equivalent.

Before proceeding we wish to mention, however, that one and the same q-bit may have different 'basic alternatives' depending on the physical context. For a spin-$\frac{1}{2}$ atom, for example, 'up' and 'down' could refer to the vertical direction or a horizontal one.

It is convenient to introduce the ladder operators τ and $\tau^\dagger$,

$$\tau = |\mathrm{Q}_2\rangle\langle\mathrm{Q}_1|, \quad \tau^\dagger = |\mathrm{Q}_1\rangle\langle\mathrm{Q}_2|, \tag{1}$$

which are obvious analogs of Pauli's $(\sigma_x \pm \mathrm{i}\sigma_y)/2$ operators for spin-$\frac{1}{2}$, and the identities

$$\tau^\dagger\tau = |\mathrm{Q}_1\rangle\langle\mathrm{Q}_1|,$$
$$\tau\tau^\dagger = |\mathrm{Q}_2\rangle\langle\mathrm{Q}_2|, \quad \tau^\dagger\tau + \tau\tau^\dagger = 1_\mathrm{Q} \tag{2}$$

are immediate consequences of the definition (1). Then

$$\begin{aligned}\rho_\mathrm{Q} &= |\mathrm{Q}_1\rangle w_1 \langle\mathrm{Q}_1| + |\mathrm{Q}_2\rangle w_2 \langle\mathrm{Q}_2| + |\mathrm{Q}_2\rangle \\ &\quad\times \sqrt{w_2 w_1}\,\epsilon\langle\mathrm{Q}_1| + |\mathrm{Q}_1\rangle\sqrt{w_1 w_2}\,\epsilon^*\langle\mathrm{Q}_2| \\ &= w_1\tau^\dagger\tau + w_2\tau\tau^\dagger + \sqrt{w_1 w_2}\,(\epsilon\tau + \epsilon^*\tau^\dagger)\end{aligned} \tag{3}$$

are equivalent expressions for the statistical operator of a q-bit – its *state* – and

$$\rho_\mathrm{Q} \mathrel{\widehat{=}} \begin{pmatrix} w_1 & \epsilon^*\sqrt{w_1 w_2} \\ \epsilon\sqrt{w_1 w_2} & w_2 \end{pmatrix} \tag{4}$$

is a helpful 2×2-matrix representation. The positivity of ρ_Q and its normalization to unit trace,

$$\rho_\mathrm{Q} \ge 0, \quad \mathrm{tr}_\mathrm{Q}\{\rho_\mathrm{Q}\} = 1, \tag{5}$$

impose the restrictions

$$0 \le w_{1,2} \le 1, \quad w_1 + w_2 = 1, \quad |\epsilon| \le 1 \tag{6}$$

on the parameters w_1, w_2, and ϵ [3]. The probabilities for finding the alternatives Q_1 and Q_2 are given by their respective weights w_1 and w_2, and in view of

$$\rho_\mathrm{Q}(1_\mathrm{Q} - \rho_\mathrm{Q}) = w_1 w_2\left(1 - |\epsilon|^2\right)1_\mathrm{Q} \ge 0 \tag{7}$$

the complex number ϵ is a measure for the purity of the state ρ_Q – with $|\epsilon| = 1$ indicating that ρ_Q is a projector, i.e., a pure state. More significantly, however, ϵ specifies the degree of coherence between the alternatives and their phase relation.

The latter statement is justified by the following consideration. Coherence manifests itself in the capability of interfering, and the relevant interference

[3] In fact, ϵ is not restricted at all if $w_1 w_2 = 0$. We shall not entertain the reader with more mathematical pedantry of this kind.

pattern is given by the probability $p(\phi)$ of finding the superposition

$$\frac{1}{\sqrt{2}}\left(|\mathrm{Q}_1\rangle + |\mathrm{Q}_2\rangle \mathrm{e}^{\mathrm{i}\phi}\right) \tag{8}$$

or, equivalently, by the expectation value of the projector $I_{\mathrm{Q}}(\phi)$ to the ket (8),

$$I_{\mathrm{Q}}(\phi) = \tfrac{1}{2}\left(1_{\mathrm{Q}} + \mathrm{e}^{\mathrm{i}\phi}\tau + \mathrm{e}^{-\mathrm{i}\phi}\tau^{\dagger}\right) \mathrel{\widehat{=}} \frac{1}{2}\begin{pmatrix} 1 & \mathrm{e}^{-\mathrm{i}\phi} \\ \mathrm{e}^{\mathrm{i}\phi} & 1 \end{pmatrix}, \tag{9}$$

so that

$$p(\phi) = \mathrm{tr}_{\mathrm{Q}}\left\{I_{\mathrm{Q}}(\phi)\,\rho_{\mathrm{Q}}\right\} = \tfrac{1}{2} + \sqrt{w_1 w_2}\ \mathrm{Re}(\epsilon \mathrm{e}^{-\mathrm{i}\phi})\,. \tag{10}$$

Its ϕ dependence exhibits interference fringes, with the locations of the crests and troughs determined by the argument of ϵ, and since the extreme values of $p(\phi)$ are $\frac{1}{2} \pm \sqrt{w_1 w_2}\,|\epsilon|$, the visibility of these fringes involves the modulus of ϵ,

$$\mathscr{V} = 2\sqrt{w_1 w_2}\ |\epsilon|\,. \tag{11}$$

These observations demonstrate the significance of ϵ stated above. We note that unit visibility requires a state that is both unbiased ($w_1 = w_2 = \frac{1}{2}$) and pure ($|\epsilon| = 1$).

2.2. Q-bits entangled with their environments

Q-bits get entangled with other degrees of freedom as they interact with them. Part of this interaction may be in accordance with the experimenter's intentions, part may consist of unwanted disturbances. In any case, the joint state of a q-bit and its environment will be of the form

$$\begin{aligned}\rho_{\mathrm{Q\&E}} &= w_1\tau^{\dagger}\tau\rho_{\mathrm{E}}^{(1)} + w_2\tau\tau^{\dagger}\rho_{\mathrm{E}}^{(2)} + \sqrt{w_1 w_2}\left(\tau\chi_{\mathrm{E}} + \tau^{\dagger}\chi_{\mathrm{E}}^{\dagger}\right) \\ &\mathrel{\widehat{=}} \begin{pmatrix} w_1\,\rho_{\mathrm{E}}^{(1)} & \sqrt{w_1 w_2}\ \chi_{\mathrm{E}}^{\dagger} \\ \sqrt{w_1 w_2}\ \chi_{\mathrm{E}} & w_2\,\rho_{\mathrm{E}}^{(2)} \end{pmatrix},\end{aligned} \tag{12}$$

where $\rho_{\mathrm{E}}^{(1)}$ and $\rho_{\mathrm{E}}^{(2)}$ are states of the environment and χ_{E} is the environment's 'cross term'. Included in the environment are only those degrees of freedom about which the experimenter has some knowledge and some control, and not everything composing the rest of the universe that surrounds the q-bit in question.

Upon tracing over the environment we get the (reduced) state of the q-bit, which is of the form (3) with ϵ given by

$$\epsilon = \mathrm{tr}_{\mathrm{E}}\{\chi_{\mathrm{E}}\}\,. \tag{13}$$

Likewise, the state of the environment is

$$\rho_{\mathrm{E}} = \mathrm{tr}_{\mathrm{Q}}\{\rho_{\mathrm{Q\&E}}\} = w_1\,\rho_{\mathrm{E}}^{(1)} + w_2\,\rho_{\mathrm{E}}^{(2)}\,; \tag{14}$$

it is a convex sum of $\rho_{\mathrm{E}}^{(1)}$ and $\rho_{\mathrm{E}}^{(2)}$ – a *mixture* blended from these ingredients with weights w_1 and w_2. The q-bit and its environment are genuinely entangled if $\rho_{\mathrm{Q\&E}}$ of (12) is different from the (tensor) product of ρ_{Q} and ρ_{E},

$$\begin{aligned}\rho_{\mathrm{Q\&E}} &= \rho_{\mathrm{Q}}\,\rho_{\mathrm{E}} : \text{no entanglement}\,,\\ \rho_{\mathrm{Q\&E}} &\neq \rho_{\mathrm{Q}}\,\rho_{\mathrm{E}} : \text{yes entanglement}\,,\end{aligned} \tag{15}$$

so that entanglement is there unless

$$\rho_{\mathrm{E}} = \rho_{\mathrm{E}}^{(1)} = \rho_{\mathrm{E}}^{(2)} \quad \text{and} \quad \chi_{\mathrm{E}} = \rho_{\mathrm{E}}\mathrm{tr}_{\mathrm{E}}\{\chi_{\mathrm{E}}\} \tag{16}$$

hold.

The correct normalization of $\rho_{\mathrm{Q\&E}}$ is ensured by construction, but its positivity imposes some restrictions on χ_{E} for given $\rho_{\mathrm{E}}^{(1)}$ and $\rho_{\mathrm{E}}^{(2)}$. Since $\rho_{\mathrm{Q\&E}} \geq 0$ just means that all probabilities derived from $\rho_{\mathrm{Q\&E}}$ are non-negative, one must simply require that

$$\mathrm{tr}_{\mathrm{E}}\{P_{\mathrm{Q\&E}}\ \rho_{\mathrm{Q\&E}}\} \geq 0 \tag{17}$$

is obeyed for any $P_{\mathrm{Q\&E}}$ that projects to a pure state. This condition is turned into a more explicit statement after noting that all such projectors are of the form

$$P_{\mathrm{Q\&E}} \mathrel{\widehat{=}} \begin{pmatrix} |\mathrm{E}_a\rangle w_a \langle \mathrm{E}_a| & |\mathrm{E}_a\rangle \mathrm{e}^{-\mathrm{i}\varphi}\sqrt{w_a w_b}\,\langle \mathrm{E}_b| \\ |\mathrm{E}_b\rangle \mathrm{e}^{\mathrm{i}\varphi}\sqrt{w_a w_b}\,\langle \mathrm{E}_a| & |\mathrm{E}_b\rangle w_b \langle \mathrm{E}_b| \end{pmatrix}, \tag{18}$$

where $|\mathrm{E}_a\rangle$ and $|\mathrm{E}_b\rangle$ are two arbitrary, normalized environment kets that need not be orthogonal to each other. The weights w_a, w_b have the usual properties [cf. Eq. (6)], and one easily verifies that $P_{\mathrm{Q\&E}}^2 = P_{\mathrm{Q\&E}}$ and $\mathrm{tr}_{\mathrm{Q\&E}}\{P_{\mathrm{Q\&E}}\} = 1$. Now, (17) has to be true for all values of the phase parameter φ as well as all allowed choices for w_a and w_b in (18), which is equivalent to requiring that

$$\langle \mathrm{E}_b|\chi_{\mathrm{E}}|\mathrm{E}_a\rangle\langle \mathrm{E}_a|\chi_{\mathrm{E}}^\dagger|\mathrm{E}_b\rangle \le \langle \mathrm{E}_a|\rho_{\mathrm{E}}^{(1)}|\mathrm{E}_a\rangle\langle \mathrm{E}_b|\rho_{\mathrm{E}}^{(2)}|\mathrm{E}_b\rangle \tag{19}$$

holds for any pair of environment kets $|\mathrm{E}_a\rangle$ and $|\mathrm{E}_b\rangle$. Summation over complete sets of $|\mathrm{E}_b\rangle$'s or $|\mathrm{E}_a\rangle$'s implies the useful operator statements

$$\chi_{\mathrm{E}}^\dagger\,\chi_{\mathrm{E}} \le \rho_{\mathrm{E}}^{(1)} \quad\text{and}\quad \chi_{\mathrm{E}}\,\chi_{\mathrm{E}}^\dagger \le \rho_{\mathrm{E}}^{(2)}\,. \tag{20}$$

We note also that $|\epsilon| \le 1$ obtains as a consequence of (19) and (13), as it should.

2.3. Labeled subensembles

Intentional entanglement of the q-bit with its environment enables the experimenter to sort q-bits into labeled subensembles. The label is provided by the outcome of a measurement performed on the environment. We take for granted that this is a standard von Neumann measurement of an environment observable O_{E} of the experimenter's choosing and that, therefore, the outcome is one of the eigenvalues of O_{E}. These eigenvalues can be degenerate or non-degenerate, so that the number of subensembles may vary from one sorting to the next. In particular, the limiting case $O_{\mathrm{E}} = 1_{\mathrm{E}}$ that corresponds to no sorting at all is included.

For the actual O_{E} under consideration, we denote the projectors to the subspaces associated with its eigenvalues by $P_{\mathrm{E}}^{(k)}$. They are, of course, mutually orthogonal and complete,

$$P_{\mathrm{E}}^{(j)}P_{\mathrm{E}}^{(k)} = \delta_{jk}P_{\mathrm{E}}^{(k)}\,, \quad \sum_k P_{\mathrm{E}}^{(k)} = 1_{\mathrm{E}}\,. \tag{21}$$

The trace of $P_{\mathrm{E}}^{(k)}$ is the multiplicity of the respective eigenvalue of O_{E}. In passing we note that one could consider measurements of a more general kind, where a complete set of positive operators would play the role of the projectors $P_{\mathrm{E}}^{(k)}$, but we wouldn't really gain anything from such a generalization.

The eigenvalue corresponding to $P_{\mathrm{E}}^{(k)}$ is found with the probability

$$\begin{aligned} p_k = \mathrm{tr}_{\mathrm{Q\&E}}\{P_{\mathrm{E}}^{(k)}\rho_{\mathrm{Q\&E}}\} &= w_1\mathrm{tr}_{\mathrm{E}}\{P_{\mathrm{E}}^{(k)}\rho_{\mathrm{E}}^{(1)}\} \\ &\quad + w_2\mathrm{tr}_{\mathrm{E}}\{P_{\mathrm{E}}^{(k)}\rho_{\mathrm{E}}^{(2)}\}\,, \end{aligned} \tag{22}$$

and the q-bit state $\rho_{\mathrm{Q}}^{(k)}$ that is conditioned on this outcome of the O_{E} measurement is given by

$$p_k\,\rho_{\mathrm{Q}}^{(k)} = \mathrm{tr}_{\mathrm{E}}\{P_{\mathrm{E}}^{(k)}\rho_{\mathrm{Q\&E}}\}\,. \tag{23}$$

Since all q-bit states are of the form (3), $\rho_{\mathrm{Q}}^{(k)}$ is characterized by positive weights w_{1k} and w_{2k} as well as by a complex ϵ_k, as specified by

$$\begin{aligned} p_k w_{1k} &= w_1\mathrm{tr}_{\mathrm{E}}\{P_{\mathrm{E}}^{(k)}\rho_{\mathrm{E}}^{(1)}\}\,, \\ p_k w_{2k} &= w_2\mathrm{tr}_{\mathrm{E}}\{P_{\mathrm{E}}^{(k)}\rho_{\mathrm{E}}^{(2)}\}\,, \\ p_k\sqrt{w_{1k}w_{2k}}\;\epsilon_k &= \sqrt{w_1w_2}\;\mathrm{tr}_{\mathrm{E}}\{P_{\mathrm{E}}^{(k)}\chi_{\mathrm{E}}\}\,. \end{aligned} \tag{24}$$

One verifies easily that $|\epsilon_k| \le 1$ is ensured by restriction (19). Fig. 1 illustrates these matters.

2.4. Which-alternative sorting and quantum erasure

Different choices for the environment observable O_{E}, or more to the point: of the projectors $P_{\mathrm{E}}^{(k)}$

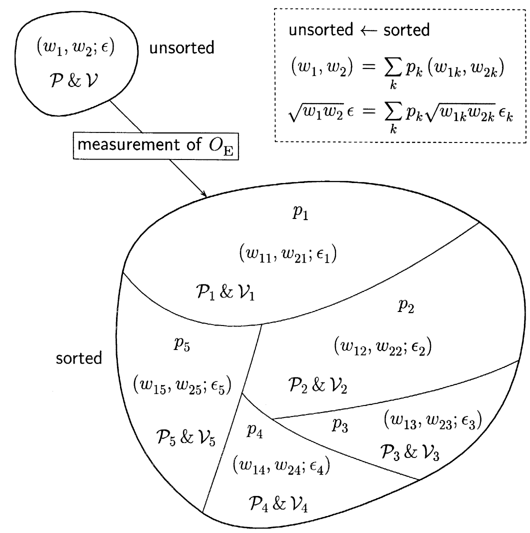

Fig. 1. Sorting q-bits into subensembles. According to the outcome of a measurement of the environment observable O_{E}, the ensemble of q-bits – characterized by the triple (w_1,w_2,ϵ) – is sorted into subensembles – characterized by $(w_{1k},w_{2k};\epsilon_k)$ with $k = 1,\dots,5$ here. The probability that a q-bit ends up in the k-th subensemble is p_k. The subensembles have predictabilities $\mathcal{P}_k = |w_{1k} - w_{2k}|$ and visibilities $\mathcal{V}_k = 2\sqrt{w_{1k}w_{2k}}\,|\epsilon_k|$. As a rule, they are different from the predictability $\mathcal{P} = |w_1 - w_2|$ and the visibility $\mathcal{V} = 2\sqrt{w_1w_2}\,|\epsilon|$ of the whole ensemble.

associated with it, result in different sortings of identically prepared and processed q-bits into subensembles. Two sorting schemes are of particular interest, namely the which-alternative (WA) sorting and the quantum-erasure (QE) sorting [4].

The WA sorting aims at getting as much information as Nature would grant us about the alternative that is actually the case. A quantitative measure is the *WA knowledge* $\mathscr{K}$,

$$\begin{aligned}\mathscr{K}(O_{\mathrm{E}}) &= \sum_k p_k |w_{1k} - w_{2k}| \\ &= \sum_k \left|\mathrm{tr}_{\mathrm{E}}\left\{P_{\mathrm{E}}^{(k)}\left(w_1\,\rho_{\mathrm{E}}^{(1)} - w_2\,\rho_{\mathrm{E}}^{(2)}\right)\right\}\right|,\end{aligned} \tag{25}$$

which is from the range $0 \le \mathscr{K} \le 1$ and has the following significance. If you make an educated guess about the actual alternative, based on knowing the outcome of the O_{E} measurement, then $\mathscr{K}(O_{\mathrm{E}})$ is the fractional excess of right guesses over wrong guesses in many experiments repeated under identical conditions. A value of $\mathscr{K} = 0.4$, say, means that the percentage of right guesses exceeds that of wrong ones by 40% – of all your guesses 70% will be right, 30% will be wrong.

The notation $\mathscr{K}(O_{\mathrm{E}})$ in (25) is a reminder that this WA knowledge depends on the choice of O_{E} (or rather the orthogonal subspaces associated with it). Now, a *WA sorting* is such that $\mathscr{K}(O_{\mathrm{E}})$ is as large as it can possibly be. The largest $\mathscr{K}$ value is the *distinguishability* $\mathscr{D}$ of the alternatives,

$$\mathscr{D} = \underset{O_{\mathrm{E}}}{\mathrm{Max}}\left\{\mathscr{K}(O_{\mathrm{E}})\right\} = \mathrm{tr}_{\mathrm{E}}\left\{\left|w_1\,\rho_{\mathrm{E}}^{(1)} - w_2\,\rho_{\mathrm{E}}^{(2)}\right|\right\}, \tag{26}$$

where the modulus $|X|$ of an operator X is given by $|X| = \sqrt{X^\dagger X}$ [5]. The value of $\mathscr{D}$ is thus obtained by summing the moduli of the eigenvalues of $w_1\,\rho_{\mathrm{E}}^{(1)} - w_2\,\rho_{\mathrm{E}}^{(2)}$.

Any choice for O_{E} for which $\mathscr{K}(O_{\mathrm{E}})$ equals $\mathscr{D}$ defines q-bit subensembles of a WA sorting. All WA sortings have this characterizing property in common:

If

$$\mathrm{tr}_{\mathrm{E}}\left\{P_{\mathrm{E}}^{(k)}\left(w_1\,\rho_{\mathrm{E}}^{(1)} - w_2\,\rho_{\mathrm{E}}^{(2)}\right)\right\} > 0,\ < 0,\ = 0,$$

then

$$P_{\mathrm{E}}^{(k)}\left(w_1\,\rho_{\mathrm{E}}^{(1)} - w_2\,\rho_{\mathrm{E}}^{(2)}\right)P_{\mathrm{E}}^{(k)} \ge 0,\ \le 0,\ = 0,$$

respectively. (27)

As a rule there is more than one WA sorting.

There is also a smallest value for the knowledge $\mathscr{K}(O_{\mathrm{E}})$, the so-called *predictability* $\mathscr{P}$ of the alternatives,

$$\mathscr{P} = \underset{O_{\mathrm{E}}}{\mathrm{Min}}\left\{\mathscr{K}(O_{\mathrm{E}})\right\} = |w_1 - w_2|. \tag{28}$$

Clearly, this minimum obtains for $O_{\mathrm{E}} = 1_{\mathrm{E}}$ and so $\mathscr{P}$ represents the WA knowledge that is available without any measurement on the environment. We note that

$$\mathscr{P} \le \mathscr{K}(O_{\mathrm{E}}) \le \mathscr{D} \tag{29}$$

is an obvious hierarchy: The distinguishability $\mathscr{D}$ represents Nature's information about the actual alternative; the knowledge $\mathscr{K}(O_{\mathrm{E}})$ is what Man can learn from a measurement of the environment observable O_{E}; and the predictability $\mathscr{P}$ is Man's knowledge before measuring O_{E}. It is worth emphasizing that the three numbers appearing in (29) quantify information or knowledge without invoking an entropic concept of some kind.

The QE sorting aims at finding subensembles with particularly large fringe visibility. Since $\mathscr{V}_k = 2\sqrt{w_{1k}w_{2k}}\,|\epsilon_k|$ is the visibility in the $\rho_{\mathrm{Q}}^{(k)}$ subensemble and p_k is the weight it carries, we get an averaged visibility $\mathscr{V}^{(\mathrm{QE})}$ for the sorting based on the O_{E} measurement that is given by

$$\begin{aligned}\mathscr{V}^{(\mathrm{QE})}(O_{\mathrm{E}}) &= \sum_k p_k\, 2\sqrt{w_{1k}w_{2k}}\,|\epsilon_k| \\ &= 2\sqrt{w_1 w_2}\sum_k \left|\mathrm{tr}_{\mathrm{E}}\left\{P_{\mathrm{E}}^{(k)}\chi_{\mathrm{E}}\right\}\right|.\end{aligned} \tag{30}$$

This p_k-weighted sum of subensemble visibilities $\mathscr{V}_k$ is, of course, quite analogous to $\mathscr{K}(O_{\mathrm{E}})$ of

---

[4] In the context of interferometers, where the alternatives are commonly called *ways*, the WA sorting is familiar as the which-way sorting, achieved by exploiting whatever serves as a which-way detection device, whose degrees of freedom make up the environment in the terminology of the present study.

[5] Operators X for which the trace of $|X|$ is finite are said to be in the *trace class*; alternatively, they are called *nuclear operators*. Their most important mathematical properties are discussed in textbooks that treat the theory of Hilbert spaces in some detail (see Refs. [8–11], for example); some basics are stated in Section 2.9. Note that Eqs. (20) state $|\chi_{\mathrm{E}}|^2 \le \rho_{\mathrm{E}}^{(1)}$ and $|\chi_{\mathrm{E}}^\dagger|^2 \le \rho_{\mathrm{E}}^{(2)}$.

(25), which is the weighted sum of subensemble predictabilities $\mathscr{P}_k = |w_{1k} - w_{2k}|$.

Optimization of $\mathscr{V}^{(\mathrm{QE})}(O_\mathrm{E})$ with respect to the choice of the environment observable O_E yields the *coherence* $\mathscr{C}$ [6],

$$\mathscr{C} = \sup_{O_\mathrm{E}} \left\{ \mathscr{V}^{(\mathrm{QE})}(O_\mathrm{E}) \right\} = 2\sqrt{w_1 w_2}\, \mathrm{tr}_\mathrm{E}\left\{ |\chi_\mathrm{E}| \right\}, \tag{31}$$

where one has to be content with a supremum and cannot count on having a maximum (see Section 2.9). The chosen name 'coherence' for the optimized QE visibility $\mathscr{V}^{(\mathrm{QE})}$ is suggested by the fact that $\mathscr{C}$ is a measure for the absolute size of the cross term χ_E and thus for the extent to which (12) contains coherent superpositions of the q-bit alternatives $|\mathrm{Q}_1\rangle$ and $|\mathrm{Q}_2\rangle$.

If $|\chi_\mathrm{E}|$ and $|\chi_\mathrm{E}^\dagger|$ have the same number of zero eigenvalues (perhaps infinitely many), and only then, they are unitarily equivalent and the supremum in (31) is a maximum. Then there is one or more environment observables O_E for which $\mathscr{V}^{(\mathrm{QE})}(O_\mathrm{E}) = \mathscr{C}$ [7]. The $P_\mathrm{E}^{(k)}$'s of such *QE sortings* can be characterized by a property that is analogous to (27) but somewhat more involved.

Since

$$\sum_k \left| \mathrm{tr}_\mathrm{E}\left\{ P_\mathrm{E}^{(k)} \chi_\mathrm{E} \right\} \right| \geq \left| \sum_k \mathrm{tr}_\mathrm{E}\left\{ P_\mathrm{E}^{(k)} \chi_\mathrm{E} \right\} \right| = \left| \mathrm{tr}_\mathrm{E}\{ \chi_\mathrm{E} \} \right|, \tag{32}$$

the QE visibility $\mathscr{V}^{(\mathrm{QE})}(O_\mathrm{E})$ of (30) is always at least as large as the visibility $\mathscr{V} = \mathscr{V}^{(\mathrm{QE})}(1_\mathrm{E})$ of the unsorted whole ensemble, given in Eq. (11) with ϵ from (13), and therefore we have also a hierarchy of visibilities,

$$\mathscr{V} \leq \mathscr{V}^{(\mathrm{QE})}(O_\mathrm{E}) \leq \mathscr{C}\,. \tag{33}$$

It supplements and complements the knowledge hierarchy of (29).

The environment observables O_E that must be measured for the optimizations of the WA knowledge $\mathscr{K}(O_\mathrm{E})$ and the QE visibility $\mathscr{V}^{(\mathrm{QE})}(O_\mathrm{E})$ are usually quite different, almost always incompatible, and sometimes complementary. But, somewhat surprisingly, it is also possible that the same O_E yields both a WA sorting and a QE sorting.

Clearly, interesting sortings are only possible if the q-bit is genuinely entangled with its environment, because the 'no' case of (16) implies $\mathscr{D} = \mathscr{P}$ and $\mathscr{C} = \mathscr{V}$. Accordingly, $(\mathscr{D}^2 + \mathscr{C}^2) - (\mathscr{P}^2 + \mathscr{V}^2) > 0$ is indicative of entanglement, and perhaps one can introduce a useful quantitative measure for the entanglement by studying the properties of this difference. As far as we know, this is unexplored territory, and we'll be content presently with the remark that $\mathscr{D}^2 + \mathscr{C}^2$ may depend on what one regards as the basic alternatives of the q-bit, whereas $\mathscr{P}^2 + \mathscr{V}^2$ does not.

2.5. Inequalities

Each of the hierarchies (29) and (33) refers to one diagonal of the 2×2 matrix (12). In addition, there are fundamental inequalities that involve the upper and lower bounds of both hierarchies. First, we recall the familiar *duality relation*

$$\mathscr{V}^2 + \mathscr{D}^2 \leq 1\,, \tag{34}$$

which is a quantitative statement about wave-particle duality because it puts a bound on the available WA knowledge, a particle-typical property, given that the unsorted whole ensemble exhibits wave-typical fringes with a certain visibility. Second, we put on record the new *erasure relation*

$$\mathscr{P}^2 + \mathscr{C}^2 \leq 1\,, \tag{35}$$

which is a quantitative statement about the effectiveness of QE because it puts a bound on the visibility that QE can recover, given that the unsorted whole ensemble possesses a certain predictability. These matters are summarized in Fig. 2.

We postpone the technical proofs of these inequalities until Section 2.7 and prefer to offer first some remarks here, then to consider simple illustrating examples in the following Section 2.6. A brief historical review of the subject follows in Sections 2.8 and 2.9 deals with more mathematical issues.

---

[6] Note that $\mathrm{tr}_\mathrm{E}\{|\chi_\mathrm{E}|\} = \mathrm{tr}_\mathrm{E}\{|\chi_\mathrm{E}^\dagger|\}$ although $|\chi_\mathrm{E}| = (\chi_\mathrm{E}^\dagger \chi_\mathrm{E})^{1/2}$ and $|\chi_\mathrm{E}^\dagger| = (\chi_\mathrm{E} \chi_\mathrm{E}^\dagger)^{1/2}$ are different, unless χ_E commutes with $\chi_\mathrm{E}^\dagger$.

[7] For example, if one simply exploits the unitary equivalence by choosing O_E such that $|\chi_\mathrm{E}^\dagger| \exp(\mathrm{i} O_\mathrm{E}) = \exp(\mathrm{i} O_\mathrm{E}) |\chi_\mathrm{E}|$, then $\mathscr{V}^{(\mathrm{QE})}(O_\mathrm{E}) = \mathscr{C}$; see Section 2.9 for more detail.

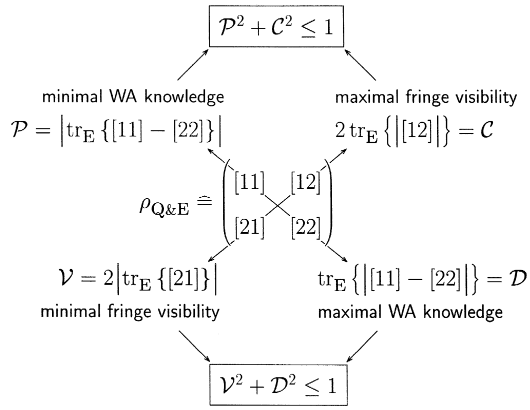

Fig. 2. Significance of the diagonals of the 2×2 matrix (12). From the main diagonal with $[11]\equiv w_1\,\rho_{\mathrm{E}}^{(1)}$ and $[22]\equiv w_2\,\rho_{\mathrm{E}}^{(2)}$ one can extract WA knowledge, whereas the skew diagonal with $[21]\equiv\sqrt{w_1w_2}\;\chi_{\mathrm{E}}$ and $[12]=[21]^\dagger$ determines the fringe visibility.

As a first remark we note that there is, of course, also an inequality for unsorted data, namely the statement

$$\mathscr{P}^{\,2}+\mathscr{V}^{\,2}\le 1\,, \tag{36}$$

which can be regarded as an immediate consequence of either the duality relation or the erasure relation since $\mathscr{P}\le\mathscr{D}$ and $\mathscr{V}\le\mathscr{C}$. Perhaps more important is the observation that one can never be sure if the optimizations required in (26) and (31) are realized by the sortings done in an actual experiment. Accordingly, experimental tests of (34) and (35) necessarily concern the inequalities

$$\mathscr{V}^{\,2}+\left[\mathscr{K}(O_{\mathrm{E}})\right]^2\le 1 \tag{37}$$

and

$$\mathscr{P}^{\,2}+\left[\mathscr{V}^{\,(\mathrm{QE})}(O_{\mathrm{E}})\right]^2\le 1\,, \tag{38}$$

respectively.

The most stringent test, however, is offered by

$$\left[\mathscr{K}(O_{\mathrm{E}})\right]^2+\left[\mathscr{V}^{\,(\mathrm{QE})}(O_{\mathrm{E}})\right]^2\le 1\,, \tag{39}$$

which relates to each other the WA knowledge and the QE visibility of one sorting, and holds for all sortings. Clearly, the inequalities (37) and (38) are immediate consequences of (39) and the optimizations of (26) and (31) then imply the duality relation (34) and the erasure relation (35). In this sense, (39) is the strongest of the various inequalities, and (36) is the weakest.

In another sense, however, inequality (36) is as strong as (39), since one proof of (39) is based on little more than the insight that (36) must be obeyed by each subensemble separately. This is so because it is the experimenter's decision which q-bits are considered members of the whole ensemble, and so any subset of the subensembles associated with the $P_{\mathrm{E}}^{(k)}$'s can serve as *the* ensemble to which (36) refers.

2.6. Simple illustrating examples

As a *first example* we consider a pure state, that is: a $\rho_{\mathrm{Q\&E}}$ of the form (18). The extreme values of the WA knowledge are

$$\mathscr{P}=\left|w_a-w_b\right|\le\left(1-4w_aw_b\left|\langle\mathrm{E}_a|\mathrm{E}_b\rangle\right|^2\right)^{1/2}=\mathscr{D}\,, \tag{40}$$

while those of the visibility are

$$\mathscr{V}=2\sqrt{w_aw_b}\;\left|\langle\mathrm{E}_a|\mathrm{E}_b\rangle\right|\le 2\sqrt{w_aw_b}\;=\mathscr{C}\,. \tag{41}$$

We see that the equal signs hold in the duality relation (34) and the erasure relation (35) whenever the q-bit and its environment are in a pure state. Therefore one cannot improve on the upper bounds in these inequalities.

A *second example* is a mixed state obtained from two pure states that have the labels a and b interchanged,

$$\rho_{\mathrm{Q\&E}}=\tfrac{1}{2}\left[P_{\mathrm{Q\&E}}+(\tau+\tau^\dagger)\,P_{\mathrm{Q\&E}}(\tau+\tau^\dagger)\right] \tag{42}$$

with $P_{\mathrm{Q\&E}}$ as in (18), for which

$$w_1=w_2=\tfrac{1}{2}\,,$$
$$\rho_{\mathrm{E}}^{(1)}=\rho_{\mathrm{E}}^{(2)}=|\mathrm{E}_a\rangle w_a\langle\mathrm{E}_a|+|\mathrm{E}_b\rangle w_b\langle\mathrm{E}_b| \tag{43}$$

and

$$\chi_{\mathrm{E}}=\chi_{\mathrm{E}}^\dagger=\sqrt{w_aw_b}\,\left(|\mathrm{E}_b\rangle\mathrm{e}^{\mathrm{i}\varphi}\langle\mathrm{E}_a|+|\mathrm{E}_a\rangle\mathrm{e}^{-\mathrm{i}\varphi}\langle\mathrm{E}_b|\right). \tag{44}$$

Here we find

$$\mathscr{P}=\mathscr{D}=0\,, \tag{45}$$

and

$$\mathscr{V}=2\sqrt{w_aw_b}\;\left|\mathrm{Re}\left(\mathrm{e}^{\mathrm{i}\varphi}\langle\mathrm{E}_a|\mathrm{E}_b\rangle\right)\right|\le 2\sqrt{w_aw_b}\times\left(1-\left[\mathrm{Im}\left(\mathrm{e}^{\mathrm{i}\varphi}\langle\mathrm{E}_a|\mathrm{E}_b\rangle\right)\right]^2\right)^{1/2}=\mathscr{C}\,, \tag{46}$$

so that, in particular,

$$\mathscr{P} = \mathscr{D} = \mathscr{V} = 0 \quad \text{and} \quad \mathscr{C} = 1 \tag{47}$$

if $w_a = w_b = \frac{1}{2}$ and $\langle E_a | E_b \rangle = 0$. This example shows that QE can recover a lot of fringe visibility ($\mathscr{C} = 1$) although the alternatives remain completely indistinguishable ($\mathscr{D} = 0$). In other words, WA distinguishability is not a precondition for QE.

Likewise, full WA knowledge may be available ($\mathscr{D} = 1$) without the possibility of recovering fringes with the aid of QE ($\mathscr{C} = 0$). This can happen in the *third example*,

$$\rho_{Q\&E} = \tfrac{1}{2}\left[P_{Q\&E} + (\tau^\dagger\tau - \tau\tau^\dagger) P_{Q\&E} (\tau^\dagger\tau - \tau\tau^\dagger) \right] \tag{48}$$

with $P_{Q\&E}$ as in (18), for which

$$w_1 = w_a\,, \quad \rho_E^{(1)} = |E_a\rangle\langle E_a|\,, \quad w_2 = w_b\,,$$
$$\rho_E^{(2)} = |E_b\rangle\langle E_b|\,, \quad \text{and} \quad \chi_E = 0\,. \tag{49}$$

Here we have $\mathscr{V} = \mathscr{C} = 0$ and the values of (40) for $\mathscr{P}$ and $\mathscr{D}$, so that

$$\mathscr{P} = \mathscr{V} = \mathscr{C} = 0 \quad \text{and} \quad \mathscr{D} = 1 \tag{50}$$

obtain for $w_a = w_b = \frac{1}{2}$ and $\langle E_a | E_b \rangle = 0$.

The *fourth example* illustrates a remark after Eq. (31), namely that the WA sorting and the QE sorting can be identical. Consider

$$\rho_{Q\&E} = \tfrac{1}{2}\Big[P_{Q\&E}\big|_{|E_a\rangle = |E_b\rangle = |E_1\rangle,\, w_a = 4/5,\, w_b = 1/5,\, \varphi = 0} + P_{Q\&E}\big|_{|E_a\rangle = |E_b\rangle = |E_2\rangle,\, w_a = 1/5,\, w_b = 4/5,\, \varphi = \pi} \Big] \tag{51}$$

with $\langle E_1 | E_2 \rangle = 0$, for which

$$w_1 = \tfrac{1}{2}\,, \quad \rho_E^{(1)} = |E_1\rangle \tfrac{4}{5} \langle E_1| + |E_2\rangle \tfrac{1}{5} \langle E_2|\,,$$
$$w_2 = \tfrac{1}{2}\,, \quad \rho_E^{(2)} = |E_1\rangle \tfrac{1}{5} \langle E_1| + |E_2\rangle \tfrac{4}{5} \langle E_2|\,,$$

and

$$\chi_E = \tfrac{2}{5}\big(|E_1\rangle\langle E_1| - |E_2\rangle\langle E_2| \big)\,, \tag{52}$$

so that

$$\mathscr{P} = \mathscr{V} = 0\,, \quad \mathscr{D} = \tfrac{3}{5}\,, \quad \mathscr{C} = \tfrac{4}{5}\,. \tag{53}$$

All sortings that have $|E_1\rangle$ and $|E_2\rangle$ among the $|E_k\rangle$'s give

$$\mathscr{K}(O_E) = \mathscr{D} \quad \text{and} \quad \mathscr{V}^{(QE)}(O_E) = \mathscr{C} \tag{54}$$

and are therefore both WA sortings and QE sortings, indeed. Incidentally, the equal sign holds in (39) for this example.

The four examples given above owe their simplicity to the circumstance that the environment is essentially another q-bit. In particular, the various values of the coherence $\mathscr{C}$ in Eqs. (41), (46), (50) and (53) are all maxima of the respective knowledge $\mathscr{K}(O_E)$. This is not the case in the final *fifth example* where we must settle for the supremum of (31). Here the environment is a harmonic oscillator (ladder operators a, $a^\dagger$; Fock states $|n\rangle$ with $n = 0,1,2,\ldots$), and we consider an entangled state $\rho_{Q\&E}$ for which

$$w_1 = \tfrac{1}{2}\,,$$
$$\rho_E^{(1)} = (1-\lambda)\,\lambda^{a^\dagger a} = \sum_{n=0}^{\infty} |n\rangle(1-\lambda)\,\lambda^n\langle n|\,,$$
$$w_2 = \tfrac{1}{2}\,,$$
$$\rho_E^{(2)} = a^\dagger \frac{1}{\sqrt{aa^\dagger}}\,\rho_E^{(1)} \frac{1}{\sqrt{aa^\dagger}}\,a = \sum_{n=0}^{\infty} |n+1\rangle(1-\lambda)\,\lambda^n\langle n+1|\,,$$

and

$$\chi_E = a^\dagger \frac{1}{\sqrt{aa^\dagger}}\,\rho_E^{(1)} = \sum_{n=0}^{\infty} |n+1\rangle(1-\lambda)\,\lambda^n\langle n|\,, \tag{55}$$

where λ is a numerical parameter from the range $0 < \lambda < 1$. We find

$$\mathscr{P} = \mathscr{V} = 0\,, \quad \mathscr{D} = 1 - \lambda\,, \quad \mathscr{C} = 1\,, \tag{56}$$

but since $|\chi_E| = \rho_E^{(1)}$ and $|\chi_E^\dagger| = \rho_E^{(2)}$ are not unitarily equivalent, the option of footnote 7 is not available, and $\mathscr{C}$ is only the supremum of $\mathscr{V}^{(QE)}(O_E)$, not its maximum.

Therefore there is no O_E with $\mathscr{V}^{(QE)}(O_E) = \mathscr{C}$. At best, we can have a series $O_E^{(N)}$ $(N = 1,2,3,\ldots)$ of environment observables, for which $\mathscr{V}^{(QE)}\big(O_E^{(N)}\big) \to \mathscr{C}$ as $N \to \infty$. One such series is specified by the unitary operators

$$\exp\big(\mathrm{i}O_E^{(N)}\big) = |0\rangle\langle N| + \sum_{n=0}^{N-1} |n+1\rangle\langle n| + \sum_{n=N+1}^{\infty} |n\rangle\langle n|\,, \tag{57}$$

for example [8]. The corresponding projectors

$$P_{\mathrm{E}}^{(k)} = \begin{cases} \dfrac{1}{N+1}\displaystyle\sum_{n,m=0}^{N} |n\rangle\langle m| + \sum_{n=N+1}^{\infty} |n\rangle\langle n| & \text{for } k=0, \\ \dfrac{1}{N+1}\displaystyle\sum_{n,m=0}^{N} |n\rangle \mathrm{e}^{\mathrm{i}\frac{2\pi}{N+1}(n-m)k}\langle m| & \text{for } k=1,2,\ldots,N, \end{cases} \tag{58}$$

yield

$$\mathscr{V}^{(\mathrm{QE})}\big(O_{\mathrm{E}}^{(N)}\big) = 1-\lambda^N, \tag{59}$$

so that the QE visibility $\mathscr{V}^{(\mathrm{QE})}$ comes arbitrarily close to the coherence $\mathscr{C}=1$ without ever being equal to it.

The fifth example of Eqs. (55) is also well suited for a simple illustration of the loss of distinguishability and coherence that results from dissipative processes in the environment. In the context of (55), dissipation modeled by the master equation

$$\frac{\partial}{\partial t}\rho_{\mathrm{E}} = -\frac{\gamma}{2}\Bigg(a^\dagger\frac{1}{aa^\dagger}a\rho_{\mathrm{E}} - 2\frac{1}{\sqrt{aa^\dagger}}a\rho_{\mathrm{E}}a^\dagger\frac{1}{\sqrt{aa^\dagger}} + \rho_{\mathrm{E}}a^\dagger\frac{1}{aa^\dagger}a\Bigg) \tag{60}$$

(with $\gamma>0$) is particularly easy to study. The steady state $\rho_{\mathrm{E}}^{(\mathrm{SS})} = |0\rangle\langle 0|$ of (60) projects to the oscillator's ground state, and the temporal evolution of $\rho_{\mathrm{E}}^{(1)}$, $\rho_{\mathrm{E}}^{(2)}$, and χ_{E} is given by

$$\begin{aligned} \rho_{\mathrm{E}}^{(1)} &\to \rho_{\mathrm{E}}^{(\mathrm{SS})} + \mathrm{e}^{-(1-\lambda)\gamma t}\big(\rho_{\mathrm{E}}^{(1)} - \rho_{\mathrm{E}}^{(\mathrm{SS})}\big), \\ \rho_{\mathrm{E}}^{(2)} &\to \rho_{\mathrm{E}}^{(\mathrm{SS})} + \mathrm{e}^{-(1-\lambda)\gamma t}\big(\rho_{\mathrm{E}}^{(2)} - \rho_{\mathrm{E}}^{(\mathrm{SS})}\big), \\ \chi_{\mathrm{E}} &\to \frac{1-\lambda}{1-2\lambda}\big(\mathrm{e}^{-\gamma t/2} - \mathrm{e}^{-(1-\lambda)\gamma t}\big)a^\dagger\rho_{\mathrm{E}}^{(\mathrm{SS})} \\ &\quad + \mathrm{e}^{-(1-\lambda)\gamma t}\chi_{\mathrm{E}}, \end{aligned} \tag{61}$$

if the expressions of (55) serve as the initial conditions at $t=0$. Then

$$\mathscr{D}(t) = (1-\lambda)\mathrm{e}^{-(1-\lambda)\gamma t}$$

and

$$\mathscr{C}(t) = \frac{1}{1-2\lambda}\big[(1-\lambda)\mathrm{e}^{-\gamma t/2} - \lambda\mathrm{e}^{-(1-\lambda)\gamma t}\big] \tag{62}$$

show the exponentially fast degradation of the distinguishability and the coherence. Thus the O_{E} measurement must be performed quickly because WA knowledge and QE visibility are available only for a limited period. Note that the initial distinguishability $\mathscr{D}(t=0)=1-\lambda$ enters the effective decay rate $\gamma_{\mathrm{eff}} = (1-\lambda)\gamma$, which is a typical feature, although probably not a generic one. Equally typical is the occurrence of different time scales in $\mathscr{D}(t)$ and $\mathscr{C}(t)$, but the possibility of a much slower decay of $\mathscr{C}(t)$ than $\mathscr{D}(t)$ seems to be an atypical property of the somewhat artificial model.

2.7. Proofs of the inequalities (34), (35), and (39)

In conjunction with simple supplementary arguments, the various inequalities of Section 2.5 imply each other and, therefore, it would suffice to demonstrate just one of them, perhaps (39). But we think that there is some value in the redundancy inherent in independent proofs that differ in mathematical sophistication. Accordingly, we report (i) proofs of the duality relation (34) and the erasure relation (35) that make use of the explicit forms for the distinguishability $\mathscr{D}$ and the coherence $\mathscr{C}$ in terms of environment traces; (ii) a first proof of (39) that exploits the positivity condition (19); and (iii) a second proof of the 'strongest' inequality (39) that is based on the validity of the 'weakest' inequality (36) for each subensemble.

The method of (i) is an extension of the one used in Refs. [13–15], the approach of (ii) follows a suggestion by Hradil [16], and (iii) is similar to the reasoning in [17]. We emphasize that none of the proofs invokes a Heisenberg–Robertson inequality [18,19] of any kind [9].

[8] These $O_{\mathrm{E}}^{(N)}$ are inspired by the Pegg–Barnett papers on the phase operator [12].

[9] As shown very recently by Dürr and Rempe [20], it is possible to prove the duality relation on the basis of uncertainty relations for suitably chosen Q&E observables. These observables are, however, somewhat unusual because they depend on the given entangled state $\rho_{\mathrm{Q\&E}}$, a feature which is beyond the traditional scope of uncertainty relations of the Heisenberg–Robertson kind.

(i) Any arbitrary entangled state $\rho_{\mathrm{Q\&E}}$ is a convex sum of pure states,

$$\rho_{\mathrm{Q\&E}}=\sum_k r_k P^{(k)}_{\mathrm{Q\&E}} \quad \text{with} \quad r_k>0\,, \quad \sum_k r_k=1\,, \tag{63}$$

which could be, but need not, the spectral representation of $\rho_{\mathrm{Q\&E}}$. Accordingly we have

$$w_1\,\rho^{(1)}_{\mathrm{E}}=\sum_k r_k\bigl(|\mathrm{E}_a\rangle w_a\langle \mathrm{E}_a|\bigr)^{(k)}\,,$$
$$w_2\,\rho^{(2)}_{\mathrm{E}}=\sum_k r_k\bigl(|\mathrm{E}_b\rangle w_b\langle \mathrm{E}_b|\bigr)^{(k)}\,, \tag{64}$$

and

$$\sqrt{w_1w_2}\;\chi_{\mathrm{E}}=\sum_k r_k\Bigl(|\mathrm{E}_b\rangle\sqrt{w_aw_b}\,\mathrm{e}^{\mathrm{i}\varphi}\langle \mathrm{E}_a|\Bigr)^{(k)} \tag{65}$$

with a self-explaining notation that builds on the form (18) of a pure state. Now, the mapping $X_{\mathrm{E}}\to \mathrm{tr}_{\mathrm{E}}\{|X_{\mathrm{E}}|\}$ has all properties of a norm including, in particular, the triangle inequality

$$\mathrm{tr}_{\mathrm{E}}\{|X_{\mathrm{E}}+Y_{\mathrm{E}}|\}\le \mathrm{tr}_{\mathrm{E}}\{|X_{\mathrm{E}}|\}+\mathrm{tr}_{\mathrm{E}}\{|Y_{\mathrm{E}}|\}\,. \tag{66}$$

As a consequence thereof we have

$$\mathscr{D}\le\sum_k r_k\mathscr{D}^{(k)} \quad \text{and} \quad \mathscr{C}\le\sum_k r_k\mathscr{C}^{(k)}\,, \tag{67}$$

whereas

$$\mathscr{P}\le\sum_k r_k\mathscr{P}^{(k)} \quad \text{and} \quad \mathscr{V}\le\sum_k r_k\mathscr{V}^{(k)} \tag{68}$$

follow immediately from the familiar triangle inequality for complex numbers.

For the pairs $\mathscr{V}^{(k)}$, $\mathscr{D}^{(k)}$ and $\mathscr{P}^{(k)}$, $\mathscr{C}^{(k)}$ the equal signs hold in (34) and (35), respectively, as we have noted in the context of Eqs. (40) and (41), and therefore the Schwarz inequality implies

$$\mathscr{V}^{(j)}\mathscr{V}^{(k)}+\mathscr{D}^{(j)}\mathscr{D}^{(k)}\le 1$$
and
$$\mathscr{P}^{(j)}\mathscr{P}^{(k)}+\mathscr{C}^{(j)}\mathscr{C}^{(k)}\le 1 \tag{69}$$

for any j,k pair of indices. Together (67)–(69) establish

$$\mathscr{V}^2+\mathscr{D}^2\le\sum_{j,k} r_jr_k\bigl(\mathscr{V}^{(j)}\mathscr{V}^{(k)}+\mathscr{D}^{(j)}\mathscr{D}^{(k)}\bigr)$$
$$\le\Bigl(\sum_k r_k\Bigr)^2=1$$

and

$$\mathscr{P}^2+\mathscr{C}^2\le\sum_{j,k} r_jr_k\bigl(\mathscr{P}^{(j)}\mathscr{P}^{(k)}+\mathscr{C}^{(j)}\mathscr{C}^{(k)}\bigr)$$
$$\le\Bigl(\sum_k r_k\Bigr)^2=1\,, \tag{70}$$

which completes the proof of the inequalities (34) and (35).

(ii) Let us employ a different method, one that makes no explicit use of the triangle inequality (66), for the first proof of the inequality in (39). We introduce the numbers a_k and b_k in accordance with

$$a_k=\Bigl(\mathrm{Max}\bigl\{w_1\mathrm{tr}_{\mathrm{E}}\{P^{(k)}_{\mathrm{E}}\rho^{(1)}_{\mathrm{E}}\},w_2\mathrm{tr}_{\mathrm{E}}\{P^{(k)}_{\mathrm{E}}\rho^{(2)}_{\mathrm{E}}\}\bigr\}\Bigr)^{1/2}$$
$$\ge\Bigl(\mathrm{Min}\bigl\{w_1\mathrm{tr}_{\mathrm{E}}\{P^{(k)}_{\mathrm{E}}\rho^{(1)}_{\mathrm{E}}\},w_2\mathrm{tr}_{\mathrm{E}}\{P^{(k)}_{\mathrm{E}}\rho^{(2)}_{\mathrm{E}}\}\bigr\}\Bigr)^{1/2}$$
$$=b_k\ge 0\,, \tag{71}$$

so that the WA knowledge appears as

$$\mathscr{K}(O_{\mathrm{E}})=\sum_k\bigl(a_k^2-b_k^2\bigr)\,. \tag{72}$$

Together with the normalization

$$\sum_k\bigl(a_k^2+b_k^2\bigr)=\mathrm{tr}_{\mathrm{E}}\bigl\{w_1\,\rho^{(1)}_{\mathrm{E}}+w_2\,\rho^{(2)}_{\mathrm{E}}\bigr\}=1\,, \tag{73}$$

this establishes

$$\bigl[\mathscr{K}(O_{\mathrm{E}})\bigr]^2=1-4\sum_j a_j^2\sum_k b_k^2\,. \tag{74}$$

Turning to the QE visibility now, we note that the positivity condition (19) implies first

$$\sqrt{w_1w_2}\;\bigl|\mathrm{tr}_{\mathrm{E}}\{P^{(k)}_{\mathrm{E}}\chi_{\mathrm{E}}\}\bigr|\le a_kb_k \tag{75}$$

and then a consequent upper bound on $\mathscr{V}^{(\mathrm{QE})}(O_{\mathrm{E}})$,

$$\mathscr{V}^{(\mathrm{QE})}(O_{\mathrm{E}})\le 2\sum_k a_kb_k\,; \tag{76}$$

with this at hand,

$$\bigl[\mathscr{V}^{(\mathrm{QE})}(O_{\mathrm{E}})\bigr]^2\le 4\sum_j a_j^2\sum_k b_k^2 \tag{77}$$

follows from the Schwarz inequality. Upon combining (77) with (74), we arrive at (39), indeed.

(iii) The second proof of (39) proceeds from noting that the predictability $\mathscr{P}_k=|w_{1k}-w_{2k}|$ and the visibility $\mathscr{V}_k = 2\sqrt{w_{1k}w_{2k}}\;|\epsilon_k|$ of each subensemble must obey (36),

$$\mathscr{P}_k^2 + \mathscr{V}_k^2 \le 1\,. \tag{78}$$

As a consequence of the Schwarz inequality we then have

$$\mathscr{P}_j\mathscr{P}_k + \mathscr{V}_j\mathscr{V}_k \le 1 \tag{79}$$

for each pair of subensembles. Since the WA knowledge and the QE visibility are given by

$$\mathscr{K}(O_\mathrm{E}) = \sum_k p_k\mathscr{P}_k$$

and

$$\mathscr{V}^{(\mathrm{QE})}(O_\mathrm{E}) = \sum_k p_k\mathscr{V}_k\;, \tag{80}$$

respectively, we get

$$\begin{aligned}&\left[\mathscr{K}(O_\mathrm{E})\right]^2 + \left[\mathscr{V}^{(\mathrm{QE})}(O_\mathrm{E})\right]^2\\ &\quad= \sum_{j,k} p_j\,p_k\big(\mathscr{P}_j\mathscr{P}_k + \mathscr{V}_j\mathscr{V}_k\big) \le \sum_j p_j \sum_k p_k = 1\,,\end{aligned} \tag{81}$$

and the validity of (39) is established again.

2.8. Brief history of quantitative wave-particle duality

Historically, wave-particle duality is associated with two-way interferometers where the observed fringe pattern demonstrates wave aspects whereas knowledge about the way through the apparatus is considered particle-typical. The binary alternatives are the two ways, and which-alternative knowledge is which-way knowledge. Most of the following remarks concern this context.

It is difficult to say when the first attempts were made at formulating quantitative statements about wave-particle duality, but surely the Wootters–Zurek paper [21] of 1979 constitutes the first major publication on the subject. They dealt with WA knowledge in Young's double-slit interferometer and stated, in particular, Eq. (36) in an implicit manner (rather than the predictability $\mathscr{P}$, an equivalent entropic measure was used).

This 'weakest inequality' can also be found in a number of other publications – sometimes explicitly, sometimes more implicitly, and in a variety of notational conventions. We are aware of Refs. [22–28] but presumably there are others. In all the physical situations considered in these papers, predictability comes about because of an intentional asymmetry of the interferometer. One example is a Young interferometer with slits of different width, as in Ref. [21]; another is a Mach–Zehnder interferometer in which one of the two partial beams is attenuated behind the symmetric beam splitter at the entry port. The neutron experiments reported in Refs. [22,23] employed this method, and so did the photon experiment of Ref. [26].

Markedly different is the scenario studied by Glauber [24], who lets a single photon pass through a double-slit and then amplifies the radiation by so much that the limit of classical electromagnetism is reached. Owing to the noise that unavoidably accompanies the amplification, the resulting classical amplitudes are, as a rule, not symmetric although the quantum input into the amplifiers is. The degree of asymmetry is, of course, not predictable and each repetition of the experiment will have a random value of the asymmetry. Each repetition will also yield a full interference pattern produced by the abundance of photons present after the amplification. The fringe visibility $\mathscr{V}$ is larger if the asymmetry is less pronounced. Upon measuring the asymmetry by a number that is the analog of the predictability $\mathscr{P}$, Glauber finds that the equal sign holds in (36) for each shot, while the values of $\mathscr{P}$ and $\mathscr{V}$ vary from shot to shot in an unpredictable fashion.

The exploitation of entanglement for gaining WA knowledge beyond the predictability $\mathscr{P}$ was a central issue in the Wootters–Zurek paper as well. In the specific context of Einstein's recoiling-slit proposal (recall Bohr's account [2]), they evaluated, in essence, the WA knowledge $\mathscr{K}(O_\mathrm{E})$ for one particular $\rho_{\mathrm{Q\&E}}$ and one particular observable O_E; their results are, of course, consistent with inequality (37). In hindsight it is certainly a pity that Wootters and Zurek did not ask if another observable could supply more WA knowledge than the momentum of the recoiling slit they considered (which, incidentally, is not the optimal observable).

A more general and more systematic treatment was not given until much later. In Ref. [29], Jaeger, Shimony, and Vaidman (JSV) considered an arbitrary ρ_Q and – assuming that its impurity originates in an entanglement of the q-bit with some other degree(s) of freedom – they wondered to which extent WA knowledge can then be available beyond the predictability $\mathscr{P}$ inherent in the ρ_{Q} of interest. In the JSV approach, ρ_{Q} is given and various $\rho_{\mathrm{Q\&E}}$'s with $\mathrm{tr}_{\mathrm{E}}\{\rho_{\mathrm{Q\&E}}\}=\rho_{\mathrm{Q}}$ are considered, which represent different 'preparations' in their terminology. Through an analysis of an exhaustive list of cases, JSV found that the total WA knowledge can never exceed $\sqrt{1-\mathscr{V}^{\,2}}$, and that this upper bound is reached if the 'preparation' $\rho_{\mathrm{Q\&E}}$ is a pure state (cf. the first example in Section 2.6) [10].

The JSV trio introduced the distinguishability $\mathscr{D}$ in the sense of the equation on the left in (26), and so they arrived at the inequalities (37) and (34). They could have also determined an observable O_{E} that actually maximizes $\mathscr{K}(O_{\mathrm{E}})$ for an arbitrary $\rho_{\mathrm{Q\&E}}$, but unfortunately they have not [11]. As a consequence, JSV did not find the equation on the right in (26), and so they missed a complete proof of the duality relation by a hair.

Somewhat later, and independent of the work by JSV, the issue was addressed anew in Ref. [13], where generic two-way interferometers were studied. The explicit expression for $\mathscr{D}$ in terms of the trace in (26) was then found and, based on this insight, the duality relation (34) was proven.

The proof of (34) in [13] makes use of a specific structure of $\rho_{\mathrm{Q\&E}}$ that obtains when a disentangled state is subject to unitary transformations of a restricted kind. This restriction is lifted in Refs. [14,15], but unitary transformations are still an element. Therefore, the validity of those proofs is not obvious in situations where dissipation or other non-unitary processes are present, as exemplified by Eqs. (60)–(62) in Section 2.6. No such unitarity assumptions enter the proofs given in Section 2.7, which are based on nothing more than the positivity of $\rho_{\mathrm{Q\&E}}$.

The entangled states $\rho_{\mathrm{Q\&E}}$ studied in the recent paper [32] by Björk and Karlsson (BK) come from disentangled ones and restricted unitary transformations, but the consequent lack of generality has no bearing on their findings. BK proceed from noting that degrees of freedom of two kinds contribute to the environment, external ones and internal ones; symbolically: $\mathrm{E}=\mathrm{E}_1\&\mathrm{E}_2$. For example, the center-of-mass degree of freedom of Einstein's recoiling slit is external; but the hyperfine degree-of-freedom used for WA detection in the atom interferometer of Refs. [33,34,17] is internal, and so is the photon's polarization in the Mach–Zehnder interferometer of Ref. [35]. BK then note that the O_{E}'s that can be measured realistically are not quite arbitrary but rather of the product form $O_{\mathrm{E}}=O_{\mathrm{E}_1}O_{\mathrm{E}_2}$, and therefore the practically available WA knowledge may be considerably less than the distinguishability $\mathscr{D}$ of (26).

The keeping apart of the external and the internal environment enables BK to perform more sophisticated sortings. In particular, a measurement on the internal E_2 part alone yields subensembles of the $\mathrm{Q\&E}_1$ kind, labeled by ν, say. Each of them carries a weight $\tilde{p}_\nu$ and has a WA knowledge $\mathscr{K}_\nu(O_{\mathrm{E}_1})$ and a QE visibility $\mathscr{V}_\nu^{\,(\mathrm{QE})}(O_{\mathrm{E}_1})$ for each E_1 observable O_{E_1}. Of course, they obey inequality (39) individually,

$$\left[\mathscr{K}_\nu\left(O_{\mathrm{E}_1}\right)\right]^2+\left[\mathscr{V}_\nu^{\,(\mathrm{QE})}\left(O_{\mathrm{E}_1}\right)\right]^2\le 1$$

$$\text{for all } \nu \text{ and all } O_{\mathrm{E}_1}, \tag{82}$$

and by summing over ν, BK derive the inequality

$$\sum_\nu\left(\left[\tilde{p}_\nu\mathscr{K}_\nu\left(O_{\mathrm{E}_1}\right)\right]^2+\left[\tilde{p}_\nu\mathscr{V}_\nu^{\,(\mathrm{QE})}\left(O_{\mathrm{E}_1}\right)\right]^2\right)\le\sum_\nu\tilde{p}_\nu^2 \tag{83}$$

[this is BK's Eq. (11) in the present notation] and some similar ones. These BK inequalities concern summed squares and, accordingly, they are different from the inequalities in Section 2.5, which make statements about squared sums; of course, the two kinds are closely related. In the limiting situation of a single ν value, (83) reduces to (39), and therefore Ref. [32] contains an implicit derivation of Eq. (39), the 'strongest inequality'.

The BK inequalities are valuable in themselves, but we think that the main merit of BK's work consists in their considerations concerning quantum erasure. They introduce the relevant quantities, namely the WA knowledge $\mathscr{K}$ (which they call

---

[10] For an explicit example of a realistic experimental situation of this kind see Fig. 3 in [30], which builds on problem 9-6 in [31].

[11] A strategy for determining an optimal O_{E} is also essential for the demonstration that the knowledge $\mathscr{K}(O_{\mathrm{E}})$ actually has a maximum, not just a supremum. But this subtlety is of lesser importance.

'measured distinguishability') and the QE visibility $\mathscr{V}^{(\mathrm{QE})}$ (the 'conditioned visibility' in their terminology), and study some of their properties. In contrast to our approach, however, BK are not interested in maximizing $\mathscr{V}^{(\mathrm{QE})}$ but rather in minimizing $\mathscr{K}$. Consistent with the hierarchy (29), they find that 'complete quantum erasure' (that is: $\mathscr{K}=0$) can only be obtained if the predictability $\mathscr{P}$ vanishes. The difference in attitude (they minimize, we maximize) is very likely the reason why BK did not get to a statement equivalent to (31).

Although the quantitative aspects of wave-particle duality have been dealt with in a fair number of theoretical papers, there are just a few experimental studies. The neutron experiments of Refs. [22,23] and the photon experiments of Ref. [26] addressed inequality (36). The duality relation (34), or rather its 'experimental version' (37), were tested in the atom interferometer of Ref. [34] and in the photon interferometer of Ref. [35]. The latter also reports an experimental realization of full quantum erasure without available WA distinguishability ($\mathscr{V}^{(\mathrm{QE})}\lesssim 1$ and $\mathscr{D}\gtrsim 0$), as in the second example of Section 2.6; for additional details consult Ref. [36]. The measurements of Ref. [17] concern inequality (39) and a subensemble variant thereof that is essentially identical with (78). All experimental findings are in good agreement with the respective theoretical expectations.

2.9. Mathematical addendum: maximum in (26), supremum in (31)

The modulus of an operator and related concepts are not part of the standard material covered in lectures on quantum mechanics, and many textbooks do not address these issues. It may, therefore, be helpful to review some basic features concisely. In particular, we shall explain why $\mathscr{D}$ is the maximum of all $\mathscr{K}(O_{\mathrm{E}})$ in (26), whereas $\mathscr{C}$ is only the supremum of the $\mathscr{V}^{(\mathrm{QE})}(O_{\mathrm{E}})$'s in (31).

Any operator X with $0<\mathrm{tr}\{|X|\}<\infty$ can be written in the form

$$X=\sum_{m=0}^{M}|\bar\psi_m\rangle x_m\langle\psi_m| \tag{84}$$

with

$$x_0\geq x_1\geq x_2\geq\ \ldots\ \geq x_M>0 \tag{85}$$

and

$$\langle\psi_\ell|\psi_m\rangle=\delta_{\ell m}\,,\quad \langle\bar\psi_\ell|\bar\psi_m\rangle=\delta_{\ell m}\,,\quad \text{for}\ \ell,m=0,1,2,\ldots,M, \tag{86}$$

where M is a non-negative integer that could be as large as $M=\infty$. The two completeness relations

$$\sum_{m=-M_1}^{M}|\psi_m\rangle\langle\psi_m|=1$$

and

$$\sum_{m=-M_2}^{M}|\bar\psi_m\rangle\langle\bar\psi_m|=1 \tag{87}$$

(with $M_1\geq 0$ and $M_2\geq 0$) may involve more $|\psi_m\rangle$'s or $|\bar\psi_m\rangle$'s than those appearing in (84). The orthonormality relations in (86) must then be extended correspondingly to cover negative indices as well. If there is a finite number of terms in (84), then M_1 and M_2 must be the same, including the possibility $M_1=M_2=\infty$. But in the case of $M=\infty$ it can happen that $M_1\neq M_2$. The χ_{E} of the fifth example in Section 2.6 is of this kind since $M=\infty$, $M_1=1$, and $M_2=0$ there.

In view of

$$|X|=\sqrt{X^\dagger X}\ =\sum_{m=0}^{M}|\psi_m\rangle x_m\langle\psi_m|\,,$$

$$|X^\dagger|=\sqrt{XX^\dagger}\ =\sum_{m=0}^{M}|\bar\psi_m\rangle x_m\langle\bar\psi_m|\,, \tag{88}$$

the number M_1 is the count of null eigenvalues of $|X|$, and M_2 that of $|X^\dagger|$. Consequently, $|X|$ and $|X^\dagger|$ are unitarily equivalent if $M_1=M_2$, and only then.

With the optimizations of (26) and (31) in mind, we now consider a complete, orthogonal set of kets $|\alpha_k\rangle$,

$$\sum_k|\alpha_k\rangle\langle\alpha_k|=1\,,\quad \langle\alpha_j|\alpha_k\rangle=\delta_{jk}\,, \tag{89}$$

and show first that

$$\sum_k\left|\langle\alpha_k|X|\alpha_k\rangle\right|\leq\mathrm{tr}\{|X|\}\,, \tag{90}$$

and establish then the conditions under which the equal sign holds. This will tell us that the optimization of $\mathscr{K}(O_{\mathrm{E}})$ in (26) yields a maximum, whereas that of $\mathscr{V}^{(\mathrm{QE})}(O_{\mathrm{E}})$ in (31) might not.

After inserting (84) on the left-hand side of (90) we get

$$\begin{aligned}
&\sum_k \left|\langle \alpha_k | X | \alpha_k \rangle\right| \\
&\quad = \sum_k \left| \sum_{m=0}^{M} x_m \left\langle \alpha_k \middle| \bar{\psi}_m \right\rangle \langle \psi_m | \alpha_k \rangle \right| \\
&\quad \le \sum_{m=0}^{M} x_m \sum_k \left| \langle \alpha_k | \bar{\psi}_m \rangle \langle \psi_m | \alpha_k \rangle \right| \\
&\quad \le \sum_{m=0}^{M} x_m \left[\sum_j \left| \langle \alpha_j | \bar{\psi}_m \rangle \right|^2 \sum_k \left| \langle \psi_m | \alpha_k \rangle \right|^2 \right]^{1/2} \\
&\quad = \sum_{m=0}^{M} x_m = \operatorname{tr}\{ |X| \}\,, \qquad (91)
\end{aligned}$$

so that (90) is indeed true for each set of $|\alpha_k\rangle$'s. For the equal sign to apply in (90), both the triangle inequality and the Schwarz inequality that are used in the intermediate steps of (91) must be equalities. This is only the case if

$$\langle \alpha_k | \bar{\psi}_m \rangle = \mathrm{e}^{\mathrm{i}\beta_k} \langle \alpha_k | \psi_m \rangle$$

$$\text{for all } k \text{ and all } m = 0,1,\ldots,M \qquad (92)$$

with real phases β_k that do not depend on the value of m.

Now, supposing that there are $\langle \alpha_k|$'s with the property (92), their completeness implies

$$|\bar{\psi}_m\rangle = U |\psi_m\rangle$$

with

$$U = (U^\dagger)^{-1} = \sum_k |\alpha_k\rangle \mathrm{e}^{\mathrm{i}\beta_k} \langle \alpha_k|\,, \qquad (93)$$

so that

$$X = U|X| = |X^\dagger| U\,. \qquad (94)$$

In other words, if there are bras $\langle \alpha_k|$ that obey (92), then $|X|$ and $|X^\dagger|$ are unitarily equivalent. The reverse is, of course, also true: If there is a unitary U that relates $|X|$ to $|X^\dagger|$, then (92) holds for the eigenbras of U.

Such is the situation in Eqs. (25) and (26), where $X = w_1 \rho_{\mathrm{E}}^{(1)} - w_2 \rho_{\mathrm{E}}^{(2)}$ is selfadjoint and $|X|$ is simply identical with $|X^\dagger|$. The optimum of $\mathscr{K}(O_{\mathrm{E}})$ is therefore a maximum, which is given by the environment trace on the right-hand side of (26). Further we note that the phase factors in (92) can only be equal to $+1$ or -1 here, and this observation is the basis for (27).

By contrast, in (30) and (31) we deal with χ_{E}, for which the case $M_1 \neq M_2$ is a real possibility (recall the fifth example in Section 2.6). A set of $\langle \alpha_k|$'s obeying (92) is not available then, and it remains to be shown that the right-hand side of (90) is the supremum to the left-hand sides, not just an upper bound.

For this purpose we consider the sequence $X^{(N)}$ defined by

$$X^{(N)} = \sum_{m=0}^{N} |\bar{\psi}_m\rangle x_m \langle \psi_m| \quad \text{for } N = 0,1,2,\ldots\,, \qquad (95)$$

and the associated sequence of bases $|\alpha_k^{(N)}\rangle$, which are such that

$$\sum_k \left| \langle \alpha_k^{(N)} | X^{(N)} | \alpha_k^{(N)} \rangle \right| = \operatorname{tr}\{ |X^{(N)}| \} = \sum_{m=0}^{N} x_m\,. \qquad (96)$$

Since

$$\begin{aligned}
&\sum_k \left| \langle \alpha_k^{(N)} | (X - X^{(N)}) | \alpha_k^{(N)} \rangle \right| < \operatorname{tr}\{ |X - X^{(N)}| \} \\
&\quad = \sum_{m=N+1}^{\infty} x_m = \operatorname{tr}\{|X|\} - \operatorname{tr}\{ |X^{(N)}| \}\,, \qquad (97)
\end{aligned}$$

we get

$$\begin{aligned}
\operatorname{tr}\{|X|\} &> \sum_k \left| \langle \alpha_k^{(N)} | X | \alpha_k^{(N)} \rangle \right| \\
&= \sum_k \left| \langle \alpha_k^{(N)} | \left[X^{(N)} + (X - X^{(N)}) \right] | \alpha_k^{(N)} \rangle \right| \\
&> \operatorname{tr}\{ |X^{(N)}| \} - \operatorname{tr}\{ |X - X^{(N)}| \} \\
&= \operatorname{tr}\{|X|\} - 2\operatorname{tr}\{ |X - X^{(N)}| \} \qquad (98)
\end{aligned}$$

after using the standard triangle inequality for complex numbers. In the limit $N \to \infty$ we have

$$\operatorname{tr}\{ |X - X^{(N)}| \} \to 0 \qquad (99)$$

and therefore

$$\sum_k \left| \langle \alpha_k^{(N)} | X | \alpha_k^{(N)} \rangle \right| \to \operatorname{tr}\{|X|\}\,, \qquad (100)$$

which shows that the right-hand side in (90) is the supremum of all possible left-hand sides, indeed. It follows that the trace in (31) is the supremum of the $\mathscr{K}(O_E)$ values for all environment observables O_E conceivable.

3. Realistic examples

The 1982 proposal by Scully and Drühl [4] became (almost [12]) a reality in 1993 [37], and has been described repeatedly, recently in Chapter 21 of [7]. Rather than applying the concepts of Section 2 to this example or to the much-discussed thought experiment of Ref. [40] (see also [41,3,42] and Chapter 20 in [7]), we treat two other physical situations that have been studied by Marlan.

The first concerns the entanglement between the spin of an atom and its center-of-mass motion that results from the interaction with a Stern–Gerlach magnet [43–46]. The second is the quantum-optical Ramsey interferometer of Ref. [47] (see also Section 20.4 in [7]), which – so we hope – will become an experimental reality eventually.

3.1. Stern–Gerlach magnets

The splitting of a beam of magnetic spin-$\frac{1}{2}$ atoms by a Stern–Gerlach (SG) apparatus is an example that is as instructive as it is elementary. The spin-up/down alternatives (in z direction, say) constitute the q-bit, and the environment consists of the center-of-mass degree(s) of freedom of the atom.

Under idealized circumstances, one regards the longitudinal motion of the atom as classical and approximates the magnetic field as a purely transverse field of constant gradient. Prior to the passage through the SG magnet the q-bit ($\equiv$ spin-$\frac{1}{2}$ degree of freedom) is not entangled with the environment ($\equiv$ motion in the transverse z direction), so that the initial state is of the product form

$$\rho^{(0)}_{\mathrm{Q\&E}} \mathrel{\hat{=}} \frac{1}{2}\begin{pmatrix} 1+s_z & s_x - \mathrm{i}s_y \\ s_x+\mathrm{i}s_y & 1-s_z \end{pmatrix}\rho^{(0)}_{\mathrm{E}}$$

with

$$s_x^2+s_y^2+s_z^2 \le 1\,, \tag{101}$$

where the numerical vector $\boldsymbol{s}=(s_x,s_y,s_z)$ is the expectation value of Pauli's spin vector operator $\boldsymbol{\sigma}$.

[12] The experiment of Ref. [37] was not performed in a one-photon-at-a-time fashion, and therefore it is not conclusive [38,39].

With the said idealizations the final state is given by

$$\rho_{\mathrm{Q\&E}} = U^{\dagger}_{\mathrm{Q\&E}}\,\rho^{(0)}_{\mathrm{Q\&E}}U_{\mathrm{Q\&E}}\,, \tag{102}$$

where

$$U_{\mathrm{Q\&E}} \mathrel{\hat{=}} \begin{pmatrix} U_{\mathrm{E}} & 0 \\ 0 & U^{\dagger}_{\mathrm{E}} \end{pmatrix} \tag{103}$$

with

$$U_{\mathrm{E}} = \mathrm{e}^{\mathrm{i}\phi_L/2}\mathrm{e}^{\mathrm{i}(p_z\Delta z - z\Delta p_z)/\hbar} \tag{104}$$

is the unitary operator that summarizes the effect of the SG magnet. In (104), the operators z and p_z are the transverse position and momentum, respectively, the numbers Δz and Δp_z are the corresponding net displacements of the spin-up component, and ϕ_L is the accumulated Larmor precession angle.

Accordingly, in the entangled final state (102) we have

$$w_1 = \tfrac{1}{2}(1+s_z)\,,\quad \rho^{(1)}_{\mathrm{E}} = U^{\dagger}_{\mathrm{E}}\,\rho^{(0)}_{\mathrm{E}}U_{\mathrm{E}}\,,$$

$$w_2 = \tfrac{1}{2}(1-s_z)\,,\quad \rho^{(2)}_{\mathrm{E}} = U_{\mathrm{E}}\,\rho^{(0)}_{\mathrm{E}}U^{\dagger}_{\mathrm{E}}\,,$$

and

$$\sqrt{w_1w_2}\,\chi_{\mathrm{E}} = \tfrac{1}{2}(s_x+\mathrm{i}s_y)U_{\mathrm{E}}\,\rho^{(0)}_{\mathrm{E}}U_{\mathrm{E}}\,. \tag{105}$$

Whereas the predictability $\mathscr{P}=|s_z|$ and the coherence $\mathscr{C}=|s_x+\mathrm{i}s_y|$ are not affected by the interaction with the SG magnet, the distinguishability $\mathscr{D}$ and the visibility $\mathscr{V}$ do depend on Δz and Δp_z (but not on ϕ_L). In particular, we have

$$\mathscr{V} = \mathscr{C}\,\left|\left\langle \mathrm{e}^{2\mathrm{i}(p_z\Delta z - z\Delta p_z)/\hbar}\right\rangle^{(0)}_{\mathrm{E}}\right|\,, \tag{106}$$

so that $\mathscr{V} \lesssim \mathscr{C}$ requires that

$$|\Delta z\,\delta p_z| \ll \hbar \quad\text{and}\quad |\Delta p_z\,\delta z| \ll \hbar \tag{107}$$

hold, where δz and δp_z are the initial spreads in position and momentum that are subject to $\delta z\delta p_z \ge \hbar/2$ in accordance with Heisenberg's uncertainty relation. The implied condition

$$|\Delta z\,\Delta p_z|\,(\ll)^2\,\hbar\,, \tag{108}$$

which refers to properties of the macroscopic SG magnet, is the reason why it is so very difficult to recover the initial spin state after a SG apparatus has split a beam of magnetic atoms in two.

The problem of recombining the partial beams of a SG experiment is treated in Refs. [45,46] in detail and reviewed in Refs. [48,49] [13]. It is found that an unavoidable loss of spin coherence (that is: $\mathscr{C} < |s_x + \mathrm{i}s_y|$) occurs in all realistic magnetic fields, although a U_E of the idealized form (103) may be a very good approximation even under realistic circumstances.

3.2. Ramsey interferometers

The phase shift in a traditional Ramsey interferometer originates in the slight detuning between the classical radiation in the separated Ramsey zones and the quantum transition under investigation. Alternatively, one could use resonant radiation and introduce a relative phase with other means. For our purposes the second option is more convenient and, following Refs. [47,14], we consider the situation that is depicted in Fig. 3(c) and specified in the figure caption.

Prior to the atom-photon interaction in the resonator, we have the initial state

$$\rho^{(0)}_{\mathrm{Q\&E}} = \tau^\dagger\tau\,\rho^{(0)}_\mathrm{E} \,\widehat{=}\, \begin{pmatrix} \rho^{(0)}_\mathrm{E} & 0 \\ 0 & 0 \end{pmatrix}, \tag{109}$$

where $\rho^{(0)}_\mathrm{E}(a^\dagger,a)$ is a function of the photon ladder operators $a^\dagger$ and a. The interaction turns $\rho^{(0)}_{\mathrm{Q\&E}}$ into the final state $\rho_{\mathrm{Q\&E}}$ according to (102) with [14]

$$U_{\mathrm{Q\&E}} \,\widehat{=}\, \begin{pmatrix} \cos\left(\varphi\sqrt{aa^\dagger}\,\right) & \sin\left(\varphi\sqrt{aa^\dagger}\,\right)(aa^\dagger)^{-1/2}a \\ -a^\dagger(aa^\dagger)^{-1/2}\sin\left(\varphi\sqrt{aa^\dagger}\,\right) & \cos\left(\varphi\sqrt{a^\dagger a}\,\right) \end{pmatrix}, \tag{110}$$

here, wherein the parameter φ is the accumulated (vacuum) Rabi angle that measures the net strength of the interaction. As a consequence, we have

$$\begin{aligned} w_1\,\rho^{(1)}_\mathrm{E} &= \cos\left(\varphi\sqrt{aa^\dagger}\,\right)\rho^{(0)}_\mathrm{E}\cos\left(\varphi\sqrt{aa^\dagger}\,\right), \\ w_2\,\rho^{(2)}_\mathrm{E} &= a^\dagger\left(aa^\dagger\right)^{-1/2}\sin\left(\varphi\sqrt{aa^\dagger}\,\right)\rho^{(0)}_\mathrm{E} \\ &\quad\times\sin\left(\varphi\sqrt{aa^\dagger}\,\right)\left(aa^\dagger\right)^{-1/2}a, \end{aligned} \tag{111}$$

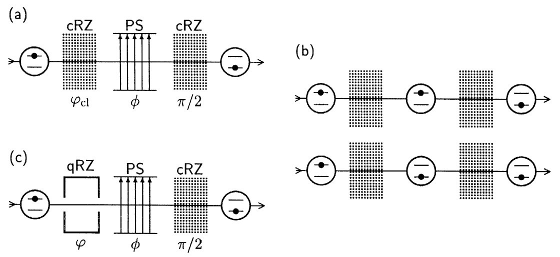

Fig. 3. Ramsey interferometer. (a) A two-level atom passes through two resonant microwave fields that play the roles of classical Ramsey zones (cRZ's). The first cRZ (of variable strength, measured by the angle parameter φ_cl) prepares the atom in a coherent superposition of the dynamically relevant levels $|\mathrm{Q}_1\rangle$ and $|\mathrm{Q}_2\rangle$. A static electric field acts as a phase shifter (PS) inasmuch as the differential Stark shift introduces a relative phase ϕ between the two levels. The PS and the second cRZ (of fixed strength $\varphi_\mathrm{cl} = \pi/2$) allow for a measurement of $I_\mathrm{Q}(\phi)$ because state-sensitive detection of the atom behind the second cRZ yields the probabilities for $I_\mathrm{Q}(\phi)=1$ and $I_\mathrm{Q}(\phi)=0$ at the intermediate stage after the first cRZ and the PS. (b) The ways through the interferometer are the basic alternatives of the q-bit in question. They are characterized by the state of the atom at the central stage. In one way the atomic transition happens in the first cRZ, in the other way it occurs in the second cRZ. (c) Upon replacing the microwave field of the first cRZ by the quantized field of a resonator, one has a quantum Ramsey zone (qRZ) instead. The atom-photon interaction entangles the $|\mathrm{Q}_1\rangle/|\mathrm{Q}_2\rangle$ q-bit with the photonic degree of freedom that constitutes the environment.

and

$$\sqrt{w_1w_2}\;\chi_\mathrm{E} = a^\dagger\left(aa^\dagger\right)^{-1/2}\sin\left(\varphi\sqrt{aa^\dagger}\,\right)\rho^{(0)}_\mathrm{E}\cos\left(\varphi\sqrt{aa^\dagger}\,\right) \tag{112}$$

for the ingredients of the final state. The predictability and the visibility

$$\begin{aligned} \mathscr{P} &= \left|\left\langle\cos\left(2\varphi\sqrt{aa^\dagger}\,\right)\right\rangle^{(0)}_\mathrm{E}\right|, \\ \mathscr{V} &= 2\left|\left\langle\cos\left(\varphi\sqrt{aa^\dagger}\,\right)a^\dagger\left(aa^\dagger\right)^{-1/2}\sin\left(\varphi\sqrt{aa^\dagger}\,\right)\right\rangle^{(0)}_\mathrm{E}\right|, \end{aligned} \tag{113}$$

involve the main diagonal and the first side-diagonal of the Fock matrix to the initial photon state $\rho^{(0)}_\mathrm{E}$,

$$\begin{aligned} \mathscr{P} &= \left|\sum_{n=0}^{\infty}\langle n|\rho^{(0)}_\mathrm{E}|n\rangle\cos\left(2\varphi\sqrt{n+1}\,\right)\right|, \\ \mathscr{V} &= 2\left|\sum_{n=1}^{\infty}\langle n-1|\rho^{(0)}_\mathrm{E}|n\rangle\sin\left(\varphi\sqrt{n}\,\right)\right. \\ &\qquad\left.\times\cos\left(\varphi\sqrt{n+1}\,\right)\right|, \end{aligned} \tag{114}$$

[13] What is called 'spin coherence' in Refs. [45,46,48,49] is the visibility $\mathscr{V}$, not the coherence $\mathscr{C}$ of the present, more general, discussion.

respectively. Generally valid expressions of this relative simplicity are not available for the distinguishability and the coherence.

If $\rho_{\mathrm{E}}^{(0)}$ is a highly excited coherent state of amplitude $r\exp(\mathrm{i}\theta)$, then the replacements

$$a \to r\mathrm{e}^{\mathrm{i}\theta}, \quad a^{\dagger} \to r\mathrm{e}^{-\mathrm{i}\theta}, \quad \sqrt{aa^{\dagger}} \simeq \sqrt{a^{\dagger}a} \to r \gg 1 \tag{115}$$

are permissible and

$$\rho_{\mathrm{Q\&E}} \simeq \rho_{\mathrm{Q}}\ \rho_{\mathrm{E}}^{(0)}$$

with

$$\rho_{\mathrm{Q}} \mathrel{\widehat{=}} \frac{1}{2}\begin{pmatrix} 1+\cos(2r\varphi) & \mathrm{e}^{\mathrm{i}\theta}\sin(2r\varphi) \\ \mathrm{e}^{-\mathrm{i}\theta}\sin(2r\varphi) & 1-\cos(2r\varphi) \end{pmatrix} \tag{116}$$

obtains. In this limit, thus, the qRZ performs like a cRZ with $\varphi_{\mathrm{cl}} = 2r\varphi$, as it should. Then the atom does not get entangled with the photon field, so that $\mathscr{P} = \mathscr{D} = |\cos(\varphi_{\mathrm{cl}})|$, $\mathscr{V} = \mathscr{C} = |\sin(\varphi_{\mathrm{cl}})|$, and $\mathscr{P}^2 + \mathscr{V}^2 = 1$.

If, however, the initial photon state has a definite photon number, $\rho_{\mathrm{E}}^{(0)} = |n\rangle\langle n|$, then

$$\begin{aligned} \mathscr{P} &= \left|\cos\left(2\varphi\sqrt{n+1}\right)\right| \le \mathscr{D} = 1, \\ \mathscr{C} &= \left|\sin\left(2\varphi\sqrt{n+1}\right)\right| \ge \mathscr{V} = 0. \end{aligned} \tag{117}$$

In this case, the equal signs hold in the duality relation (34) and the erasure relation (35).

Another possibility worth mentioning is an initial state that is diagonal in the photon number,

$$\rho_{\mathrm{E}}^{(0)} = f(a^{\dagger}a), \tag{118}$$

for which the thermal state is an example. Here we have, for instance,

$$\mathscr{V} = 0 \le \mathscr{C} = \sum_{n=0}^{\infty} f(n)\left|\sin\left(2\varphi\sqrt{n+1}\right)\right|, \tag{119}$$

but the loss of visibility is not compensated for by a large distinguishability because

$$\begin{aligned} \mathscr{D} = \sum_{n=0}^{\infty} \Big| & f(n)\cos^2\left(\varphi\sqrt{n+1}\right) \\ & - f(n-1)\sin^2\left(\varphi\sqrt{n}\right)\Big| \end{aligned} \tag{120}$$

could be rather small. If the probabilities $f(n)$ are positive for arbitrarily large n values, then one has the ‘supremum, not maximum’ situation of the fifth example in Section 2.6.

One should not conclude, however, that a small spread in the photon number is needed for a good distinguishability. A counter example is given by states of definite parity, characterized by

$$(-1)^{a^{\dagger}a}\rho_{\mathrm{E}}^{(0)} = \rho_{\mathrm{E}}^{(0)}(-1)^{a^{\dagger}a} = \pm\rho_{\mathrm{E}}^{(0)}, \tag{121}$$

for which we always have $\mathscr{D} = 1$ and $\mathscr{V} = 0$, whether the photon number has a large variance or a small one.

The actual experimental realization of the Ramsey interferometer of Fig. 3(c) would be highly interesting, but the challenge has not been met so far. The dissipation that results from the finite quality factor of the resonator would lead to time dependent values of $\mathscr{D}$ and $\mathscr{C}$, similarly to the fifth example in Section 2.6, and experimental studies of this time dependence could be very instructive in addition. Another scenario, in which a non-resonant atom-photon interaction established entanglement, was realized to some extent in the experiment reported in Ref. [50], and the systematic decrease of fringe visibility was observed. No WA knowledge was extracted, however, and therefore the inequalities of Section 2.5 are still waiting for a test in such an experiment [14].

4. Summary

We have presented a fully general treatment of q-bits entangled with other degrees of freedom that make up their environment. Measurements of environment observables enable the experimenter to sort the q-bits into subensembles. The particular sorting scheme can yield a lot of WA knowledge or a large QE visibility. There are bounds on the possible values for either quantity, and their extreme values must obey the familiar duality relation (34) and the new erasure relation (35). We have given proofs for these inequalities and related ones. An experimental scheme that seems to be sufficiently flexible to allow for tests of many different aspects of the various

[14] The limitations imposed by the experimental restriction of the O_{E}'s to the ones derived from the interaction of single atoms with the photon field make it very difficult to get to the equal sign in (39), but it is possible to get close [51].

inequalities is given by a Ramsey interferometer for atoms with a quantum Ramsey zone at the initial stage.

Acknowledgements

B.G.E. is grateful to H. Rauch and collaborators for their hospitality at the Atominstitut in Vienna, where part of this work was done, and he thanks the Technical University of Vienna for financial support. B.G.E. would also like to thank P. Zoller and the University of Innsbruck for their hospitality and financial support during a later stage of this project. We acknowledge helpful discussions with H.-J. Briegel, S. Dürr, D.A. Harmin, Z. Hradil, P.G. Kwiat, J. Martinez-Linares, G. Rempe, G. Süßmann, and, of course, M.O. Scully.

References

[1] H.J. Groenewold in a discussion recorded in Observation and Interpretation, S. Körner (Ed.), Academic Press, New York, 1957, p. 139.
[2] N. Bohr, Discussions with Einstein on Epistemological Problems in Atomic Physics, in: P.A. Schilpp (Ed.), Albert Einstein: Philosopher-Scientist, Library of Living Philosophers, Evanston, 1949.
[3] S.M. Tan, D.F. Walls, Phys. Rev. A 47 (1993) 4663.
[4] M.O. Scully, K. Drühl, Phys. Rev. A 25 (1982) 2208.
[5] M. Hillery, M.O. Scully, On State Reduction and Observation in Quantum Optics: Wigner's Friends and Their Amnesia, in: P. Meystre, M.O. Scully (Eds.), Quantum Optics, Experimental Gravity, and Measurement Theory, NATO ASI Series, Vol. B94, Plenum Press, New York, 1983.
[6] P. Meystre, M.O. Scully, Observation, State Reduction, and the Quantum Eraser in Quantum Optics, in: F.T. Arecchi, F. Strumia, H. Walther (Eds.), Advances in Laser Spectroscopy, NATO ASI Series, Vol. B95, Plenum Press, New York, 1983.
[7] M.O. Scully, M.S. Zubairy, Quantum optics, Cambridge University Press, Cambridge, 1997.
[8] N. Akhieser, I.M. Glazman, Theory of Linear Operators in Hilbert Space, Ungar, New York, 1961.
[9] M. Reed, B. Simon, Methods of Modern Mathematical Physics, Vol. 1, Academic Press, New York, 1972.
[10] W. Thirring, A Course in Mathematical Physics: Quantum Mechanics of Large Systems, Vol. 4, Springer, Vienna, 1980.
[11] J. Glimm, A. Jaffe, Quantum physics. A functional integral point of view, Springer, Berlin, 1987.
[12] D.T. Pegg, S.M. Barnett, Europhys. Lett. 6 (1988) 483.
[13] B.-G. Englert, Phys. Rev. Lett. 77 (1996) 2154.
[14] B.-G. Englert, Acta Phys. Slov. 46 (1996) 249.
[15] B.-G. Englert, Wave-particle duality quantified in: D. Han, J. Janszky, Y.S. Kim, V.I. Man'ko (Eds.), Fifth International Conference on Squeezed States and Uncertainty Relations (Balatonfüred 1997), NASA/CP-1998-206855, Greenbelt, 1998.
[16] Z. Hradil, private communication (1998).
[17] S. Dürr, G. Rempe, Opt. Commun. (same issue).
[18] W. Heisenberg, Z. Phys. 43 (1927) 172.
[19] H.P. Robertson, Phys. Rev. 34 (1929) 163.
[20] S. Dürr, G. Rempe, Can wave-particle duality be based on the uncertainty relation?, Am. J. Phys. (submitted).
[21] W.K. Wootters, W.H. Zurek, Phys. Rev. D 19 (1979) 473.
[22] H. Rauch, J. Summhammer, Phys. Lett. A104 (1984) 44.
[23] J. Summhammer, H. Rauch, D. Tuppinger, Phys. Rev. A 36 (1987) 4447.
[24] R. Glauber, Ann. N.Y. Acad. Sci. 480 (1986) 336.
[25] Ph. Grangier, doctoral thesis (Université de Paris-Sud, Orsay, 1986).
[26] P. Mittelstaedt, A. Prieur, R. Schieder, Found. Phys. 17 (1987) 891.
[27] D.M. Greenberger, A. Yasin, Phys. Lett. A128 (1988) 391.
[28] L. Mandel, Opt. Lett. 16 (1991) 1882.
[29] G. Jaeger, A. Shimony, L. Vaidman, Phys. Rev. A 51 (1995) 54.
[30] B.-G. Englert, Z. Naturforsch. 54a (1999) 11.
[31] L. Ballentine, Quantum Mechanics (Prentice Hall, Englewood Cliffs ^{1}1990; World Scientific, Singapore ^{2}1998).
[32] G. Björk, A. Karlsson, Phys. Rev. A 58 (1998) 3477.
[33] S. Dürr, T. Nonn, G. Rempe, Nature 395 (1998) 33.
[34] S. Dürr, T. Nonn, G. Rempe, Phys. Rev. Lett. 81 (1998) 5705.
[35] P.D.D. Schwindt, P.G. Kwiat, B.-G. Englert, Phys. Rev. A (in print).
[36] P.G. Kwiat, P.D.D. Schwindt, B.-G. Englert, What Does a Quantum Eraser Really Erase? in Mysteries, Puzzles, and Paradoxes in Quantum Mechanics, edited by R. Bonifacio (CP461, The American Institute of Physics, Woodbury, 1999).
[37] U. Eichmann, J.C. Bergquist, J.J. Bollinger, J.M. Gilligan, W.M. Itano, D.J. Wineland, M.G. Raizen, Phys. Rev. Lett. 70 (1993) 2359.
[38] Y. Abranyos, M. Jakob, J.A. Bergou, Acta Phys. Slov. 48 (1998) 255.
[39] Y. Abranyos, M. Jakob, J.A. Bergou, Phys. Rev. A 60 (1999) 2618.
[40] M.O. Scully, B.-G. Englert, H. Walther, Nature (London) 351 (1991) 111.
[41] B.-G. Englert, M.O. Scully, H. Walther, Sci. Am. (Int. Ed.) 271 (1994) 56.
[42] B.-G. Englert, M.O. Scully, H. Walther, Am. J. Phys. 67 (1999) 329.
[43] M.O. Scully, R. Shea, J.M. McCullen, Phys. Rep. 43 (1978) 487.
[44] M.O. Scully, W.E. Lamb, A. Barut, Found. Phys. 17 (1987) 575.

[45] B.-G. Englert, J. Schwinger, M.O. Scully, Found. Phys. 18 (1988) 1045.
[46] J. Schwinger, M.O. Scully, B.-G. Englert, Z. Phys. D 10 (1988) 135.
[47] B.-G. Englert, H. Walther, M.O. Scully, Appl. Phys. B54 (1992) 366.
[48] B.-G. Englert, Spin Coherence in Stern-Gerlach Interferometers in: A.O. Barut (Ed.), New Frontiers in Quantum Electrodynamics and Quantum Optics, NATO ASI Series, Vol. B232, Plenum Press, New York, 1990.
[49] H.-J. Briegel, B.-G. Englert, M.O. Scully, H. Walther, Atom Interferometry and the Quantum Theory of Measurement in: P.R. Berman (Ed.), Atom Interferometry, Academic Press, New York, 1997.
[50] M. Brune, E. Hagley, J. Dreyer, X. Maître, A. Maali, C. Wunderlich, J.M. Raimond, S. Haroche, Phys. Rev. Lett. 77 (1996) 4887.
[51] J.A. Bergou, B.-G. Englert, D.A. Harmin, J. Martinez-Linares (unpublished).