A brief introduction to

Quantum Measurement & Quantum Information

Nicholas Wheeler
21 August 2012

Introduction. I have for a long entertained a smouldering interest in the quantum threory of measurement, and—more particularly—in how the established principles of quantum dynamics (as embodied in, for example, the quantum theory of open systems) might be used to illuminate the physical basis of the idealized propositions in terms of which that theory is conventionally phrased. I have had many occasions to write accounts of the standard (von Neumann) formalism for the benefit of students,[1] but have always been disturbed by the circumstance that the von Neumann formalism achieves its elegant simplicity by neglect of some inescapable aspects of the physical procedures it attempts into address. Thus was I motivated in 1999 to devise a simple theory of "Quantum measurement with imperfect devices."[2]

I am inspired now (in my retirement) to attempt to look more closely to this fundamental subject by two circumstances: (*i*) I am, for the moment, sick of the subject (diverse methods for constructing—and statistical properties of the associated spectra—of random density matrices) that has most recently engaged my attention, and (*ii*) I have happened upon a splendid text[3] in which the author devotes his Chapter 4 ("Generalized measurements") to a richly detailed modern account of the subject to which I have alluded.

Here—though I may borrow also from other sources—my frankly derivative objective will be simply to write out an annotated account of what Barnett has to say.

---

[1] See, for example, "Rudiments of the quantum theory of measurement," pages 8–12 in Chapter 0 of *Advanced Quantum Topics* (2009).

[2] See the notes from the Reed College Physics Seminar of that title that was presented on 16 February 2000.

[3] Stephen M. Barnett, *Quantum Information* (Oxford UP, 2009).

Formal context, and a fussy quibble. We work within the standard formulation of orthodox (non-relativistic) quantum mechanics,[4] wherein the states of a quantum system $\mathcal{S}$ are identified with (described by) complex unit vectors $|\psi)$ that live in a complex inner-product space (Hilbert space) $\mathcal{H}_\mathcal{S}$. For expository convenience, I restrict my explicit attention to n-state systems—systems with n-dimensional state spaces,[5] and will often write $\mathcal{H}_n$ in place of $\mathcal{H}_\mathcal{S}$.

Let $\{|e_k) : k = 1, 2, \dots, n\}$ be some arbitrarily-selected orthonormal basis in $\mathcal{H}_n$:

$$(e_j|e_k) = \delta_{jk} \quad \text{and} \quad \sum_{k=1}^{n} |e_k)(e_k| = \mathbb{I}$$

Given such a basis, one has $|\psi) = \sum |e_k)(e_k|\psi) = \sum |e_k)\psi_k$ where the complex numbers $\psi_k = (e_k|\psi)$ are the "coordinates" of $|\psi)$ with respect to the given basis, and abstract $|\psi)$-vectors acquire column-vector representation:[6]

$$|\psi) \longleftrightarrow \Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{pmatrix}$$

Similarly, if O is a linear operator on $\mathcal{H}_n$ then

$$\mathsf{O} = \sum_j \sum_k |e_j)(e_j|\,\mathsf{O}\,|e_k)(e_k| = \sum_j \sum_k |e_j)O_{jk}(e_k|$$

and O acquires the matrix representation

$$\mathsf{O} \longleftrightarrow \mathbb{O} = \begin{pmatrix} O_{11} & O_{12} & \dots & O_{1n} \\ O_{21} & O_{22} & \dots & O_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ O_{n1} & O_{n2} & \dots & O_{nn} \end{pmatrix}$$

Description of the action of O becomes in representation an exercise in matrix

---

[4] "Standard" entails exclusion of (for example) the phase-space formalism of Wigner, Weyl and Moyal, while "orthodox" entails exclusion of (for example) Robert Griffiths' "consistent quantum theory" (erected on the premise that measurement should be denied a fundamental role), the Bohm formalism, the "many worlds interpretation," *etc.*

[5] Such systems $\mathcal{S}$ are too impoverished to support the commutation relation $\mathsf{x}\,\mathsf{p} - \mathsf{p}\,\mathsf{x} = i\hbar\mathsf{I}$ that underlies much of applied quantum mechanics.

[6] The advantages afforded by Dirac notation are too valuable to give up, so I will frequently allow context to determine whether $|\psi)$ is to be read as an abstract ket vector or its column vector representation with respect to an unstated basis.

albegra:

$$|\psi) \rightarrow |\phi) = \mathbf{O}\,|\psi) \quad \text{represented} \quad \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{pmatrix} \rightarrow \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_n \end{pmatrix} = \mathbb{O} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{pmatrix}$$

It is in the language of matrices, rather than the language of abstract operators, that I will phrase most of my remarks. I will more often speak of density matrices $\rho\!\!\!\rho$ than of density operators $\boldsymbol{\rho}$.

Authors frequently ask us to "Suppose $\mathcal{S}$ is in a mixed state..." That is a supposition to which I take exception, for a reason I digress now to explain:

States of systems *vs* **states of ensembles of systems.** The physical action of quantum measurement devices ("perfect meters") can—in the idealized world contemplated by von Neumann—be represented by the mathematical action of self-adjoint linear operators $\mathbf{A}$, which in reference to an orthonormal basis becomes the action of hermitian matrices $\mathbb{A}$.

Meters are, according to von Neumann, *state-preparation* devices endowed with the special property that they are equipped to *announce the identity of the state they have prepared*. But quantum theory permits one to speak only probabilistically about how the meter will respond in any specific instance. Looking to the solutions of

$$\mathbb{A}|a_k) = a_k|a_k) \quad : \quad k = 1, 2, \ldots, n$$

we know the eigenvalues of $\mathbb{A}$ to be (by hermiticity) necessarily real. Assume for the moment that they are distinct (*i.e.*, that the spectrum $\{a_1, a_2, \ldots, a_n\}$ of $\mathbb{A}$ is non-degenerate). The eigenvectors are then automatically orthogonal, and an be assumed to have been normalized: $(a_j|a_k) = \delta_{jk}$. The claim—the essential upshot of the **von Neumann projection hypothesis**—is that

$$|\psi) \xrightarrow[\mathbf{A}\text{-measurement}]{} \begin{cases} |a_1) \text{ with probability } |(a_1|\psi)|^2 = (a_1|\psi)(\psi|a_1) \\ |a_2) \text{ with probability } |(a_2|\psi)|^2 \\ \quad\vdots \\ |a_k) \text{ with probability } |(a_k|\psi)|^2 \\ \quad\vdots \\ |a_n) \text{ with probability } |(a_n|\psi)|^2 \end{cases}$$

Note that

$$\sum \text{probabilities} = \sum (\psi|a_k)(a_k|\psi) = (\psi|\psi) = 1$$

While quantum theory speaks only probabilistically about the outcome of individual measurements, it speaks with certitude about the mean of many such measurements:

$$\begin{aligned} \text{expected mean } \langle \mathbf{A} \rangle_\psi &= \sum_k a_k |(a_k|\psi)|^2 \\ &= (\psi|\mathbf{A}|\psi) \quad \text{by} \quad \mathbf{A} = \sum_k |a_k)a_k(a_k| \end{aligned}$$

But to perform many such measurements we must possess an *ensemble* $\mathcal{E}(\mathcal{S}_\psi)$ of systems $\mathcal{S}$, each of which has—by *preselection*—been placed in state $|\psi)$. Preselection (or "state preparation") is accomplished by a gated measurement from which the output states are sometimes $|\psi)$ but more typically states $|\psi)_{\text{undesired}}$ that are discarded. Schematic illustration of the preselection process is provided by the following diagram:

$$|\text{unknown state}) \xrightarrow[\mathsf{G}\text{-measurement}]{\hspace{3cm}} \begin{cases} |\psi) \xrightarrow[\text{prompt }\mathsf{A}\text{-measurement}]{\hspace{4cm}} \begin{cases} |a_1) \\ |a_2) \\ \vdots \end{cases} \\ \\ \\ |\psi)_{\text{undesired}} : \text{discarded} \end{cases}$$

Such ensembles $\mathcal{E}(\mathcal{S}_\psi)$ are said to be "pure."

But if either the preparatory G-meter or its associated output filter/gate function imperfectly, or if the systems $\mathcal{S}$ are drawn from (say) a thermalized population, then the ensemble can be expected to present a variety of states:

$$\mathcal{E}(\mathcal{S}_{\{\psi_1,\psi_2,\dots\}}) \text{ presents } \begin{cases} |\psi_1) \text{ with probability } p_1 \\ |\psi_2) \text{ with probability } p_2 \\ \ \vdots \end{cases}$$

A-measurement (performed with a perfect A-meter) can be expected to produce a_k with probability $\sum_\nu p_\nu|(a_k|\psi_\nu))|^2$. The sum of those probabilities is

$$\sum_k \sum_\nu p_\nu(\psi_\nu|a_k)(a_k|\psi_\nu) = \sum_\nu p_n(\psi_\nu|\psi_\nu) = \sum_\nu p_\nu = 1$$

while the expected mean of many such measurements (by nature the ordinary mean of a set of quantum means) becomes

$$\begin{aligned}
\langle \mathsf{A}\rangle_{\mathcal{E}} &= \sum_\nu p_\nu \langle \mathsf{A}\rangle_{\psi_\nu} \\
&= \sum_k \sum_\nu p_\nu(\psi_\nu|a_k)a_k(a_k|\psi_\nu) \\
&= \sum_j \sum_k \sum_\nu p_\nu(\psi_\nu|a_k)a_k(a_k|e_j)(e_j|\psi_\nu) \\
&= \sum_j \sum_k \sum_\nu (e_j|\psi_\nu)p_\nu(\psi_\nu|a_k)a_k(a_k|e_j) \\
&= \text{tr}(\wp_{\mathcal{E}}\,\mathbb{A}) \quad \text{with} \quad \wp_{\mathcal{E}} = \sum_\nu |\psi_\nu)p_\nu(\psi_\nu|
\end{aligned}$$

The point to which I alluded at the end of the preceding section—and now emphasize—is that the density matrix $\wp_{\mathcal{E}}$ refers <u>not to the state of a system $\mathcal{S}$ but to the observationally relevant features of an *ensemble*</u> (in the present instance an "impure" or "mixed" ensemble) of such systems. Ensembles become unmixed or "pure" when one of the p_ν is unity and the others vanish.[7] In such cases one has

$$\langle \mathsf{A} \rangle_\psi = \text{tr}(\mathbb{A}\rho_\psi) \quad \text{with} \quad \rho_\psi = |\psi)(\psi|$$

It would be misleadingly redundant to say of a quantum system $\mathcal{S}$ that "it is in a pure state" (as opposed to what? *all* systems are in "pure"—if possibly unknown—states $|\psi)$). And it would—however tempting—be a potentially misleading use of a preempted word to speak of the "state" of an ensemble.

Properties of density matrices. The defining construction[8]

$$\rho = \sum_\nu |\psi_\nu)p_\nu(\psi_\nu|$$

presents ρ as a real linear combination of hermitian projection matrices $\mathbb{P}_\nu = |\psi_\nu)(\psi_\nu|$. Density matrices ρ are therefore manifestly *hermitian*. And from

$$(\alpha|\rho|\alpha) = \begin{cases} \sum_\nu p_\nu|(\psi_\nu|\alpha)|^2 \geq 0 & : \quad \text{all } |\alpha) \\ \\ 0 \quad \text{iff} \quad \mathbb{P}_\nu|\alpha) = 0 & : \quad \text{all } |\alpha) \end{cases}$$

we see that all such matrices are *positive semi-definite*. Writing

$$\rho|r_k) = r_k|r_k)$$

we infer that the eigenvalues r_k—necessarily real by hermiticity—are all non-negative. Assuming the eigenvectors to have been normalized and the spectrum $\{r_1, r_2, \ldots, r_n\}$ to be non-degenerate, we have

$$\rho = \sum_{k=1}^n |r_k)r_k(r_k|$$

which even in spectrally degenerate cases[9] can be written

$$\rho = \sum_{\kappa=1}^\mu r_\kappa \mathbb{P}_\kappa \quad : \quad \mathbb{P}_\kappa \text{ projects onto } \kappa^{\text{th}} \text{ eigenspace}$$

where the $\mathbb{P}_\kappa$ matrices are

---

[7] The p_ν are positive reals that sum to unity, so this is equivalent to the condition $\sum p_\nu^2 = 1$.

[8] Here I omit the subscript $_\mathcal{E}$ that was attached to emphasize a point that will henceforth be taken for granted.

[9] Let the distinct eigenvalues be denoted r_κ $\{\kappa = 1, 2, \ldots, \mu\}$ and let the multiplicity of r_κ be n_κ, with $\sum n_\kappa = n$.

- hermitian
- projective: $\mathbb{P}_\kappa\mathbb{P}_\kappa = \mathbb{P}_\kappa$
- positive-definite: $(\alpha|\mathbb{P}_\kappa|\alpha) > 0$, all $|\alpha)$
- complete: $\sum \mathbb{P}_\kappa = \mathbb{I}$
- orthogonal: $\mathbb{P}_\kappa\mathbb{P}_\lambda = \delta_{\kappa\lambda}\mathbb{P}_\kappa$

Density matrices ρ manage to become sometimes positive *semi*-definite—even though the projectors $\mathbb{P}_\kappa$ are all positive-*definite*—when one of the r_κ vanishes. In particular, all density matrices that refer to *pure* ensembles ($n \geq 2$) are positive semi-definite:

$$\rho_\psi = 1 \cdot \mathbb{P}_\psi + 0 \cdot (\mathbb{I} - \mathbb{P}_\psi) \quad \text{with} \quad \mathbb{P}_\psi = |\psi)(\psi|$$

All density matrices have *unit trace*

$$\begin{aligned}
\text{tr}\rho &= \sum_{k=1}^{n}(e_k|\sum_\nu |\psi_\nu)p_\nu(\psi_\nu|e_k) \\
&= \sum_\nu \sum_{k=1}^{n} p_\nu(\psi_\nu|e_k)(e_k|\psi_\nu) \\
&= \sum_\nu p_\nu(\psi_\nu|\psi_\nu) = \sum_\nu p_\nu = 1
\end{aligned}$$

from which follows—by

$$\text{tr}\rho = \sum_\kappa r_\kappa \text{tr}\mathbb{P}_\kappa = \sum_\kappa r_\kappa n_\kappa = \sum_k r_k = 1$$

—the equivalent statement that

The eigenvalues of every density matrix ρ sum to unity

It follows moreover that

$$\text{tr}\rho^2 = \sum_k r_k^2 \leq 1, \text{with equality iff } \rho \text{ is pure}$$

We conclude that all density matrices are *positive semi-definite hermitian matrices with unit trace* and, conversely, that all such matrices admit of interpretation as density matrices. The set $\mathcal{R}$ of such $n \times n$ matrices serves to describe the set of all possible ensembles $\mathcal{E}_\mathcal{S}$ of n-state systems $\mathcal{S}$. It is easliy seen that if $\rho_1 \in \mathcal{R}$ and $\rho_2 \in \mathcal{R}$, and if the parameter x ranges on $[0,1]$, then so also are all members of the interpolating set $x\rho_1 + (1-x)\rho_2$ contained within $\mathcal{R}$, which is to say: the set $\mathcal{R}$ is *convex*, and so also therefore is the set of ensembles. The density matrices that refer to pure states live on the boundary of $\mathcal{R}$.

Mixtures of quantum states are unlike mixtures of (say) red and blue balls in that they *do not admit of unique resolution into component parts*. This becomes clear on comparison of

$$\begin{aligned} \rho &= \sum_\nu |\psi_\nu)p_\nu(\psi_\nu| \quad &:& \quad p_\nu\text{-weighted mixture of } |\psi_\nu)\text{-states} \\ &= \sum_{k=1}^n |r_k)r_k(r_k| \quad &:& \quad r_k\text{-weighted mixture of } |r_k)\text{-states} \end{aligned}$$

I have, however, nothing to say about how one "stirs" a quantum mixture to produce all possible equivalent mixtures. It was with non-uniqueness in mind that I was careful earlier to distinguish the observationally relevant features from the "composition" of quantum mixtures: the latter notion makes no objective sense.[10]

Measurement-induced density matrix transformations. Orthodox quantum theory conventionally recognizes a sharp distinction between two modes of state transformation:

- temporal unitary dynamic evolution;
- instantaneous projective measurement.

Processes that achieve state-transformation must also transform ensembles of states, and the density matrices that describe them. Here I will be concerned with transformations of the form

$$\rho_{\text{in}} \xrightarrow[\textbf{A}\text{-measurement}]{} \rho_{\text{out}}$$

where the idealized perfect **A**-meter is represented by an $n\times n$ hermitian matrix, the spectral properties of which are defined/denoted

$$\mathbb{A}|a_k) = a_k|a_k)$$

$$\mathbb{A} = \begin{cases} \sum_k |a_k)a_k(a_k| = \sum_k a_k\mathbb{P}_k & : \quad \text{non-degenerate spectrum} \\ \sum_\kappa a_\kappa \mathbb{P}_\kappa & : \quad \text{degenerate spectrum} \end{cases}$$

As indicated, I use subscripts $_k$ and $_\kappa$ to distinguish non-degenerate from degenerate spectra; $\mathbb{P}_k$ projects onto the k^{th} eigenray, $\mathbb{P}_\kappa$ projects onto the n_κ-dimensional κ^{th} eigenspace.

In the simplest instance the von Neumann projection hypothesis supplies

$$\rho_{\text{in}} = |\psi)(\psi| \xrightarrow[\textbf{A}\text{-meter reads } a_k]{} \rho_{\text{out},k} = \mathbb{P}_k$$

which occurs with probability $(a_k|\psi)(\psi|a_k) = \text{tr}(\rho_{\text{in}}\mathbb{P}_k)$. We observe with

[10] One is reminded in this connection that $|\psi)$ and $e^{i\alpha}|\psi)$ provide equivalent descriptions of the same quantum state; quantum theory attaches great importance to *relative* phase, but assigns no objective meaning to absolute phase. We note also that $\rho = |\psi)(\psi|$ is invariant under $|\psi) \longrightarrow e^{i\alpha}|\psi)$.

Barnett that $\rho_{\text{out},k}$ can in this instance be written

$$\rho_{\text{out},k} = |a_k)(a_k| = \frac{|a_k)(a_k|\psi)(\psi|a_k)(a_k|}{(a_k|\psi)(\psi|a_k)} = \frac{\mathbb{P}_k\,\rho_{\text{in}}\mathbb{P}_k}{\text{tr}(\mathbb{P}_k\,\rho_{\text{in}}\mathbb{P}_k)} \tag{1}$$

since $\text{tr}(\mathbb{P}_k\,\rho_{\text{in}}\mathbb{P}_k) = \text{tr}(\rho_{\text{in}}\mathbb{P}_k) = (a_k|\psi)(\psi|a_k)$. But *if the* A*-meter remains unread* we get

$$\begin{aligned}\rho_{\text{in}} = |\psi)(\psi| \xrightarrow[\mathsf{A}\text{-meter remains unread}]{} \rho_{\text{out}} &= \sum_k (a_k|\psi)(\psi|a_k)\mathbb{P}_k \\ &= \sum_k |a_k)(a_k|\psi)(\psi|a_k)(a_k| \\ &= \sum_k \mathbb{P}_k\,\rho_{\text{in}}\mathbb{P}_k \end{aligned} \tag{2}$$

Barnett remarks that (2) is "one of the reasons why quantum key distribution works," and that the distinction between (1) and (2) "highlights the signficance of information in quantum theory; the two density matrices are different because in the former case we know something extra (the measurement outcome); the state we assign to the post-measurement system depends upon the amount of information available to us."[11]

If, on the other hand, the A-spectrum is degenerate we have

$$|\psi)_{\text{in}} \xrightarrow[\mathsf{A}\text{-meter reads } a_\kappa]{} |\psi)_{\text{out},\kappa} = \frac{\mathbb{P}_\kappa|\psi)_{\text{in}}}{\sqrt{{}_{\text{in}}(\psi|\mathbb{P}_\kappa\mathbb{P}_\kappa|\psi)_{\text{in}}}}$$

which in density matrix language becomes

$$\rho_{\text{in}} = |\psi)(\psi| \xrightarrow[\mathsf{A}\text{-meter reads } a_\kappa]{} \rho_{\text{out},\kappa} = \frac{\mathbb{P}_\kappa\,\rho_{\text{in}}\mathbb{P}_\kappa}{\text{tr}(\mathbb{P}_\kappa\,\rho_{\text{in}}\mathbb{P}_\kappa)} \tag{3}$$

and occurs with probability $\text{tr}(\mathbb{P}_\kappa\,\rho_{\text{in}}\mathbb{P}_\kappa)$. It follows from this result that

$$\rho_{\text{in}} = |\psi)(\psi| \xrightarrow[\mathsf{A}\text{-meter remains unread}]{} \rho_{\text{out}} = \sum_\kappa \mathbb{P}_\kappa\,\rho_{\text{in}}\mathbb{P}_\kappa \tag{4}$$

These statements give back (1) and (2) in the absence of degeneracy.

Working from (3) we find that if $\rho_{\text{out},\kappa}$ is promptly presented to a second A-meter, the second meter (use $\mathbb{P}_\kappa\mathbb{P}_\lambda = \delta_{\kappa\lambda}\mathbb{P}_\kappa$) reads

$$\begin{aligned} a_\kappa \text{ with probability } \text{tr}\left(\mathbb{P}_\kappa\left[\frac{\mathbb{P}_\kappa\,\rho_{\text{in}}\mathbb{P}_\kappa}{\text{tr}(\mathbb{P}_\kappa\,\rho_{\text{in}}\mathbb{P}_\kappa)}\right]\mathbb{P}_\kappa\right) &= 1 \\ a_{\lambda\neq\kappa} \text{ with probability } \text{tr}\left(\mathbb{P}_\lambda\left[\frac{\mathbb{P}_\kappa\,\rho_{\text{in}}\mathbb{P}_\kappa}{\text{tr}(\mathbb{P}_\kappa\,\rho_{\text{in}}\mathbb{P}_\kappa)}\right]\mathbb{P}_\lambda\right) &= 0 \end{aligned}$$

The second meter simply confirms the result reported by the first meter, which is gratifying, since if confirmation were impossible it would be difficult to argue that the first meter had taught us anything.

[11] I take exception to Barnett's use here of the words "state" and "system."

Using (3) and (4) in combination we find that if the first meter remains unread and the second announces a_κ then the output from the second meter

$$\rho_{\text{out;out},\kappa} = \frac{\mathbb{P}_\kappa \sum_\lambda \mathbb{P}_\lambda \rho_{\text{in}} \mathbb{P}_\lambda \mathbb{P}_\kappa}{\text{tr}(\mathbb{P}_\kappa \sum_\lambda \mathbb{P}_\lambda \rho_{\text{in}} \mathbb{P}_\lambda \mathbb{P}_\kappa)} = \frac{\mathbb{P}_\kappa \rho_{\text{in}} \mathbb{P}_\kappa}{\text{tr}(\mathbb{P}_\kappa \rho_{\text{in}} \mathbb{P}_\kappa)}$$

is precisely what it would have been had the first meter been omitted. And if both meters remain unread it follows similarly that either meter could be deleted without affecting the final result.

It is evident that (3) and (4) retain their validity even when the assumption that ρ_{in} refers to a pure ensemble is abandoned. And from both descriptions of ρ_{out} it follows readily that—as required—

$$\text{tr}\rho_{\text{out}} = \text{tr}\rho_{\text{in}} = 1$$

From (4) it follows moreover (use orthogonality and completeness of the $\mathbb{P}_\kappa$) that

$$\text{tr}\,\rho^2_{\text{out}} = \sum_{\kappa,\lambda} \text{tr}\big(\mathbb{P}_\kappa \rho_{\text{in}} \mathbb{P}_\kappa \rho_{\text{in}} \mathbb{P}_\lambda\big) = \sum_\kappa \text{tr}\,\rho^2_{\text{in}}$$

so *unread meters do not alter the purity/impurity* of ensembles presented to them. This differs profoundly from the situation when the meter is read and found to announce a_κ, for we then have

$$\text{tr}\,\rho^2_{\text{out},\kappa} = \frac{\text{tr}(\mathbb{Q}\cdot\mathbb{Q})}{\text{tr}\,\mathbb{Q}\cdot\text{tr}\mathbb{Q}} \quad \text{with} \quad \mathbb{Q} = \mathbb{P}_\kappa \rho_{\text{in}} \mathbb{P}_\kappa$$

$\mathbb{P}_\kappa$ and ρ_{in} are both semi-positive hermitian, and it is known that products of semi-positive hermitian matrices are semi-positive hermitian, so $\mathbb{Q}$ is. It is known also[12] that if $\mathbb{A}$ and $\mathbb{B}$ are semi-positive hermitian then

$$0 \le \text{tr}\,\mathbb{A}\mathbb{B} \le \text{tr}\,\mathbb{A}\cdot\text{tr}\,\mathbb{B}$$

It follows that—as is required on general grounds—

$$0 \le \text{tr}\,\rho^2_{\text{out},\kappa} \le 1$$

Lacking theorems that would permit me to discuss the more interesting question of how $\text{tr}\,\rho^2_{\text{out},\kappa}$ relates to $\text{tr}\,\rho^2_{\text{in}}$ I proceeded experimentally: Let $\mathbb{M}$ be a 4×4 matrix with random complex elements the real/imaginary parts of which range independently but uniformly on the interval $[-1,+1]$. Construct $\mathbb{W} = \mathbb{M}^+\mathbb{M}$

[12] I was led to this result by *Mathematica*-based numerical experimentation, but am informed by quick web search that a proof can be found on page 269 of E. H. Lieb & W. Thirring, *Studies in Mathematical Physics: Essays in Honor of Valetine Bargmann* (Princeton, 1976). See also K. M. Abadin & J. R. Magnus, *Matrix Algebra* (Cambridge, 2005), pages 329 & 338.

and from $\mathbb{W}$ construct

$$\rho_{\rm in} = \mathbb{W}/\mathrm{tr}\mathbb{W}$$

Construct a large population of such random density matrices. Assume $\mathbb{A}$ to have been diagonalized and the eigenvalue a_1 to be non-degenerate. The spectral representation of $\mathbb{A}$ will have then the form

$$\mathbb{A} = a_1\begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix} + \text{orthogonal diagonal terms}$$

while if a_1 is double/triply degenerate we have

$$\mathbb{A} = a_1\begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix} + \text{orthogonal diagonal terms}$$

$$\mathbb{A} = a_1\begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0\end{pmatrix} + \text{an orthogonal diagonal term}$$

Let those projection matrices be denoted $\mathbb{P}_1$, $\mathbb{P}_2$, $\mathbb{P}_3$, respectively. Setting

$$\mathbb{Q} = \mathbb{P}_\kappa\,\rho_{\rm in}\mathbb{P}_\kappa \quad \text{with}\begin{cases} \kappa = 1 & : \quad \text{non-degenerate case}\\ \kappa = 2 & : \quad \text{doubly degenerate case}\\ \kappa = 3 & : \quad \text{trebly degenerate case}\end{cases}$$

we first ask *Mathematica* whether it is `True` that

$$\mathrm{tr}\,\rho^2_{\mathrm{out},\kappa} = \frac{\mathrm{tr}(\mathbb{Q}\cdot\mathbb{Q})}{\mathrm{tr}\,\mathbb{Q}\cdot\mathrm{tr}\mathbb{Q}} = \mathrm{tr}\,\rho^2_{\rm in}$$

and are informed that it is *never* true that $\mathrm{tr}\,\rho^2_{\mathrm{out},\kappa} = \mathrm{tr}\,\rho^2_{\rm in}$ (yet invariably true when we trivialize the meter by setting $\mathbb{A} = \mathbb{I}$). In short: interaction with a meter—*if the meter reading is recorded*—invariably serves to alter the structure of the mixture. We can quantify that broad assertion by looking to the mean values of the data upon which it is based:

$$\begin{aligned} &\big\langle \mathrm{tr}\,\rho^2_{\rm in}\big\rangle &&= 0.76 \quad : \quad \text{variance} = 0.0054\\ &\big\langle \mathrm{tr}\,\rho^2_{\mathrm{out},\kappa}\big\rangle_{\text{treble degeneracy}} &&= 0.79\\ &\big\langle \mathrm{tr}\,\rho^2_{\mathrm{out},\kappa}\big\rangle_{\text{double degeneracy}} &&= 0.84\\ &\big\langle \mathrm{tr}\,\rho^2_{\mathrm{out},k}\big\rangle_{\text{non-degenerate}} &&= 1.00 \end{aligned}$$

Evidently the purity of the output ensemble tends to be increased as the degeneracy of the observed eigenvalue is reduced (*i.e.*, as the meter reading becomes more sharply informative). But individual measurements may violate that trend; when we ask *Mathematica* whether it is `True` that

$$\mathrm{tr}\,\rho^2_{\rm out} = \frac{\mathrm{tr}(\mathbb{Q}\cdot\mathbb{Q})}{\mathrm{tr}\,\mathbb{Q}\cdot\mathrm{tr}\mathbb{Q}} > \mathrm{tr}\,\rho^2_{\rm in}$$

we get affirmative responses
in $\approx 68\%$ of trebly degenerate cases,
in $\approx 85\%$ of doubly degenerate cases,
but in 100% of non-degenerate cases.

To summarize: while interaction with a meter tends to increase the purity of an ensemble, exceptions to that general tendency become ever more likely as the degeneracy of the reported eigenvalue increases.

Imperfect meters. The action of an ideal **A**-meter[13] can be diagramed (see again page 7)

$$\rho_{\text{in}} = |\psi)(\psi| \xrightarrow[\textsf{A}\text{-meter reads } a_k]{} \rho_{\text{out},k} = \mathbb{P}_k \equiv |a_k)(a_k|$$

If the meter is "imperfect" (or "non-ideal") we on the other hand have

$$\rho_{\text{in}} = |\psi)(\psi| \xrightarrow[\textsf{A}\text{-meter reads } a_k]{} \begin{cases} \dots \\ \mathbb{P}_{k-1} & \text{but reads } a_k \text{ with cp } w_{k|k-1} \\ \mathbb{P}_k & \text{but reads } a_k \text{ with cp } w_{k|k} \\ \mathbb{P}_{k+1} & \text{but reads } a_k \text{ with cp } w_{k|k+1} \\ \dots \end{cases}$$

where "cp $w_{k|j}$" signifies "conditional probability of k, given j." In short: ideal meters—*upon announcement of the meter reading*—produce pure ensembles, while <u>imperfect meters produce mixtures</u>:

$$\rho_{\text{out},k} = \sum_j w_{k|j}\mathbb{P}_j \quad : \quad \sum_k w_{k|j} = 1 \text{ (all } j) \tag{5}$$

Observe that $\text{tr}\,\rho_{\text{out},k} = \sum_j w_{k|j}\text{tr}\mathbb{P}_j = 1$; also $\text{tr}\,\rho^2_{\text{out},k} = \sum_j w_{k|j}^{\ 2} < 1$ unless the meter is in fact ideal. We expect "good imperfect meters" to be "fuzzy" but not to be flagrant liers; *i.e.*, we expect $\max(w_{k|j}) = w_{k|k}$.

In "Quantum measurement with imperfect devices"[2]—which was written in shameful ignorance of the relevant literature, as a hasty supplement to some class notes—I use the simplest of means to trace out some of the implications of the elementary notion just sketched. Barnett devotes his pages 92–111 to a more sophisticated discussion of its formal/physical ramifications, and it is from Barnett that I borrow the central thread of what now follows.

Present ρ_{in} to an imperfect **A**-meter. From (5) we infer that the probability that the meter will register a_k is

$$\begin{aligned} \text{prob}(\rho_{\text{in}}, a_k) &= \sum_j w_{k|j}\text{tr}\big(\rho_{\text{in}}\mathbb{P}_j\big) \\ &= \text{tr}\big(\rho_{\text{in}}\tilde{\mathbb{P}}_k\big) \quad \text{where} \quad \tilde{\mathbb{P}}_k = \sum_j w_{k|j}\mathbb{P}_j \end{aligned} \tag{6}$$

---

[13] It is convenient for the purposes of the present discussion to assume that the spectrum of $\mathbb{A}$ is non-degenerate, an assumption which we will ultimately find easy to relax.

The imperfect meter is described now not by a hermitian matrix $\mathbb{A}$ but by the *set* $\mathcal{A} = \{\tilde{\mathbb{P}}_k : k = 1, 2, \ldots, n\}$ of matrices which collectively comprise a POVM, a "**positive operator-valued measure**."[14] The elements of $\mathcal{A}$ are obviously *hermitian* and *positive-definite*. They are, moveover, *complete*:

$$\sum_k \tilde{\mathbb{P}}_k = \sum_k \sum_j w_{k|j} \mathbb{P}_j = \sum_j \mathbb{P}_j = \mathbb{I}$$

The orthogonal projectivity of the $\mathbb{P}$-matrices

$$\mathbb{P}_j \mathbb{P}_k = \delta_{jk} \mathbb{P}_j$$

has, however, been lost:

$$\begin{aligned}
\tilde{\mathbb{P}}_j \tilde{\mathbb{P}}_k &= \sum_{a,b} w_{j|a} w_{k|b} \mathbb{P}_a \mathbb{P}_b \\
&= \sum_{a,b} w_{j|a} w_{k|b} \delta_{ab} \mathbb{P}_a \\
&= \sum_a w_{j|a} w_{k|a} \mathbb{P}_a \\
&\Downarrow \\
\tilde{\mathbb{P}}_k \tilde{\mathbb{P}}_k &= \sum_a w_{k|a}^2 \mathbb{P}_a \\
&= \tilde{\mathbb{P}}_k \quad \text{if and only if} \quad w_{k|a}^2 = w_{k|a} \quad (\text{all } a)
\end{aligned} \tag{7}$$

Just as the set of all ideal **A**-meters can be identified with the set of all suitably-dimensioned hermitian matrices $\mathbb{A}$ (each of which can be thought of as a projector-set $A = \{\mathbb{P}_1, \mathbb{P}_2, \ldots, \mathbb{P}_{\nu \leq n}\}$ to each element of which a "meter mark" (eigenvalue) has been attached), so can the set of non-ideal meters be identified with the set of all similarly-dimensioned "marked POVM"s

$$\begin{array}{rcl}
\text{imperfect } \mathbf{A}\text{-meter} & \longleftrightarrow & \mathcal{A} = \{\tilde{\mathbb{P}}_1, \tilde{\mathbb{P}}_2, \ldots\} \\
& & \qquad\ \ \updownarrow \quad \updownarrow \\
& & \qquad\ \ a_1 \quad a_2
\end{array} \tag{8}$$

We saw at (7) the nature of the highly restrictive circumstance that must prevail if the $\tilde{\mathbb{P}}$-matrices are to be projective, and ideal meters to emerge as special instances of non-ideal meters.

It is important to notice that while the number ν of projectors $\mathbb{P}_\kappa \in A$ is dimension-limited,[15] the number of $\tilde{\mathbb{P}}_\kappa \in \mathcal{A}$ is—because the projectivity requirement has been relaxed—unlimited.

The association (8) provides a fair summary of the argument which led us from a tentative sense of "how imperfect meters work" to invention of the POVM concept. But it remains to be established that "most general quantum

[14] John Preskill[16] remarks that "The term measure is a bit heavy-handed in our finite-dimensional context; it becomes more apt [when the dimension becomes infinite]."

[15] We have $\nu = n$ if and only if the $\mathbb{A}$-spectrum is non-degenerate.

measurements" can be accomplished by imperfect meters, and fall within the rubric of the POVM formalism. To that end, Barnett sketches a *model* of the "most general quantum measurement.

Model of the quantum measurement process. The quantum system $\mathcal{S}$ under study is initially in the unknown state $|\psi) \in \mathcal{H}_{\mathcal{S}}$. The meter—also a quantum system $\mathcal{M}$ (traditionally called the "ancilla" by writers in this field) is initially in the known state $|\alpha) \in \mathcal{H}_{\mathcal{M}}$. The initial state $|\psi) \otimes |\alpha)$ of the composite system lives in $\mathcal{H} = \mathcal{H}_{\mathcal{S}} \otimes \mathcal{H}_{\mathcal{M}}$, wherein

$$\{|e_i) \otimes |f_j)\} \quad : \quad \begin{cases} i = 1, 2, \ldots, n \\ j = 1, 2, \ldots, m \end{cases}$$

comprises an orthonormal basis. Brief dynamical system-meter interaction sends

$$\big[|\psi) \otimes |\alpha)\big]_{\text{unentangled}} \longrightarrow \mathbb{U}\big[|\psi) \otimes |\alpha)\big]_{\text{entangled}}$$

where $\mathbb{U}$ is a presumably known $mn \times mn$ unitary matrix. The probability that von Neuman measurement (henceforth called a "projection-valued measurement, or PVM) will show the composite system to be in state $|e_i) \otimes |f_j)$ is

$$\text{Prob}_{ij} = \Big|\big[(e_i| \otimes (f_j|\big]\mathbb{U}\big[|\psi) \otimes |\alpha)\big]\Big|^2$$

The situation is clarified by notational adjustment: write $|\psi) = \sum_k |e_k)\psi_k$ and introduce mn-dimensional vectors

$$|E_{ij}) = |e_i) \otimes |f_j) \quad \text{and} \quad |A_k) = |e_k) \otimes |\alpha)$$

Then

$$\text{Prob}_{ij} = \Big|\textstyle\sum_k (E_{ij}|\mathbb{U}|A_k)\,\psi_k\Big|^2 \tag{9}$$

Now introduce the ij-indexed n-dimensional bra vectors

$$(\pi_{ij}| = \big(\,(E_{ij}|\mathbb{U}|A_1) \quad (E_{ij}|\mathbb{U}|A_2) \quad \ldots \quad (E_{ij}|\mathbb{U}|A_n)\,\big) \tag{10}$$

and obtain

$$\begin{aligned} \text{Prob}_{ij} &= \big|(\pi_{ij}|\psi)\big|^2 \\ &= (\psi|\tilde{\mathbb{P}}_{ij}|\psi) \quad \text{with} \quad \tilde{\mathbb{P}}_{ij} = |\pi_{ij})(\pi_{ij}| \end{aligned} \tag{11}$$

Compare this result with (9), which can be written

$$\text{Prob}_{ij} = (A|\mathbb{Q}_{ij}|A) \tag{12}$$

where $|A) = |\psi) \otimes |\alpha)$ and the $mn \times mn$ matrix $\mathbb{Q}_{ij} = \mathbb{U}^{+}|E_{ij})(E_{ij}|\mathbb{U}$ projects onto the entangled state $\mathbb{U}^{+}|E_{ij})$.

The $n \times n$ matrices $\tilde{\mathbb{P}}_{ij}$, which are mn in number, are clearly hermitian. That they are at least positive *semi*-definite (but not necessarily positive definite) follows from the circumstance that Prob_{ij}—though it may vanish—cannot be negative. And from

$$\sum_{ij} |E_{ij})(E_{ij}| = \Big(\textstyle\sum_i |e_i)(e_i|\Big) \otimes \Big(\textstyle\sum_j |f_j)(f_j|\Big) = \mathbb{I}_n \otimes \mathbb{I}_m = \mathbb{I}_{mn}$$

it follows (essentially from the completeness of the $\{|e_i)\}$ and $\{|f_j)\}$ bases) that the $\tilde{\mathbb{P}}_{ij}$-matrices are complete:

$$\Big[\sum_{ij} \tilde{\mathbb{P}}_{ij}\Big]_{pq} = \Big[(A|\mathbb{U}^{+}\mathbb{I}_{mn}\mathbb{U}|A)\Big]_{pq} = (e_p|e_q) \otimes (\alpha|\alpha) = \mathbb{I}_n$$

Barnett would have us believe that the strict positivity required if the $\tilde{\mathbb{P}}_{ij}$ are to comprise a POVM was established at (11), but that cannot be the case. For look to the case $\mathbb{U} = \mathbb{I}_{mn}$; (12) then supplies $\text{Prob}_{ij} = (A|E_{ij})(E_{ij}|A)$ and it is entirely possible that $(E_{ij}|A) = (e_i|\psi) \cdot (f_j|\alpha)$ may vanish. Evidently strict positivity imposes a difficult-to-describe constraint on the unitary matrix $\mathbb{U}$ that serves to entangle the initially unentangled states of system and meter.

Equation (11) looks superficially like the description of the expected result of a projective measurement (POM). It fails to be so because the n-vector $|\pi_{ij})$ is not a unit vector ($\tilde{\mathbb{P}}_{ij}$ is not projective, does not have unit trace). Generally

$$(\pi_{ij}|\pi_{ij}) = \big[(\psi| \otimes (\alpha|\big]\mathbb{U}^{+}\big[|e_i) \otimes |f_j)\big]\big[(e_i| \otimes (f_j|\big]\mathbb{U}\big[|\psi) \otimes |\alpha)\big]$$

where $|\psi) = \sum_q |e_q)\psi_q$. To demonstrate that $(\pi_{ij}|\pi_{ij}) \neq 1$ it is sufficient to look to the trivial case $\mathbb{U} = \mathbb{I}_{mn}$, where we have

$$= \bar{\psi}_i(\alpha|f_j)(f_j|\alpha)\psi_i$$

which equals one only under circumstances so special that if satisfied for some specified values of i and j *cannot* be satisfied for any other values.

The preceding discussion serves to demonstrate how it comes about that PVM measurements on $\mathcal{H}_{\mathcal{S}} \otimes \mathcal{H}_{\mathcal{M}}$ come to be realized as POVM measurements on $\mathcal{H}_{\mathcal{S}}$. John Preskill[16] elects to "follow a somewhat different procedure that, while not as well motivated physically, is simpler and more natural from a mathematical point of view." By working not in the mn-dimensional space $\mathcal{H}_{\mathcal{S}} \otimes \mathcal{H}_{\mathcal{M}}$ but in the $(m+n)$-dimensional space $\mathcal{H}_{\mathcal{S}} \oplus \mathcal{H}_{\mathcal{M}}$ (with $\mathcal{H}_{\mathcal{S}} \perp \mathcal{H}_{\mathcal{M}}$) he manages to avoid the notational and other complexities latent in the Kronecker

[16] John Preskill is the Feynman Professor of Theoretical Physics at Caltech. His "Lecture Notes for Physics 229: Quantum Information & Computation"—prepared in 1997-98 and available at http://www.theory.caltech.edu/people/preskill/ph229—are a widely quoted source, and (though not cited by him) pretty clearly influenced Barnett. The present topic is developed in Preskill's Chapter 3.

product. Barnett's line of argument relates in a more natural way to what one might mean by a "quantum dynamical theory of quantum measurement," but is—as it stands—very much less than such a theory, for he has nothing to say about the construction of the Hamiltonian $\mathbb{H}$ that generates the meter-system interaction $\mathbb{U}$. Nor has he anything to say about how—physically—one is to *perform* a PVM on a composite system. The relevant hermitian matrix

$$\mathbb{Z} = \sum_{ij} w_{ij}\mathbb{Q}_{ij}$$

is structurally quite unlike the meters $\mathbb{A}\otimes\mathbb{I}_n$ and $\mathbb{I}_m\otimes\mathbb{B}$ employed by Alice and Bob when they examine their respective components of an entangled composite system.

POVM-induced density matrix transformations. The issue before us

$$\rho_{\text{in}} \xrightarrow[\text{imperfect } \mathsf{A}\text{-meter reads } a_k]{} \rho_{\text{out,k}}$$

wa addressed in the old material previously cited,[2] but we are now in position to appoach thte issue in quite another way, which appears to be in several respects much more elegantly efficient and useful.

Barnett (who takes up this issue in his §4.5) by asking this general question: "What is the most general way in which we can change a density matrix?" He argues on linearity grounds—quantum mechanics being an exercise in linear algebra—that it must have the form $\rho \longrightarrow \mathbb{A}\rho\mathbb{B}$, where hermiticity-preservation forces $\mathbb{A} = \mathbb{B}^{+}$. Positivity-preservation is then automatic: $(\psi|\rho|\psi) \geq 0$ (all $|\psi)$) implies $(\phi|\mathbb{B}^{+}\rho\mathbb{B}|\phi) \geq 0$ (all $|\phi)$). More generally, one might have

$$\rho \longrightarrow \sum_i \mathbb{B}_i^{+}\rho\mathbb{B}_i \tag{13.1}$$

Then $\text{tr}\rho \longrightarrow \text{tr}\Big[\big(\sum_i \mathbb{B}_i\mathbb{B}_i^{+}\big)\rho\Big]$ and unit-trace-preservation (all ρ) is seen to require

$$\sum_i \mathbb{B}_i\mathbb{B}_i^{+} = \mathbb{I} \tag{13.2}$$

The (generally non-projective) hermitian matrices $\mathbb{W}_i = \mathbb{B}_i\mathbb{B}_i^{+}$ are positive because of their Wishart structure. And since they sum to the identity $\sum_i \mathbb{W}_i = \mathbb{I}$ the set $\{\mathbb{W}_i\}$ possesses all the defining properties of a POVM.

Transformations of the form general (13) are called "operations," and the study of their ramifications is called "operational quantum theory." The matrices $\mathbb{B}_i$ are matrix representations of "Kraus operators."

I have recently had occasion to remark[17] that, by the Schur decomposition theorem, *any* real or complex square matrix $\mathbb{W}$ can be rendered

$$\mathbb{W} = \mathbb{U}\,\mathbb{T}\,\mathbb{U}^{-1}$$

where $\mathbb{U}$ is orthogonal (unitary or rotational: $\mathbb{U}^{-1} = \mathbb{U}^{+}$) and $\mathbb{T}$—the "Schur

[17] "Populations of random density matrices," (August 2012).

form" of $\mathbb{W}$—is upper triangular:

$$\mathbb{T} = \begin{pmatrix} \bullet & \bullet & \bullet & \cdots & \bullet \\ 0 & \bullet & \bullet & \cdots & \bullet \\ 0 & 0 & \bullet & \cdots & \bullet \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \bullet \end{pmatrix}$$

Since $\mathbb{W}$ and $\mathbb{T}$ are similar they have identical spectra, and since $\mathbb{T}$ is triangular its eigenvalues are precisely the numbers that appear on its principal diagonal. When $\mathbb{W}$ is hermitian the off-diagonal elements of $\mathbb{T}$ vanish, and the Schur decomposition of $\mathbb{W}$ assumes the form

$$\mathbb{W} = \mathbb{U}\mathbb{D}\mathbb{U}^{+} \quad \text{with} \quad \mathbb{D} = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

which is readily seen[17] to amount simply to a formulation of the spectral decomposition of $\mathbb{W}$. If the hermitian matrix $\mathbb{W}$ is positive then $\sqrt{\mathbb{D}}$ (all roots $\sqrt{\lambda_k}$ real, and taken to be positive) is a well-defined positive hermitian matrix, and we have

$$\mathbb{W} = \mathbb{B}\mathbb{B}^{+} \quad \text{with} \quad \mathbb{B} = \mathbb{U}\sqrt{\mathbb{D}}$$

We have established that *every positive hermitian matrix can be rendered as a Wishart product*. Note, however, that the Wishart factors of $\mathbb{W}$ are *not unique*:

$$\mathbb{W} = \mathbb{B}\mathbb{B}^{+} = \mathbb{A}\mathbb{A}^{+} \quad \text{if} \quad \mathbb{A} = \mathbb{B}\mathbb{V} \;:\; \mathbb{V} \text{ unitary}$$

So much for generalities, which have been seen to acquire spontaneously a distinctly POVM odor.

Suppose now that the elements of the POVM $\{\tilde{\mathbb{P}}_1, \tilde{\mathbb{P}}_2, \ldots, \tilde{\mathbb{P}}_\nu\}$ that describes the action of an imperfect $\mathbf{A}$-meter have been presented in factored form

$$\tilde{\mathbb{P}}_j = \mathbb{A}_j\mathbb{A}_j^{+}$$

and that the associated meter readings are $\{a_1, a_2 \ldots, a_\nu\}$. Barnett claims that

$$\rho_{\text{in}} \xrightarrow[\text{imperfect } \mathbf{A}\text{-meter reads } a_\kappa]{} \rho_{\text{out},\kappa} = \frac{\mathbb{A}_\kappa^{+}\rho_{\text{in}}\mathbb{A}_\kappa}{\text{tr}(\mathbb{A}_\kappa^{+}\rho_{\text{in}}\mathbb{A}_\kappa)} \tag{14}$$

occurs with probability $\text{tr}(\mathbb{A}_\kappa^{+}\rho_{\text{in}}\mathbb{A}_\kappa)$ and entails

$$\rho_{\text{in}} \xrightarrow[\text{imperfect } \mathbf{A}\text{-meter remains unread}]{} \rho_{\text{out}} = \sum_\kappa \mathbb{A}_\kappa^{+}\rho_{\text{in}}\mathbb{A}_\kappa \tag{15}$$

The claim is supported by the observations that $\text{tr}\rho_{\text{out},\kappa} = 1$ (trivially) and $\text{tr}\rho_{\text{out}} = 1$ (by $\sum_\kappa \mathbb{A}_\kappa\mathbb{A}_\kappa^{+} = \mathbb{I}$), but rests mainly that circumstance that (trivially) every projection matrix can be written $\mathbb{P} = \mathbb{A}\mathbb{A}^{+}$ with $\mathbb{A} = \mathbb{A}^{+} = \mathbb{P}$, so (3) and (4)—which are conceptually secure—can be recovered as special instances of (14) and (15).

There is, however, a problem: the probabilities $\text{tr}(\mathbb{A}_\kappa^+ \rho_{\text{in}} \mathbb{A}_\kappa)$ are invariant under $\mathbb{A} \longrightarrow \mathbb{A}\mathbb{V}$ ($\mathbb{V}$ unitary) but the matrices that stand on the right sides of (13) and (14) are not. Evidently the characterization of an imperfect **A**-meter resides not in the elements $\mathbb{P}_\kappa$ of a POVM but in their explicit "Kraus factors" $\mathbb{A}_\kappa$. Note in this connection that the POVM constructed at (6)

$$\tilde{\mathbb{P}}_k = \sum_j w_{k|j} \mathbb{P}_j$$

is *not* presented in factored form, while the one constructed at (11)

$$\tilde{\mathbb{P}}_{ij} = |\pi_{ij})(\pi_{ij}|$$

is explicitly factored.

Neumark's dilation theorem: from POVM to PVM. Barnett's model of the quantum measurement process served to illustrate how PVMs in $\mathcal{H}_\mathcal{S} \otimes \mathcal{H}_\mathcal{M}$ acquire the character of POVMs in $\mathcal{H}_\mathcal{S}$. "Neumark's dilation theorem"[18] proceeds in the opposite direction: it established that every PVOM in $\mathcal{H}$ can be realized as a PVM in a Hilbert space of higher dimension.

Though Neumark's paper was only three pages long, the Wikipedia article presumes command of a lot of fairly abstruse mathematics, which, I suppose, is why Barnett is content to illustrate how Newmark's theorem pertains to a single illustrative example.[19] Let $\mathcal{S}$ be a qubit (let $\mathcal{H}_\mathcal{S}$ be 2-dimensional, denoted $\mathcal{H}_2$). To describe an ideal (PVM) **A**-meter in such a setting one has simply to present a 2×2 hermitian matrix. . . which might be done in this uncommon way: let $|A_1)$ and $|A_2)$ be a pair of unnormalized but orthogonal vectors, and write

$$\mathbb{A} = |A_1)(A_1| + |A_2)(A_2|$$

When we set $|A_k) = \sqrt{a_k}\,|a_k)$ with $a_k = (A_K|A_k)$ we recover the familiar spectral decomposition $\mathbb{A} = |1_0)a_1(a_1| + |a_2)a_2(a_2|$. Notice that for ideal meters the number of "meter marks" $\{a_1, a_2\}$ is in this simple context equal to the dimension of $\mathcal{H}$.[20] The number of marks displayed on the dial of a imperfect

[18] A. Neumark, "On a representation of additive operator set functions," Acad.Sci. USRR **41**, 359-361 (1943). The Neumark dilation theorem can be obtained as a consequence of the "Stinespring dilation/factorization theorem": W. F. Stinespring, "Positive functions on C^* algebras," Proc. Amer. Math. Soc. **6**, 211-216 (1955). A standard source for information about such matters is V. Paulsen, *Completely Bounded Maps and Operator Algebras* (2003).

[19] Preskill presents a one-page proof, which he illustrates with an example similar to Barnett's. Mario Flory's lecture notes from *Foundations of Quantum Mechanics*, a course presented at the Arnold Sommerfeld Center for Theoretical Physics, Ludwig-Maximilians-Universität München (2010) are also helpful. See the chapter "POVMs and superoperators."

[20] If the dimension of $\mathcal{H}_\mathcal{S}$ is $N > 2$ the number of distinct marks may be less than N, which simply signals the presence of degeneracy. Degeneracy trivializes the case $N = 2$.

A-meter can, however—and typically will—exceed the dimension of $\mathcal{H}$. Such meters are described by manifestly positive-definite hermitian matrices

$$\begin{aligned}\tilde{\mathbb{A}} &= |A_1)(A_1| \;\; + |A_2)(A_2| \;\; + |A_3)(A_3| \;\; + \cdots + |A_n)(A_n| \\ &= |a_1)a_1(a_1| + |a_2)a_2(a_2| + |a_3)a_3(a_3| + \cdots + |a_n)a_n(a_n|\end{aligned}$$

where the non-orthogonal unit 2-vectors $\{|a_1), |a_2), \dots, |a_n)\}$ are subject to the requirement

$$|a_1)(a_1| + |a_2)(a_2| + |a_3)(a_3| + \cdots + |a_n)(a_n| = \mathbb{I}_2$$

which by $|a_k) = \begin{pmatrix} a_{k1} \\ a_{k2} \end{pmatrix}$ can be written

$$\sum_{k=1}^{n} a_{ki}\bar{a}_{kj} = \delta_{ij} \quad : \quad i,j \text{ range on } \{1,2\}$$

Look upon $\mathcal{H}_2$ as a subspace of $\mathcal{H}_n$, and let $\{|k) : k = 1,2,\dots,n\}$ be an orthonormal basis in $\mathcal{H}_n$. Construct a pair of n-vectors

$$|b_1) = \sum_k a_{k1}|k), \quad |b_2) = \sum_k a_{k2}|k)$$

that embody all the information written into the 2-vectors $\{|a_1), |a_2), \dots, |a_n)\}$ and observe that those vectors are orthonormal:

$$(b_i|b_j) = \sum_{k,l=1}^{n} \bar{a}_{ki}a_{lj}(k|l) = \sum_{k=1}^{n} \bar{a}_{ki}a_{kj} = \delta_{ij} \quad : \quad i,j \text{ range on } \{1,2\}$$

Consider $|b_1)$ and $|b_2)$ to be leading elements of a complete orthonormal basis $\{|b_1), |b_2), |b_3), \dots, |b_n)\}$ in $\mathcal{H}_n$,[21] and write

$$|b_q) = \sum_k a_{kq}|k) \quad : \quad q = 3,4,\dots,n$$

to describe the n-vectors that have been introduced to complete the basis, the orthonormality of which entails

$$\sum_{k=1}^{n} \bar{a}_{ki}a_{kj} = \delta_{ij} \quad : \quad i,j \text{ range now on } \{1,2,3,\dots,n\}$$

where $\mathbb{U} = \|a_{kj}\|$ is an $n \times n$ unitary matrix in which the first two columns present information derived from the prescribed structure of $\tilde{\mathbb{A}}$ and the remaining columns derive from the basis completion process.

[21] Completion can be accomplished in infinitely many ways