Generalized Quantum Measurement

Imperfect meters and POVMs

Nicholas Wheeler
September 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 have happened recently upon a splendid text[3] in which the author devotes his Chapter 4 ("Generalized measurements") to a richly detailed but very readable survey of the modern essentials of the subject to which I have alluded, and which I am inspired now to revisit. Notational obscurities and some confusing errors (?) in Barnett's text have led me to consult also other sources, particularly some lecture notes by John Preskill[4] and notes from a talk entitled "POVMs and superoperators" by one Mario Flory to fellow students in a *Foundations of Quantum Mechanics* course at the Arnold Sommerfeld Center for Theoretical Physics, Ludwig-Maximilians-Universität München (2010).

[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).

[4] 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. The present topic is developed in Preskill's Chapter 3.

Here, after review of some introductory material, it is Flory's elegant essay that will serve as one of my primary sources.

States of systems *vs* **states of ensembles of systems.** We work within the standard formulation of orthodox (non-relativistic) quantum mechanics,[5] 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,[6] and will often write $\mathcal{H}_n$ (or simply $\mathcal{H}$)in place of $\mathcal{H}_\mathcal{S}$.

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 $\mathbb{A}$. Such matrices—in non-degenerate cases—can be developed

$$\mathbb{A} = \sum_{k=1}^{n} a_k \mathbb{P}_k \quad \text{where} \quad \mathbb{P}_k = |a_k)(a_k|$$

projects onto the 1-dimensional k^{th} eigenspace of $\mathbb{A}$. In degenerate cases we have

$$\mathbb{A} = \sum_{\kappa} a_\kappa \mathbb{P}_\kappa$$

where the a_κ are distinct, a_κ has multiplicity ν_κ ($\sum_\kappa \nu_\kappa = n$) and $\mathbb{P}_\kappa$ projects onto the ν_κ-dimensional κ^{th} eigenspace of $\mathbb{A}$. In either case[7] the $\mathbb{P}$-matrices are

- hermitian
- positive: $(\alpha|\mathbb{P}_i|\alpha) \geq 0$ (all $|\alpha)$)
- complete: $\sum_i \mathbb{P}_i = \mathbb{I}$
- orthogonal: $\mathbb{P}_i\mathbb{P}_j = \delta_{ij}\mathbb{P}_i$

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* (*or at last of the eigenspace that contains the state*) *they have prepared*. But quantum theory permits one to speak only probabilistically about how the meter will respond in any specific instance. The claim—the essential upshot of the **von Neumann projection hypothesis**—is that, in non-degenerate cases,

[5] "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*.

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

[7] In what follows I will usually restrict my explicit remarks and notation to the non-degenerate case.

repeated presentation of $|\psi)$ to an ideal A-meter results—*so long as the meter remains* *unread*— in the serial production of a weighted assortment—briefly: a "mixture"—of states

$$|\psi) \xrightarrow[\mathsf{A}\text{-measurement}]{} \begin{cases} |a_1) \text{ with probability } |(a_1|\psi)|^2 = (\psi|\mathbb{P}_1|\psi) \\ |a_2) \text{ with probability } |(a_2|\psi)|^2 = (\psi|\mathbb{P}_2|\psi) \\ \vdots \\ |a_k) \text{ with probability } |(a_k|\psi)|^2 = (\psi|\mathbb{P}_k|\psi) \\ \vdots \\ |a_n) \text{ with probability } |(a_n|\psi)|^2 = (\psi|\mathbb{P}_n|\psi) \end{cases} \tag{1.1}$$

If the meter *is* read (and a_k is non-degenerate) one has

$$|\psi) \xrightarrow[\mathsf{A}\text{-meter reads } a_k]{} |a_k)$$

But if a_k *is* degenerate one is again left with a mixture

$$|\psi) \xrightarrow[\mathsf{A}\text{-meter reads } a_k]{} \begin{cases} |a_{k,1}) \text{ with probability } |(a_{k,1}|\psi)|^2 = (\psi|\mathbb{P}_{k,1}|\psi) \\ |a_{k,2}) \text{ with probability } |(a_{k,2}|\psi)|^2 = (\psi|\mathbb{P}_{k,2}|\psi) \\ \vdots \\ |a_{k,\nu}) \text{ with probability } |(a_{k,\nu}|\psi)|^2 = (\psi|\mathbb{P}_{k,\nu}|\psi) \end{cases} \tag{1.2}$$

where $\{\mathbb{P}_{k,1}, \mathbb{P}_{k,2}, \ldots, \mathbb{P}_{k,\nu}\}$ project onto some/any orthonormal basis within the ν-dimensional k^{th} eigenspace of $\mathbb{A}$.[8]

Note that

$$\begin{aligned}(1.1) &\Longrightarrow \sum \text{probabilities} = (\psi|\big(\textstyle\sum_k \mathbb{P}_k\big)|\psi) = (\psi|\psi) = 1 \\ (1.2) &\Longrightarrow \sum \text{probabilities} = (\psi|\big(\textstyle\sum_j \mathbb{P}_{k,j}\big)|\psi) = (\psi|\mathbb{P}_k|\psi)\end{aligned}$$

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

$$\text{expected mean } \langle \mathsf{A} \rangle_\psi = \sum_k a_k (\psi|\mathbb{P}_k|\psi) = (\psi|\mathbb{A}|\psi)$$

irrespective of whether the spectrum is non-degenerate or degenerate.

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 measurement

[8] It is tempting *but would be incorrect* to say that (1.1) and (1.2) describe mixtures of $\mathbb{A}$ eigenstates, for—as will emerge—quantum mixtures, unlike classical mixtures of (say) colored balls, do not admit of unambiguous resolution into constituent parts.

from which the output states are sometimes $|\psi)$ (as announced by the G-meter) but are more typically states $|\psi)_{\text{undesired}}$ that a filter or gate serves to discard. Schematic illustration of the preselection process is provided by the following diagram:

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

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 to the A-meter:

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

A-measurement (performed with a perfect A-meter) can be expected in such a cirvcumstance to announce a_k with probability $\sum_j p_j(\psi_j|\mathbb{P}_k|\psi_j)$. The sum of those probabilities is

$$\sum_k \sum_j p_j(\psi_j|\mathbb{P}_k|\psi_j) = \sum_j p_j(\psi_j|\psi_j) = \sum_j p_j = 1$$

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

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

It will be appreciated that the **density matrix** $\wp_{\mathcal{E}}$ refers not to the state of a system $\mathcal{S}$ but to the observationally relevant features of an *ensemble* $\mathcal{E}$ (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.[9] 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* individual quantum 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. But—awkwardly—it would, as previously remarked, be equally misleading to speak of the "composition" of an ensemble. . . which may account for the fact that both of the fussy points just mentioned are commonly ignored in relaxed quantum discourse.

Properties of density matrices. It must be emphasized at the outset that the states $|\psi_\nu)$ that enter into the construction

$$\rho = \sum_\nu |\psi_\nu)p_\nu(\psi_\nu| \tag{2}$$

are not required to be orthogonal, or the eigenvectors of anything (though at (1.1) they happened to be). From (2) it follows immediately that all density matrices are

- hermitian
- positive: $(\alpha|\rho|\alpha) \geq 0$ (all $|\alpha)$)
- have unit trace: $\text{tr}\rho = \sum_\nu p_\nu = 1$

Conversely, every marix ρ with those properties admits of interpretation as a density matrix.

Hermiticity entails the possibility of spectral decomposition[10]

$$\rho = \sum_k |r_k)r_k(r_k| = \sum_\kappa r_\kappa \mathbb{P}_\kappa$$

where the r_κ are distinct eigenvalues of ρ and the $\mathbb{P}_\kappa$ project onto the associated eigenspaces. Positivity implies that all eigenvalues r_κ are non-negative, while the unit trace condition asserts that they sum to unity. The twice-mentioned *non-uniqueness of quantum mixtures* is illustrated by the observation that

$$\rho = \begin{cases} \sum_\nu |\psi_\nu)p_\nu(\psi_\nu| \text{ displays } \rho \text{ as the } p_\nu\text{-weighted mixture of } |\psi_\nu)\text{-states} \\ \\ \sum_k |r_k)r_k(r_k| \text{ displays } \rho \text{ as the } r_k\text{-weighted mixture of eigenstates } |r_k) \end{cases}$$

---

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

[10] Of which we encountered an instance already at (1.1):

$$\rho = \sum_k |a_k)r_k(a_k| \quad \text{with} \quad r_k = (\psi|\mathbb{P}_k|\psi)$$

By easy argument we have

$$\operatorname{tr}\rho^2 = \textstyle\sum_k p_\nu^2 = \sum_k r_k^2 \;\le\; 1, \quad \text{with equalty iff } \rho \text{ is pure}$$

Relatedly, a density matrix ρ is projective

$$\rho^2 = \rho \quad \text{iff } \rho \text{ is pure}$$

Let ρ_1 and ρ_2 be density matrices, and look to

$$\rho(x) = x\rho_1 + (1-x)\rho_2 \quad : \quad 0 \le x \le 1$$

which interpolates linearly between them. It is immediately evident that $\rho(x)$ possesses all of the properties (hermiticity, positivity, unit trace) characterisitic of density matrices. We conclude that the set $\{\rho\}$ of all $n\times n$ density matrices is *convex*, and that so also therefore is the set $\{\mathcal{E}\}$ of ensembles to which they refer. Moreover,

$$\operatorname{tr}\rho^2(x) = x^2 \operatorname{tr}\rho_1^2 + 2x(1-x)\operatorname{tr}\rho_1\rho_2 + (1-x)^2 \operatorname{tr}\rho_2^2$$

in which connection we write

$$\rho_1 = \sum_{k=1}^n |r_k)r_k(r_k|, \quad \rho_2 = \sum_{k=1}^n |s_k)s_k(s_k|$$

and observe that

$$\begin{aligned}
\tfrac{1}{n} \le \operatorname{tr}\rho_1^2 &= \sum r_k^2\,(\le 1), \text{ with equality iff all } r_k \text{ are equal}\\
\tfrac{1}{n} \le \operatorname{tr}\rho_2^2 &= \sum s_k^2\,(\le 1), \text{ with equality iff all } s_k \text{ are equal}\\
\operatorname{tr}\rho_1\rho_2 &= \sum_{j,k} r_j s_k (r_j|s_k)(s_k|r_j) \le \sum_{j,k} r_j s_k (r_j|r_j)(s_k|s_k) \text{ by Cauchy-Schwarz}\\
&= \sum_{j,k} r_j s_k = 1
\end{aligned}$$

I am satisfied on the basis of numerical experimentation (but don't know how to prove) that in fact

$$\operatorname{tr}\rho_1\rho_2 \le \max\big\{\operatorname{tr}\rho_1^2, \operatorname{tr}\rho_2^2\big\} \text{ with equality iff } \rho_1 = \rho_2$$

and am brought to the conclusion (supported again by experimental evidence) that

$$\rho(x) \text{ is pure iff } \rho_1 = \rho_2 \text{ is pure}$$

which is to say: *pure density matrices live on the boundary* of $\{\rho\}$.

It is obvious/trivial that every pure density matrix, since (hermitian and) projective, can be *factored*

$$\rho_{\text{pure}} = \mathbb{P}^+\mathbb{P} \quad \text{with} \quad \mathbb{P} = \rho_{\text{pure}}$$

but not quite so obvious that *every* density matrix—whether pure or impure—can be factored. I discuss how this comes about. The Schur decomposition theorem asserts that any real or complex square matrix $\mathbb{M}$ can be rendered

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

where $\mathbb{U}$ is unitary and $\mathbb{T}$—the "Schur form" of $\mathbb{M}$—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 $\mathbb{M}$ 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 $\mathbb{M}$ is hermitian the off-diagonal elements of $\mathbb{T}$ vanish, and the Schur decomposition of $\mathbb{M}$ assumes the form

$$\mathbb{M} = \underbrace{\begin{pmatrix} \boxed{\mathbf{e}_1} & \boxed{\mathbf{e}_2} & \cdots & \boxed{\mathbf{e}_n} \end{pmatrix}}_{\mathbb{U}} \underbrace{\begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}}_{\mathbb{D}} \underbrace{\begin{pmatrix} \boxed{\mathbf{e}_1{}^+} \\ \boxed{\mathbf{e}_2{}^+} \\ \vdots \\ \boxed{\mathbf{e}_n{}^+} \end{pmatrix}}_{\mathbb{U}^+}$$

where $\mathbb{M}\,\mathbf{e}_i = \lambda_i\,\mathbf{e}_i$ serves to establish our notation. The orthonormality statements $\mathbf{e}_i^+\mathbf{e}_j = \delta_{ij}$ can be rendered $\mathbb{U}^+\mathbb{U} = \mathbb{I}$. It is less obvious—but follows from the circumstance that in the present context left inverses are also right inverses—that

$$\mathbb{U}\,\mathbb{U}^+ = \mathbb{I}$$

We conclude that if $\mathbb{M}$ is hermitian then it is always possible to write

$$\mathbb{M}^p = \mathbb{U}\begin{pmatrix} \lambda_1^p & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2^p & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3^p & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n^p \end{pmatrix}\mathbb{U}^+ \quad : \quad \text{all } \lambda_k \text{ real}$$

which in the case $p = 0$ gives back the preceding equation. It is by now evident that the Schur decomposition of hermitian matrices amounts simply to a reformulation of the spectral decomposition.[11] If $\mathbb{M}$ is *positive* hermitian then it makes sense to write $\mathbb{D} = (\sqrt{\mathbb{D}})^2$

$$\sqrt{\mathbb{D}} = \begin{pmatrix} \sqrt{\lambda_1} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & 0 & \cdots & 0 \\ 0 & 0 & \sqrt{\lambda_3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{\lambda_n} \end{pmatrix}$$

which becomes also positive hermitian if we adopt the convention that all $\sqrt{\lambda_k}$ are non-negative (which is to dismiss $2^n - 1$ of the possible sign allocations). We conclude that <u>every positive hermitian $\mathbb{M}$ can be written</u>

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

Notice that the preceding factorization of such matrices $\mathbb{M}$ is *not unique*, for it is invariant under

$$\mathbb{W} \longrightarrow \mathbb{W}' = \mathbb{V}\mathbb{W} \quad : \quad \mathbb{V} \text{ arbitrary unitary}$$

Density matrices ρ are distinguished from the generality of positive hermitian matrices only by the circumstance that they satisfy the unit-trace condition:

$$\sqrt{\mathbb{D}} = \begin{pmatrix} \sqrt{r_1} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{r_2} & 0 & \cdots & 0 \\ 0 & 0 & \sqrt{r_3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{r_n} \end{pmatrix} \quad \text{with} \quad \textstyle\sum_k r_k = 1$$

DIGRESSION: Wishart matrices

Let $\mathbb{W}$ be an $m \times n$ complex matrix. Then

$$\mathbb{M}_1 = \mathbb{W}\,\mathbb{W}^+$$
$$\mathbb{M}_2 = \mathbb{W}^+\,\mathbb{W}$$

are manifestly square hermitian, of dimensions m and n respectively. Such matrices are called "Wishart matrices," after John Wishart (1898–1956), who in 1928 was motivated to introduce such matrices into the biometric statistical literature. We note in passing that $\mathbb{M}_1 = \mathbb{M}_2$ (which requires $m = n$) is precisely the condition that the square matrix $\mathbb{W}$ be "normal."[12]

[11] In *Mathematica*-based numerical work the command `SchurDecomposition` provides an efficient way to construct the spectral decomposition. For positive hermitian matrices the Schur decomposition becomes identical to the SVD.

[12] We note also that "supersymmetric quantum mechanics" arises from the interplay of systems with hamiltonians of the specialized forms $\mathbf{H}_1 = \mathbf{A}\,\mathbf{A}^+$ and $\mathbf{H}_1 = \mathbf{A}^+\mathbf{A}$.

Wishart matrices possess a number of notable spectral properties. The eigenvalues of $\mathbb{M}_1$ (similarly $\mathbb{M}_2$) are, by hermiticity, necessarily real. From

$$(\alpha|\mathbb{M}_1|\alpha) = \|\mathbb{W}^+|\alpha)\|^2 \geq 0 \quad : \quad \text{all } |\alpha)$$

we discover that $\mathbb{W}_1$ (similarly $\mathbb{W}_2$) is *positive semi-definite* (all eigenvalues non-negative). From

$$\begin{aligned} \operatorname{tr}\mathbb{M}_1 &= \operatorname{tr}\mathbb{W}\mathbb{W}^+ = \operatorname{tr}\mathbb{W}^+\mathbb{W} = \operatorname{tr}\mathbb{M}_2 \\ \operatorname{tr}\mathbb{M}_1^2 &= \operatorname{tr}\mathbb{W}\mathbb{W}^+\mathbb{W}\mathbb{W}^+ = \operatorname{tr}\mathbb{W}^+\mathbb{W}\mathbb{W}^+\mathbb{W} = \operatorname{tr}\mathbb{M}_2^2 \\ &\vdots \end{aligned}$$

we learn that $\operatorname{tr}\mathbb{M}_1^p = \operatorname{tr}\mathbb{M}_2^p$ $(p = 0, 1, 2, \ldots)$. Since the coefficients that appear in the characteristic polynomial of a matrix can be assembled from powers of traces of powers of the matrix, and the formulæ that accomplish that assembly are universal, the matrices $\mathbb{M}_1$ and $\mathbb{M}_2$ must have identical (reduced) characteristic polynomials. Assume $m < n$ and let the spectrum of $\mathbb{M}_1$ be denoted $\{\mu_1, \mu_2, \ldots, \mu_m\}$. The spectrum of $\mathbb{M}_2$ has then the structure $\{\mu_1, \mu_2, \ldots, \mu_m, 0, \ldots, 0\}$ with $n - m$ dangling zeros. The intimate relationship between the spectra of $\mathbb{M}_1$ and $\mathbb{M}_2$ is reflected in a similarly intimate relationship between their eigenvectors, though the eigenvectors $\{|\alpha_1), |\alpha_2), \ldots, |\alpha_m)\}$ of $\mathbb{M}_1$ are m-dimensional, while those $\{|\beta_1), |\beta_2), \ldots\ldots |\beta_n)\}$ of $\mathbb{M}_2$ are n-dimensional:

$$\begin{aligned} \mathbb{M}_1|\alpha) = \mu|\alpha) \quad &\Longrightarrow \quad \mathbb{M}_2|\beta) = \mu|\beta) \quad \text{with } |\beta) = \mathbb{W}^+|\alpha) \\ \mathbb{M}_2|\beta) = \mu|\beta) \quad &\Longrightarrow \quad \mathbb{M}_1|\alpha) = \mu|\alpha) \quad \text{with } |\alpha) = \mathbb{W}\ |\beta) \end{aligned}$$

Those associations (which, by the way, play a central role in supersymmetric quantum mechanics) fail, however, for eigenvectors that lie in the null-spaces of either $\mathbb{M}$-matrix, so the pattern becomes complete only if both $\mathbb{M}$-matrices are strictly positive.

Wishart matrices have been an unremarked familiar part of our quantum mechanical lives. If, for example, we identify $|\psi)$ with its matrix representation then

$$\begin{aligned} (\psi|\psi) &\text{ is the } 1\times 1 \text{ unit Wishart matrix} \\ |\psi)(\psi| &\text{ is the } n\times n \text{ Wishart projector onto } |\psi) \end{aligned}$$

Matrices of the $n \times n$ Wishart form

$$\rho = \mathbb{R}^+\mathbb{R} \quad \text{with} \quad \mathbb{R} = \frac{\mathbb{W}}{\sqrt{\operatorname{tr}(\mathbb{W}^+\mathbb{W})}}\ ,\ \ \mathbb{W}\ m\times n \text{ complex} \tag{3}$$

possess all of the properties required of a density matrix, and indeed: every density matrix can be written (in many ways) as such a product of factors. We can—as has been demonstrated—always arrange for the factors to be square; that done, they are determined only to within "gauge transformations" $\mathbb{R} \longrightarrow \mathbb{R}' = \mathbb{V}\mathbb{R}$ ($\mathbb{V}$ unitary).The preceding equation provides a convenient way to generate random density matrices for use in computational experiments such as the one to which I referred on page 6.

end of digression

First steps toward a theory of imperfect meters. In "Quantum measurement with imperfect devices"[2]—which was written in shameful ignorance of the relevant literature (classic texts provided no hint that such a literature even existed), as a hasty supplement to some class notes[1]—I use the simplest of means to trace out some of the implications of the elementary notion which I will sketch in a moment. I worked there on the assumption that perfect meters can be represented by self-adjoint operators **A** with continuous non-degenerate spectra—assumptions which I now relax: the prevailing assumption that the state space $\mathcal{H}_\mathcal{S}$ is finite dimensional entails replacement of operators **A** by finite-dimensional hermitian matrices $\mathbb{A}$ with discrete spectra. While it proves convenient to assume initially that the $\mathbb{A}$-spectrum is non-degenerate, I will be at pains later (and find it easy) to relax that assumption.

The action of an ideal **A**-meter can be diagramed (see again (1.1))

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

Here I emphasize that the important thing about $\mathbb{A}$ is the complete set $\{\mathbb{P}_1, \mathbb{P}_2, \ldots, \mathbb{P}_n\}$ of orthogonal projectors to (by spectral decomposition) it gives rise; the associated "meter marks" $\{a_1, a_2, \ldots, a_n\}$ could be replaced by any other set $\{\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_n\}$ of distinct real numbers without affecting the essential physics of the meter.

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

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

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

$$\rho_{\text{out},k} = \sum_j w_{j|k}\mathbb{P}_j \quad : \quad \sum_j w_{j|k} = 1 \text{ (all } k) \tag{4}$$

Observe that

$$\text{tr}\,\rho_{\text{out},k} = \sum_j w_{k|j}\text{tr}\mathbb{P}_j = 1$$

while

$$\text{tr}\,\rho^2_{\text{out},k} = \sum_j w_{k|j}{}^2 \leq 1$$

with equality if and only if the meter is in fact ideal. We expect "good imperfect meters" to be "fuzzy" but not to be flagrant liers; *i.e.*, we expect to have

$$\max(w_{j|k}) = w_{k|k} \quad ; \quad \text{all } k$$

Note finally that—by (4)—when ρ_{in} is presented to an imperfect **A**-meter the probability that the meter will register a_k is

$$\begin{aligned}\text{prob}(\rho_{\text{in}}, a_k) &= \sum_j w_{j|k}\text{tr}\big(\rho_{\text{in}}\mathbb{P}_j\big)\\ &= \text{tr}\big(\rho_{\text{in}}\tilde{\mathbb{P}}_k\big) \quad \text{where} \quad \tilde{\mathbb{P}}_k = \sum_j w_{j|k}\mathbb{P}_j \end{aligned}\tag{5.1}$$

where

$$\sum_k \text{prob}(\rho_{\text{in}}, a_k) = 1 \quad \Longrightarrow \quad \sum_k \tilde{\mathbb{P}}_k = \mathbb{I} \tag{5.2}$$

In my seminar[2] I used von Neumann entropy in an attempt to quantify the information gained by imperfect measurement. I defer discussion of that aspect of our subject.

Sophisticated reformulation & ramifications: POVMs. It is to Barnett,[3] Preskill,[4] Flory and other modern authors that I owe the realization that my seminar would have been more valuable had I recognized and traced out the implications of (5), which I now undertake to do.

Just as $\{\mathbb{P}_1, \mathbb{P}_2, \ldots, \mathbb{P}_n\}$ provides—as previously remarked—a complete description of the essential features of an ideal meter, so do $\{\tilde{\mathbb{P}}_1, \tilde{\mathbb{P}}_2, \ldots\}$ provide a complete characterization of a non-ideal meter. But while the $\mathbb{P}$-matrices are hermitian, positive, complete and orthogonal, the $\tilde{\mathbb{P}}$-matrices are seen to be

- hermitian
- positive: $(\alpha|\tilde{\mathbb{P}}_i|\alpha) \geq 0$ (all $|\alpha)$)
- complete: $\sum_i \tilde{\mathbb{P}}_i = \mathbb{I}$
- typically non-projective and non-orthogonal: $\tilde{\mathbb{P}}_i\tilde{\mathbb{P}}_j \neq \delta_{ij}\tilde{\mathbb{P}}_i$

Note also that while on the dials of ideal meters the number of "meter marks" a_k is constrained to be $\nu \leq n$ (with equality in non-degenerate cases), the dials of non-ideal meters are subject to no such constraint

Sets $\{\tilde{\mathbb{P}}_1, \tilde{\mathbb{P}}_2, \ldots\}$ of $n\times n$ matrices endowed with the properties listed above are called "positive operator-valued measures," or POVMs, and the generalized measurements to which they give rise are called "POVM measurements."[13] The idealized von Neumann measurements that proceed from specification of $\{\mathbb{P}_1, \mathbb{P}_2, \ldots, \mathbb{P}_\nu\}$ are within this enlarged context called "projection-valued measurements," or PVMs. Every element $\tilde{\mathbb{P}}_k$, by positive hermiticity, admits of spectral development $\tilde{\mathbb{P}}_k = \sum_j p_{j,k}\mathbb{P}_{j,k}$, where $\{\mathbb{P}_{1,k}, \mathbb{P}_{2,k}, \ldots, \mathbb{P}_{\nu\leq n,k}\}$ is a complete set of orthogonal projectors and the $p_{j,k}$ are non-negative reals. If every $\tilde{\mathbb{P}}_k$ gives rise to the *same* set of projectors then we have $\tilde{\mathbb{P}}_k = \sum_j p_{j,k}\mathbb{P}_j$ which when $\sum_j p_{j,k} = 1$ (all k) becomes an instance of (5). Evidently the theory of imperfect devices[2] that I sketched in 2000 is a *special case* of the POVM-based theory of generalized quantum measurement.

---

[13] 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]."

In density matrix language the PVM scheme (1.1) reads

$$
\begin{aligned}
&\rho_{\text{in}} = |\psi)(\psi| \\
&\quad\Big\downarrow \text{ ideal } \mathsf{A}\text{-meter reads } a_k \\
&\rho_{\text{out},k} = \mathbb{P}_k \equiv |a_k)(a_k| \quad \text{with probability} \quad \text{tr}\big(\mathbb{P}_k^+\rho_{\text{in}}\mathbb{P}_k\big)
\end{aligned}
$$

from which it follows that if the state presented to the ideal meter is drawn from a *mixed* ensemble

$$
\begin{aligned}
&\rho_{\text{in}} \\
&\quad\Big\downarrow \text{ ideal } \mathsf{A}\text{-meter reads } a_k \\
&\rho_{\text{out},k} = \frac{\mathbb{P}_k^+\rho_{\text{in}}\mathbb{P}_k}{\text{tr}\big(\mathbb{P}_k^+\rho_{\text{in}}\mathbb{P}_k\big)} \quad \text{with probability} \quad \text{tr}\big(\mathbb{P}_k^+\rho_{\text{in}}\mathbb{P}_k\big)
\end{aligned}
\tag{6.1}
$$

The probability of such an outcome can—by the hermitian projectivity of the $\mathbb{P}_k$-matrices—be described in several equivalent ways:

$$
\text{tr}\big(\mathbb{P}_k^+\rho_{\text{in}}\mathbb{P}_k\big) = \text{tr}\big(\rho_{\text{in}}\mathbb{P}_k\mathbb{P}_k\big) = \text{tr}\big(\rho_{\text{in}}\mathbb{P}_k\big)
$$

If, however, the meter remains unread we have

$$
\begin{aligned}
&\rho_{\text{in}} \\
&\quad\Big\downarrow \text{ ideal } \mathsf{A}\text{-meter remains unread} \\
&\rho_{\text{out},k} = \sum_k \text{tr}\big(\mathbb{P}_k^+\rho_{\text{in}}\mathbb{P}_k\big)\cdot\frac{\mathbb{P}_k^+\rho_{\text{in}}\mathbb{P}_k}{\text{tr}\big(\mathbb{P}_k^+\rho_{\text{in}}\mathbb{P}_k\big)} = \sum_k \mathbb{P}_k^+\rho_{\text{in}}\mathbb{P}_k
\end{aligned}
\tag{6.2}
$$

In either case, we have $\text{tr}\big(\rho_{\text{out},k}\big) = \text{tr}\big(\rho_{\text{in}}\big) = 1$ (in the latter case by completeness: $\sum_k \mathbb{P}_k = \mathbb{I}$). From (6.1) if follows (by the orthogonality of the $\mathbb{P}$-matrices) that prompt repetition of such a measurement will yield a_j with probability

$$
\text{tr}\big(\mathbb{P}_j^+\rho_{\text{out},k}\mathbb{P}_j\big) = \frac{\text{tr}\big(\mathbb{P}_j^+\mathbb{P}_k^+\rho_{\text{in}}\mathbb{P}_k\mathbb{P}_j\big)}{\text{tr}\big(\mathbb{P}_k^+\rho_{\text{in}}\mathbb{P}_k\big)} = \delta_{ij}
$$

which is to say: prompt repetition serves to "verify" the preceding meter reading. It is sometimes held that verifiability of this order (exact reproducability) is essential to the very *meaning* of quantum measurement —necessary if we are to claim that the measurement taught us something— though that is a standard to which common laboratory measurements (of length, mass, *etc.*) do not rise, and as will soon emerge it is violated when quantum mesurements are performed with realistically imperfect devices. More significantly, prompt verifiability is the source of the troublesome notion — central to the Copenhagen interpretation though it has been and remains—that (projective) measurement "causes the wave function to 'collapse'."

It is plausible, in view of (6.1), that the result of a POVM measurement might be described

$$\begin{array}{l} \rho_{\text{in}} \\ \Big\downarrow \text{ imperfect } \mathsf{A}\text{-meter reads } a_k \\ \rho_{\text{out},k} = \dfrac{\tilde{\mathbb{P}}_k^+ \rho_{\text{in}} \tilde{\mathbb{P}}_k}{\text{tr}\big(\tilde{\mathbb{P}}_k^+ \rho_{\text{in}} \tilde{\mathbb{P}}_k\big)} \quad \text{with probability} \quad \text{tr}\big(\tilde{\mathbb{P}}_k^+ \rho_{\text{in}} \tilde{\mathbb{P}}_k\big) \end{array} \tag{$\star$}$$

but this would present a problem, for because the $\tilde{\mathbb{P}}$-matrices are typically not projective the conjectured probabilities do not sum to unity:

$$\sum_k \text{tr}\big(\tilde{\mathbb{P}}_k^+ \rho_{\text{in}} \tilde{\mathbb{P}}_k\big) = \sum_k \text{tr}\big(\tilde{\mathbb{P}}_k^2 \rho_{\text{in}}\big) \neq \text{tr}\Big(\big(\textstyle\sum_k \tilde{\mathbb{P}}_k\big) \rho_{\text{in}}\Big) = \text{tr}\, \rho_{\text{in}} = 1$$

To overcome this difficuty we look to the squre Wishart factors of $\tilde{\mathbb{P}}_k$, writing

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

and in place of ($\star$) write

$$\begin{array}{l} \rho_{\text{in}} \\ \Big\downarrow \text{ imperfect } \mathsf{A}\text{-meter reads } a_k \\ \rho_{\text{out},k} = \dfrac{\mathbb{A}_k^+ \rho_{\text{in}} \mathbb{A}_k}{\text{tr}\big(\mathbb{A}_k^+ \rho_{\text{in}} \mathbb{A}_k\big)} \quad \text{with probability} \quad \text{tr}\big(\mathbb{A}_k^+ \rho_{\text{in}} \mathbb{A}_k\big) \end{array} \tag{7.1}$$

whence

$$\begin{array}{l} \rho_{\text{in}} \\ \Big\downarrow \text{ imperfect } \mathsf{A}\text{-meter remains unread} \\ \rho_{\text{out},k} = \displaystyle\sum_k \text{tr}\big(\mathbb{A}_k^+ \rho_{\text{in}} \mathbb{A}_k\big) \cdot \frac{\mathbb{A}_k^+ \rho_{\text{in}} \mathbb{A}_k}{\text{tr}\big(\mathbb{A}_k^+ \rho_{\text{in}} \mathbb{A}_k\big)} = \sum_k \mathbb{A}_k^+ \rho_{\text{in}} \mathbb{A}_k \end{array} \tag{7.2}$$

The probabilities now (by the postulated completeness of the $\tilde{\mathbb{P}}$-matrices) *do* sum to unity :

$$\sum_k \text{tr}\big(\mathbb{A}_k^+ \rho_{\text{in}} \mathbb{A}_k\big) = \sum_k \text{tr}\big(\mathbb{A}_k \mathbb{A}_k^+ \rho_{\text{in}}\big) = \text{tr}\Big(\big(\textstyle\sum_k \tilde{\mathbb{P}}_k\big) \rho_{\text{in}}\Big) = \text{tr}\, \rho_{\text{in}} = 1$$

Moreover, for ideal meters $\tilde{\mathbb{P}}_k \longrightarrow \mathbb{P}_k = \mathbb{P}_k^2 = \mathbb{A}_j \mathbb{A}_j^+$ with $\mathbb{A}_j = \mathbb{P}_j$ and (7) give back (6). Note, however, that

$$\text{tr}\big(\mathbb{A}_j^+ \rho_{\text{out},k} \mathbb{A}_j\big) = \frac{\text{tr}\big(\mathbb{A}_j^+ \mathbb{A}_k^+ \rho_{\text{in}} \mathbb{A}_k \mathbb{A}_j\big)}{\text{tr}\big(\mathbb{A}_k^+ \rho_{\text{in}} \mathbb{A}_k\big)} \neq \delta_{ij}$$

since typically $\mathbb{A}_k \mathbb{A}_j \neq \delta_{jk} \mathbb{A}_k$. Prompt repetition with an imperfect meter does *not* serve to "verify" the preceding meter reading. With this development the already-murky notion that "measurement causes the wave function to collapse" becomes even more murky.[14]

The POVM schemes (7) indicate that from a procedural standpoint it is not $\{\tilde{\mathbb{P}}_1, \tilde{\mathbb{P}}_2, \ldots\}$ but the set of Wishart factors $\{\mathbb{A}_1, \mathbb{A}_2, \ldots\}$ that most properly serves to describe the action of an imperfect **A**-meter. Gauge transformations

$$\{\mathbb{A}_1, \mathbb{A}_2, \ldots\} \longrightarrow \{\mathbb{A}_1\mathbb{V}_1, \mathbb{A}_2\mathbb{V}_2, \ldots\} \quad : \quad \text{all } \mathbb{V}\text{-matrices unitary}$$

leave unaltered both $\{\tilde{\mathbb{P}}_1, \tilde{\mathbb{P}}_2, \ldots\}$ and the probability that the imperfect meter announces a_k

$$\text{tr}\big(\mathbb{A}_k^+ \rho_{\text{in}} \mathbb{A}_k\big) = \text{tr}\big(\mathbb{V}_k^+ \mathbb{A}_k^+ \rho_{\text{in}} \mathbb{A}_k \mathbb{V}_k\big)$$

but subjects the output ensemble to a unitary similarity transformation:

$$\rho_{\text{out},k} \longrightarrow \mathbb{V}_k^+ \rho_{\text{out},k} \mathbb{V}_k$$

The PVM-POVM relationship is in several respects deeper and more interesting than is suggested by the fact—noted above—that POVM-theory gives back PVM-theory as a special case, for there are important contexts within which PVM measurements give rise spontaneously to PVOM measurements in spaces of reduced dimension, while Neumark's theorem (of which more later) asserts that every PVOM measurement can be realized as a PVM measurement in a space of augmented dimension. I explore those claims in the order stated.

Model of the quantum measurement process.[15] 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 (generated in time τ by an $mn \times mn$ Hamiltonian $\mathbb{H}_{\text{interaction}}$) 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

---

[14] The question arises (discussion of which I will defer): *How nearly* can repeated measurements with the same POVM device be expected to agree? Relatedly, what becomes of the "quantum Zeno effect" (Alan Turing (1954), George Sudarshan (1974)) if the rapidly-repeated measurements are performed not (as commonly assumed) with an ideal PVM device but (more realistically) with a non-ideal POVM device? Sudarshan himself drew attention to this question in the last sentence of his original paper.

[15] My primary source here has been Barnett's §4.3. Closely related material can be found in Preskill's §3.1.2, Flory's §3.1 and §2.2.8 of M. A. Nielson & I. I. Chuang's *Quantum Computation & Quantum Information* (2000).

PVM measurement 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

$$\text{Prob}_{ij} = \Big|\textstyle\sum_k (E_{ij}|\mathbb{U}|A_k)\,\psi_k\Big|^2 \tag{8}$$

Now introduce the ij-indexed n-dimensional bra vectors

$$(\pi_{ij}| = \Big((E_{ij}|\mathbb{U}|A_1)\;\;(E_{ij}|\mathbb{U}|A_1)\;\;\ldots\;\;(E_{ij}|\mathbb{U}|A_1)\Big)$$

and obtain

$$\begin{aligned}\text{Prob}_{ij} &= \big|(\pi_{ij}|\psi)\big|^2\\ &= (\psi|\tilde{\mathbb{P}}_{ij}|\psi) \quad\text{with}\quad \tilde{\mathbb{P}}_{ij} = |\pi_{ij})(\pi_{ij}|\end{aligned} \tag{9}$$

Compare this result with (8), which can be written

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

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 nm in number, are clearly positive hermitian. 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:

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

Equation (9) looks superficially like a description of the probability that an **A**-meter—represented by the $n\times n$ hermitian matrix $\mathbb{A}=\sum_{ij}a_{ij}\tilde{\mathbb{P}}_{ij}$—will announce a_{ij}, in which case it would refer to the result of subjecting $\mathcal{S}$ to a PVM measurement. But no PVM-meter can have so many marks on its dial ($mn>n$), nor can so many projectors $\tilde{\mathbb{P}}_{ij}$ appear in the spectral representation of such an $\mathbb{A}$. In fact the matrices $\tilde{\mathbb{P}}_{ij}$ are *not* projective because the n-vectors $|\pi_{ij})$ are not unit vectors: 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, as before, $|\psi)=\sum_q|e_q)\psi_q$. 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 it *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 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 product. Barnett's line of argument (rehearsed above) 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 Barnett has nothing to say about the specific construction of the Hamiltonian 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

$$\mathbb{A}_\text{super}=\sum_{ij}a_{ij}\mathbb{Q}_{ij}$$

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

From POVM to PVM: Neumark's dilation theorem. Though Neumark's paper[16] is only three pages long, the Wikipedia article presumes command of a lot of fairly abstruse mathematics, which, I suppose, is why several authors have been content to dismiss the subject with an illustrative example. Preskill, however, has managed to capture the essence of Neumark's theorem quite simply:[4] from the unnormalized n-vectors

$$|\phi_1)=\begin{pmatrix}\phi_{11}\\ \phi_{12}\\ \vdots\\ \phi_{1n}\end{pmatrix},\ |\phi_2)=\begin{pmatrix}\phi_{21}\\ \phi_{22}\\ \vdots\\ \phi_{2n}\end{pmatrix},\dots,\ |\phi_N)=\begin{pmatrix}\phi_{N1}\\ \phi_{N2}\\ \vdots\\ \phi_{Nn}\end{pmatrix}\quad:\quad N\geq n$$

construct the N-element POVM

$$\{\tilde{\mathbb{P}}_1,\tilde{\mathbb{P}}_2,\dots,\tilde{\mathbb{P}}_N\}\quad\text{with}\quad\tilde{\mathbb{P}}_a=|\phi_a)(\phi_a|$$

that operates in $\mathcal{H}_n$. When written out in component form the condition $\sum_a\tilde{\mathbb{P}}_a=\mathbb{I}$ reads

$$\sum_a(\tilde{\mathbb{P}}_a)_{ij}=\sum_{a=1}^N\phi_{ai}\bar{\phi}_{aj}=\delta_{ij}$$

[16] 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).

Now interpret ϕ_{ai} to be not the i^{th} element of the a^{th} member $|\phi_a)$ of a set of n-vectors but to be the a^{th} element of the i^{th} member $|\varPhi_i)$ of a set of N-vectors. Then

$$\sum_{a=1}^{N}\phi_{ai}\bar{\phi}_{aj}=\delta_{ij}\quad\text{reads}\quad\sum_{a=1}^{N}\varPhi_{ia}\bar{\varPhi}_{ja}=\delta_{ij}$$

which is simply the statement that $|\varPhi_i)$ and $|\varPhi_j)$ are orthonormal. Complete the $|\varPhi)$-basis in $\mathcal{H}_N$ by adjoining to $\{|\varPhi_1),|\varPhi_2),\ldots,|\varPhi_n)\}$ orthogonal vectors $\{|\varPhi_{n+1}),|\varPhi_{n+2}),\ldots,|\varPhi_N)\}$. Feed the elements of $|\varPhi_i)$ into the i^{th} row of an $N\times N$ matrix

$$\mathbb{U}=\begin{pmatrix}\varPhi_{11} & \varPhi_{12} & \cdots & \varPhi_{1n} & \varPhi_{1,n+1} & \cdots & \varPhi_{1N}\\ \varPhi_{21} & \varPhi_{22} & \cdots & \varPhi_{2n} & \varPhi_{2,n+1} & \cdots & \varPhi_{2N}\\ \vdots & \vdots & & \vdots & \vdots & & \vdots\\ \varPhi_{n1} & \varPhi_{n2} & \cdots & \varPhi_{nn} & \varPhi_{n,n+1} & \cdots & \varPhi_{nN}\\ \varPhi_{n+1,1} & \varPhi_{n+1,2} & \cdots & \varPhi_{n+1,n} & \varPhi_{n+1,n+1} & \cdots & \varPhi_{n+1,N}\\ \vdots & \vdots & & \vdots & & & \\ \varPhi_{N1} & \varPhi_{N2} & \cdots & \varPhi_{Nn} & \varPhi_{N,n+1} & \cdots & \varPhi_{NN}\end{pmatrix}$$

$$=\begin{pmatrix}\phi_{11} & \phi_{21} & \cdots & \phi_{n1} & \phi_{n+1,1} & \cdots & \phi_{N1}\\ \phi_{12} & \phi_{22} & \cdots & \phi_{n2} & \phi_{n+1,2} & \cdots & \phi_{N2}\\ \vdots & \vdots & & \vdots & \vdots & & \vdots\\ \phi_{1n} & \phi_{2n} & \cdots & \phi_{nn} & \phi_{n+1,n} & \cdots & \phi_{Nn}\\ \varPhi_{n+1,1} & \varPhi_{n+1,2} & \cdots & \varPhi_{n,n+1} & \varPhi_{n+1,n+1} & \cdots & \varPhi_{n+1,N}\\ \vdots & \vdots & & \vdots & & & \\ \varPhi_{N1} & \varPhi_{N2} & \cdots & \varPhi_{Nn} & \varPhi_{N,n+1} & \cdots & \varPhi_{NN}\end{pmatrix}$$

Row-orthonormality $\iff$ $\mathbb{U}$-unitarity $\iff$ column-orthonormality. Let

$$|E_a)=a^{\text{th}}\text{ column in }\mathbb{U}$$

Then $(E_a|E_b)=\delta_{ab}$ and $|E_a)(E_a|$ assumes the form

$$\begin{pmatrix}\phi\\ \phi\\ \phi\\ \varPhi\\ \varPhi\end{pmatrix}\begin{pmatrix}\bar{\phi} & \bar{\phi} & \bar{\phi} & \bar{\varPhi} & \bar{\varPhi}\end{pmatrix}=\begin{pmatrix}\phi\bar{\phi} & \phi\bar{\phi} & \phi\bar{\phi} & \phi\bar{\varPhi} & \phi\bar{\varPhi}\\ \phi\bar{\phi} & \phi\bar{\phi} & \phi\bar{\phi} & \phi\bar{\varPhi} & \phi\bar{\varPhi}\\ \phi\bar{\phi} & \phi\bar{\phi} & \phi\bar{\phi} & \phi\bar{\varPhi} & \phi\bar{\varPhi}\\ \varPhi\bar{\phi} & \varPhi\bar{\phi} & \varPhi\bar{\phi} & \varPhi\bar{\varPhi} & \varPhi\bar{\varPhi}\\ \varPhi\bar{\phi} & \varPhi\bar{\phi} & \varPhi\bar{\phi} & \varPhi\bar{\varPhi} & \varPhi\bar{\varPhi}\end{pmatrix}\equiv\mathbb{P}_a$$

When written out in detail

$$\begin{pmatrix}\phi\bar{\phi} & \phi\bar{\phi} & \phi\bar{\phi}\\ \phi\bar{\phi} & \phi\bar{\phi} & \phi\bar{\phi}\\ \phi\bar{\phi} & \phi\bar{\phi} & \phi\bar{\phi}\end{pmatrix}=\begin{pmatrix}\phi_{a1}\bar{\phi}_{a1} & \phi_{a1}\bar{\phi}_{a2} & \cdots & \phi_{a1}\bar{\phi}_{an}\\ \phi_{a2}\bar{\phi}_{a1} & \phi_{a2}\bar{\phi}_{a2} & \cdots & \phi_{a2}\bar{\phi}_{an}\\ \vdots & \vdots & \ddots & \vdots\\ \phi_{an}\bar{\phi}_{a1} & \phi_{an}\bar{\phi}_{a2} & \cdots & \phi_{an}\bar{\phi}_{an}\end{pmatrix}=\tilde{\mathbb{P}}_a\,!$$

The implication is that if we write

$$|\psi) = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_m \end{pmatrix} \quad \text{and} \quad |\Psi) = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_m \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

then the results of N-element POVM measurements in $\mathcal{H}_n$ can be realized as PVM measurements in $\mathcal{H}_N$ ($N > n$), which is the upshot of Neumark's theorem.

Preskill and Flory provide instructive simple illustrations of Neumark's theorem in action. I will discuss variants of both in the next two sections, but preface that discussion with review of some general principles available for the construction of POVMs on $\mathcal{H}_2$.

POVMs for generalized qubit measurements. It is well known[17] that the most general traceless 2×2 hermitian matrix can be described

$$\mathbb{H} = h_1\boldsymbol{\sigma}_1 + h_2\boldsymbol{\sigma}_2 + h_3\boldsymbol{\sigma}_3 = \begin{pmatrix} h_3 & h_1 - ih_2 \\ h_1 + ih_2 & -h_3 \end{pmatrix}$$

The eigenvalues of such a matrix are $\pm\sqrt{h_1^2 + h_2^2 + h_3^2}$, and become ± 1 if

$$\mathbf{h} = \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix} \quad \text{is a unit vector: } \mathbf{h}\cdot\mathbf{h}=1$$

The associated orthonormal eigenvectors

$$\mathbb{H}\,|\mathbf{h}_\pm) = \pm|\mathbf{h}_\pm)$$

can (to within uninteresting phase factors) be described

$$|\mathbf{h}_\pm) = \begin{pmatrix} \sqrt{\frac{1\pm h_3}{2}} \\ \pm\sqrt{\frac{1}{2(1\pm h_3)}}(h_1 + ih_2) \end{pmatrix}$$

$$|\mathbf{h}_+) = \begin{pmatrix} \cos\frac{1}{2}\theta \\ +\sin\frac{1}{2}\theta\cdot e^{i\phi} \end{pmatrix}, \quad |\mathbf{h}_-) = \begin{pmatrix} \sin\frac{1}{2}\theta \\ -\cos\frac{1}{2}\theta\cdot e^{i\phi} \end{pmatrix}$$

[17] For detailed arguments see *Advanced Quantum Topics* (2000), Chapter 1, pages 2–5.

We have particular interest at the moment in the associated projectors, which read

$$\mathbb{P}_+(\mathbf{h}) = |\mathbf{h}_+)(\mathbf{h}_+| = \begin{pmatrix} \frac{1}{2}(1 + h_3) & +\frac{1}{2}(h_1 - ih_2) \\ +\frac{1}{2}(h_1 + ih_2) & \frac{1}{2}(1 - h_3) \end{pmatrix} = \tfrac{1}{2}(\mathbb{I} + \mathbf{h}\cdot\boldsymbol{\sigma})$$

$$= \begin{pmatrix} \cos^2\frac{1}{2}\theta & +\frac{1}{2}e^{-i\phi}\sin\theta \\ +\frac{1}{2}e^{+i\phi}\sin\theta & \sin^2\frac{1}{2}\theta \end{pmatrix}$$

$$\mathbb{P}_-(\mathbf{h}) = |\mathbf{h}_-)(\mathbf{h}_-| = \begin{pmatrix} \frac{1}{2}(1 - h_3) & -\frac{1}{2}(h_1 - ih_2) \\ -\frac{1}{2}(h_1 + ih_2) & \frac{1}{2}(1 + h_3) \end{pmatrix} = \tfrac{1}{2}(\mathbb{I} - \mathbf{h}\cdot\boldsymbol{\sigma})$$

$$= \begin{pmatrix} \sin^2\frac{1}{2}\theta & -\frac{1}{2}e^{-i\phi}\sin\theta \\ -\frac{1}{2}e^{+i\phi}\sin\theta & \cos^2\frac{1}{2}\theta \end{pmatrix}$$

Both are manifestly positive hermitian (as is made obvious also from their shared spectra: $\{1,0\}$), and both have unit traces (as indeed they must, since they project onto 1-spaces). Collectively they are orthogonal and manifestly complete. And they are very simply related:

$$\mathbb{P}_-(\mathbf{h}) = \mathbb{P}_+(-\mathbf{h})$$

The completeness relation can therefore be written

$$\mathbb{P}_+(\mathbf{h}) + \mathbb{P}_+(-\mathbf{h}) = \mathbb{I}$$

which indicates that $\{\mathbb{P}_+(\mathbf{h}), \mathbb{P}_+(-\mathbf{h})\}$ might be considered to comprise a 2-element "qubit POVM." But $\mathbb{P}_+(\pm\mathbf{h})$ are projective, so $\{\mathbb{P}_+(\mathbf{h}), \mathbb{P}_+(-\mathbf{h})\}$ is actually a 2-element "qubit PVM" (degenerate POVM).

We are led to ask "For what unit vectors $\mathbf{a}$ and $\mathbf{b}$ is $\{\mathbb{P}_+(\mathbf{a}), \mathbb{P}_+(\mathbf{b})\}$ a (non-degenerate) POVM?" From

$$\mathbb{P}_+(\mathbf{a}) + \mathbb{P}_+(\mathbf{b}) = \begin{pmatrix} \frac{1}{2}(2 + [a_3 + b_3]) & \frac{1}{2}([a_1 + b_1] - i[a_2 + b_2]) \\ \frac{1}{2}([a_1 + b_1] + i[a_2 + b_2]) & \frac{1}{2}(2 - [a_3 + b_3]) \end{pmatrix}$$

$$= \mathbb{I} \quad \text{iff} \quad \mathbf{a} + \mathbf{b} = \mathbf{0}$$

we conclude that *all* 2-element qubit POVMs are actually PVMs. How about constructions of the form $\mathbb{P}_+(\mathbf{a}) + \mathbb{P}_+(\mathbf{b}) + \mathbb{P}_+(\mathbf{c})$? Arguing as before, we see that

$$\mathbb{P}_+(\mathbf{a}) + \mathbb{P}_+(\mathbf{b}) + \mathbb{P}_+(\mathbf{c}) = \tfrac{3}{2}\mathbb{I} \quad \iff \quad \mathbf{a} + \mathbf{b} + \mathbf{c} = \mathbf{0}$$

so $\big\{\frac{2}{3}\mathbb{P}_+(\mathbf{a}), \frac{2}{3}\mathbb{P}_+(\mathbf{b}), \frac{2}{3}\mathbb{P}_+(\mathbf{c})\big\}$—the elements of which are positive hermitian but *not* projective—constitutes a 3-element qubit POVM, a qubit POVM with the least possible number of elements. Similarly,

$$\big\{\tfrac{2}{5}\mathbb{P}_+(\mathbf{a}), \tfrac{2}{5}\mathbb{P}_+(\mathbf{b}), \tfrac{2}{5}\mathbb{P}_+(\mathbf{c}), \tfrac{2}{5}\mathbb{P}_+(\mathbf{d}), \tfrac{2}{5}\mathbb{P}_+(\mathbf{e})\big\}$$

—here $\{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}\}$ are unit 3-vectors subject to the constraint

$$\mathbf{a}+\mathbf{b}+\mathbf{c}+\mathbf{d}+\mathbf{e}=\mathbf{0}$$

—consitutes a 5-element qubit POVM. It is interesting in this connection to notice that

- $\mathbf{a}+\mathbf{b}=\mathbf{0}$ forces the vectors to be colinear;
- $\mathbf{a}+\mathbf{b}+\mathbf{c}=\mathbf{0}$ forces the unit vectors to be coplanar, and to bound an equilateral triangle (a rigid structure);
- $\mathbf{a}+\mathbf{b}+\mathbf{c}+\mathbf{d}=\mathbf{0}$ forces the unit vectors to bound a parallelogram, but coplanarity and rigidity are both relaxed;
- $\mathbf{a}+\mathbf{b}+\mathbf{c}+\mathbf{d}+\cdots+\mathbf{z}=\mathbf{0}$ imposes a condition so relaxed that the vectors can articulate space curves (knots) of arbitrary complexity.

On can—while retaining

$$\sum_{i=1}^{n}\mathbf{a}_i=\mathbf{0}$$

—relax the requirement that the vectors $\mathbf{a}_i$ be unit vectors and still have

$$\sum_{i=1}^{n}\tfrac{2}{n}\mathbb{P}_+(\mathbf{a}_i)=\mathbb{I}$$

But since the eigenvalues of $\mathbb{P}_+(\mathbf{a})$ are $\lambda_\pm=\frac{1}{2}(1\pm\sqrt{\mathbf{a}\cdot\mathbf{a}})$ one must require that all $\mathbf{a}_i$ have length $\sqrt{\mathbf{a}_i\cdot\mathbf{a}_i}\le 1$ to preserve the positivity of the POVM elements $\tilde{\mathbb{P}}_i=\frac{2}{n}\mathbb{P}_+(\mathbf{a}_i)$.

Relatedly, let $\mathbf{a}$ be a unit vector and k a positive real number. Then the "dilated projector"

$$\mathbb{K}(\mathbf{a},k)=k\mathbb{P}_+(\mathbf{a})=\tfrac{1}{2}(k\mathbb{I}+k\mathbf{a}\cdot\boldsymbol{\sigma})$$

has eigenvalues $\lambda_\pm=\frac{1}{2}(k\pm k)$, so is positive hermitian. Immediately

$$\sum_i\mathbb{K}(\mathbf{a}_i,k_i)=\tfrac{k}{2}\,\mathbb{I}+\big(\textstyle\sum_i k_i\mathbf{a}_i\big)\cdot\boldsymbol{\sigma}\quad\text{where}\quad k=\textstyle\sum_i k_i$$

so the matrices $\tilde{\mathbb{P}}_i=\frac{2}{k}\mathbb{K}(\mathbf{a}_i,k_i)$ become elements of a POVM when $\sum_i k_i\mathbf{a}_i=\mathbf{0}$.

Relaxation of the unit length constraint permits an additional mode of POVM construction. Introduce hermitian matrices

$$\mathbb{Q}_i=\mathbb{Q}(\mathbf{a}_i,a_i)=\tfrac{1}{2}a_i\mathbb{I}+\mathbb{P}_+(\mathbf{a}_i)$$

where the assumption $\sum_i\mathbf{a}_i=\mathbf{0}$ remains in force and the a_i are real numbers of either sign. Then

$$\sum_{i=1}^{n}\mathbb{Q}_i=\tfrac{1}{2}(a+n)\mathbb{I}\quad\text{with}\quad a=\textstyle\sum_i a_i$$

The eigenvalues of $\mathbb{Q}=\frac{1}{2}a\mathbb{I}+\mathbb{P}_+(\mathbf{a})$ are $\lambda_\pm=\frac{1}{2}(1+a\pm\sqrt{\mathbf{a}\cdot\mathbf{a}})$ so to ensure positivity we must impose the constraints $a_i\ge\sqrt{\mathbf{a}_i\cdot\mathbf{a}_i}-1$ (all i). That done, we have an n-element qubit POVM $\big\{\frac{2}{a+n}\mathbb{Q}_1,\frac{2}{a+n}\mathbb{Q}_2,\ldots,\frac{2}{a+n}\mathbb{Q}_n\big\}$.

Preskill's Example

Preskill takes from his demo kit a 3-element qubit PVOM of the simplest possible construction: $\{\tilde{\mathbb{P}}_1, \tilde{\mathbb{P}}_2, \tilde{\mathbb{P}}_3\} = \{\frac{2}{3}\mathbb{P}_+(\mathbf{a}), \frac{2}{3}\mathbb{P}_+(\mathbf{b}), \frac{2}{3}\mathbb{P}_+(\mathbf{c})\}$ where the unit vectors $\mathbf{a}+\mathbf{b}+\mathbf{c}=\mathbf{0}$ and therefore mark the vertices of an equilateral triangle. Specifically, he works from

$$\begin{pmatrix}\sin\theta\cos\phi\\ \sin\theta\sin\phi\\ \cos\theta\end{pmatrix} \quad \text{with} \quad \theta = \{0, \tfrac{2\pi}{3}, \tfrac{4\pi}{3}\},\ \phi = 0$$

to obtain

$$\mathbf{a} = \begin{pmatrix}0\\0\\1\end{pmatrix},\ \mathbf{b} = \begin{pmatrix}+\frac{\sqrt{3}}{2}\\0\\-\frac{1}{2}\end{pmatrix},\ \mathbf{c} = \begin{pmatrix}-\frac{\sqrt{3}}{2}\\0\\-\frac{1}{2}\end{pmatrix}$$

giving

$$\tilde{\mathbb{P}}_1 = \begin{pmatrix}\frac{2}{3} & 0\\ 0 & 0\end{pmatrix},\ \tilde{\mathbb{P}}_2 = \begin{pmatrix}\frac{1}{6} & +\frac{1}{2\sqrt{3}}\\ +\frac{1}{2\sqrt{3}} & \frac{1}{2}\end{pmatrix},\ \tilde{\mathbb{P}}_2 = \begin{pmatrix}\frac{1}{6} & -\frac{1}{2\sqrt{3}}\\ -\frac{1}{2\sqrt{3}} & \frac{1}{2}\end{pmatrix}$$

which clearly sum to $\mathbb{I}$. Those matrices (which are non-projective because of the $\frac{2}{3}$-factors which entered into their definitions) can be developed

$$\begin{aligned}
\tilde{\mathbb{P}}_1 &= |a_+)(a_+| \quad \text{with} \quad |a_+) = \sqrt{\tfrac{2}{3}}\begin{pmatrix}\cos(0/2)\\ \sin(0/2)\end{pmatrix} = \begin{pmatrix}\sqrt{\frac{2}{3}}\\ 0\end{pmatrix}\\
\tilde{\mathbb{P}}_2 &= |b_+)(b_+| \quad \text{with} \quad |b_+) = \sqrt{\tfrac{2}{3}}\begin{pmatrix}\cos(\frac{2\pi}{3}/2)\\ \sin(\frac{2\pi}{3}/2)\end{pmatrix} = \begin{pmatrix}+\frac{1}{\sqrt{6}}\\ \frac{1}{\sqrt{2}}\end{pmatrix}\\
\tilde{\mathbb{P}}_3 &= |c_+)(c_+| \quad \text{with} \quad |c_+) = \sqrt{\tfrac{2}{3}}\begin{pmatrix}\cos(\frac{4\pi}{3}/2)\\ \sin(\frac{4\pi}{3}/2)\end{pmatrix} = \begin{pmatrix}-\frac{1}{\sqrt{6}}\\ \frac{1}{\sqrt{2}}\end{pmatrix}
\end{aligned}$$

Spreading the top elements of those 2-vectors along the top row of a 3×3 matrix, and the bottom elements along the second row, we obtain

$$\begin{pmatrix}\sqrt{\frac{2}{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}}\\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ x & y & z\end{pmatrix}$$

The top rows are clearly orthonormal, and we fix the values of $\{x, y, z\}$ by Gram-Schmidt; *i.e.*, by solving the system

$$\begin{aligned}
\sqrt{\tfrac{2}{3}}\,x - \tfrac{1}{\sqrt{6}}\,y + \tfrac{1}{\sqrt{6}}\,z &= 0\\
\tfrac{1}{\sqrt{2}}\,y + \tfrac{1}{\sqrt{2}}\,z &= 0\\
x^2 + y^2 + z^2 &= 1
\end{aligned}$$

and obtain

$$\mathbb{U} = \begin{pmatrix} \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

Let the (automatically orthnormal) columns of $\mathbb{U}$ be denoted $\{|E_1), |E_2), |E_3)\}$. The projectors onto those vectors are

$$\mathbb{P}_1 = |E_1)(E_1| = \begin{pmatrix} \frac{2}{3} & 0 & \frac{\sqrt{2}}{3} \\ 0 & 0 & 0 \\ \frac{\sqrt{2}}{3} & 0 & \frac{1}{3} \end{pmatrix}$$

$$\mathbb{P}_2 = |E_2)(E_2| = \begin{pmatrix} \frac{1}{6} & \frac{1}{2\sqrt{3}} & \frac{1}{3\sqrt{2}} \\ \frac{1}{2\sqrt{3}} & \frac{1}{2} & \frac{1}{\sqrt{6}} \\ \frac{1}{3\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{3} \end{pmatrix}$$

$$\mathbb{P}_3 = |E_3)(E_3| = \begin{pmatrix} \frac{1}{6} & -\frac{1}{2\sqrt{3}} & -\frac{1}{3\sqrt{2}} \\ -\frac{1}{2\sqrt{3}} & \frac{1}{2} & \frac{1}{\sqrt{6}} \\ -\frac{1}{3\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{3} \end{pmatrix}$$

where the red submatrices are seen to be precisely the elements $\{\tilde{\mathbb{P}}_1, \tilde{\mathbb{P}}_2, \tilde{\mathbb{P}}_3\}$ of the qubit POVM which provided Preskill with his point of departure. Writing

$$|\psi) = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \quad \text{and} \quad |\Psi) = \begin{pmatrix} \psi_1 \\ \psi_2 \\ 0 \end{pmatrix}$$

we have

$$(\psi|\tilde{\mathbb{P}}_i|\psi) = (\Psi|\mathbb{P}_i|\Psi) \quad : \quad i = 1, 2, 3$$

We have, in Preskill's phrase, *realized POVM measurements on a qubit as PVM measurements on a "tribit."*

Theory of qubit discrimination. Suppose Alice sends Bob one or the other of a pair of non-orthogonal states $\{|\alpha), |\beta)\}$ in random sequence (the sequence may convey an encoded message). Bob, by prearrangement, knows the identity of the states Alice intends to employ for this purpose. His assignment is to "read the message" as best he can. PVM measurement will not suffice, since $|\alpha)$ and $|\beta)$—because of their non-orthogonality—cannot be simultaneous eigenvectors of a hermitian matrix $\mathbb{A}$.[18] So Bob looks to see what might be accomplished by POVM measurement. Some inspired tinkering leads him to to construct

$$\mathbb{E}_1 = \mathbb{I} - |\alpha)(\alpha|, \quad \mathbb{E}_2 = \mathbb{I} - |\beta)(\beta|, \quad \mathbb{E}_3 = \mathbb{I} - \mathbb{E}_1 - \mathbb{E}_2$$

[18] Barnett (page 99) provides an elegant formal proof (if proof be needed) of this important fact.

These matrices are manifestly hermitian, and the first pair—which might be written $\mathbb{E}_1 = |\alpha_\perp)(\alpha_\perp|$, $\mathbb{E}_2 = |\beta_\perp)(\beta_\perp|$—are clearly positive. But $\mathbb{E}_3$ is *not* positive; the eigenvalues of $\mathbb{E}_3$ are[19] $\lambda_{3\pm} = \pm|(\alpha|\beta)|$. Bob (after some more inspired tinkering) is led thus to the constructions

$$\begin{aligned}\mathbb{F}_1 &= k\,\mathbb{E}_1\\ \mathbb{F}_2 &= k\,\mathbb{E}_2\\ \mathbb{F}_3 &= \mathbb{I} - \mathbb{F}_1 - \mathbb{F}_2\end{aligned}$$

the spectra of which are $\{k, 0\}$, $\{k, 0\}$ and $\{1-(1-|(\alpha|\beta)|)k,\ 1-(1+|(\alpha|\beta)|)k\}$ respectively. To achieve the positivity required for $\{\mathbb{F}_1, \mathbb{F}_2, \mathbb{F}_3\}$ to comprise a POVM we are obliged to impose upon k the following constraints:

$$0 < k \leq k_{\max} = \frac{1}{1 + |(\alpha|\beta)|}$$

Each time $|\alpha)$ else $|\beta)$ is presented to Bob's meter one or another of its display lights flashes, with probabilities provided by the following table:

state	#1	#2	#3
$\vert\alpha)$	$(\alpha\vert\tilde{\mathbb{F}}_1\vert\alpha)$	$(\alpha\vert\tilde{\mathbb{F}}_3\vert\alpha)$	$(\alpha\vert\tilde{\mathbb{F}}_3\vert\alpha)$
$\vert\beta)$	$(\beta\vert\tilde{\mathbb{F}}_1\vert\beta)$	$(\beta\vert\tilde{\mathbb{F}}_2\vert\beta)$	$(\beta\vert\tilde{\mathbb{F}}_3\vert\beta)$

So long as the multiplier k remains unspecified, we have

state	#1	#2	#3
$\vert\alpha)$	0	$k(1-x^2)$	$1-k(1-x^2)$
$\vert\beta)$	$k(1-x^2)$	0	$1-k(1-x^2)$

with $x = |(\alpha|\beta)|$. If Alice were sending orthogonal states to Bob (case $x = 0$) then the probability table would assume the form

state	#1	#2	#3
$\vert\alpha)$	0	k	$1-k$
$\vert\beta)$	k	0	$1-k$

with this consequence: when #1 flashes Bob knows he has received a $|\beta)$, when #2 flashes he knows he has received an $|\alpha)$, when #3 flashes he knows that he might with equal probability have received either $|\alpha)$ or $|\beta)$. This circumstance introduces confusion/noise into Bob's transcription of Alice's message. If Bob tweaks his meter, setting $k = k_{\max} = \frac{1}{1+x} \to 1$ at $x = 0$ the table becomes

[19] This I know only on the basis of numerical experimentation; I have not yet managed to construct a formal demonstration.

state	#1	#2	#3
$\vert\alpha)$	0	1	0
$\vert\beta)$	1	0	0

#3 now never flashes; Bob is able to read Alice's message with perfect fidelity; the PVOM meter has become a PVM meter.

If—more interestingly—Alice dispatches non-orthogonal states ($0 < x$) to Bob then it remains the case that when #1 flashes Bob knows he has received a $|\beta)$, when #2 flashes he knows he has received an $|\alpha)$, when #3 flashes he knows that he might with equal probability have received either $|\alpha)$ or $|\beta)$. The latter circumstance occurs with probability

$$\text{Prob}_? = 1 - k(1 - x^2)$$

Bob sets $k \to k_{\max} = \frac{1}{1+x}$ to minimize the likelihood of such uninformative events, and obtains

$$\begin{aligned}\text{Prob}_? &= x\\ &\downarrow\\ &= \begin{cases} 1 \text{ as } |\alpha) \text{ and } |\beta) \text{ become parallel (indistinguishable)}\\ 0 \text{ as } |\alpha) \text{ and } |\beta) \text{ become orthogonal}\end{cases}\end{aligned}$$

This result was first obtained I. D. Ivanovic, D. Dieks and A. Peres, working independently (1987 & 1988), and is known as the "IDP limit."

The short of it: imperfect (PVOM) meters can be designed to exploit facts known *a priori* (Bob's knowledge of the states Alice has elected to employ) to provide information that remains forever beyond the reach of perfect (PVM) meters.

Suppose, to make matters more concrete, that Alice is shipping *qubits* to Bob (we have here pulled back from $\mathcal{H}_n$ to $\mathcal{H}_2$), in states[20]

$$|\alpha) = \begin{pmatrix}\cos u\\ \sin u\end{pmatrix}, \quad |\beta) = \begin{pmatrix}\cos v\\ \sin v\end{pmatrix}$$

Straightforward calculation then supplies

$$\begin{aligned}\mathbb{F}_1 &= \begin{pmatrix} k\sin^2 u & -k\cos u \sin u\\ -k\cos u\sin u & k\cos^2 u\end{pmatrix}\\ \mathbb{F}_2 &= \begin{pmatrix} k\sin^2 v & -k\cos v \sin v\\ -k\cos v\sin v & k\cos^2 v\end{pmatrix}\\ \mathbb{F}_3 &= \mathbb{I} - \mathbb{F}_1 - \mathbb{F}_2\end{aligned}$$

[20] To achieve complete generality we would have to set

$$|\alpha) = \begin{pmatrix}\cos u\\ e^{i\xi}\sin u\end{pmatrix}, \quad |\beta) = \begin{pmatrix}\cos v\\ e^{i\zeta}\sin v\end{pmatrix}$$

I have purchased some relative simplicity by setting $\xi = \zeta = 0$.

which by appeal to $\frac{1}{2}\text{tr}\boldsymbol{\sigma}_i\boldsymbol{\sigma}_j = \delta_{ij} : \{i,j\} \in \{0,1,2,3\}$ assume the form

$$\begin{aligned}\mathbb{F}_1 &= a\mathbb{I} + \mathbf{a}\cdot\boldsymbol{\sigma}\\ \mathbb{F}_2 &= b\mathbb{I} + \mathbf{b}\cdot\boldsymbol{\sigma}\\ \mathbb{F}_3 &= c\mathbb{I} + \mathbf{c}\cdot\boldsymbol{\sigma}\end{aligned}$$

with $a = b = \frac{1}{2}k,\ c = 1 - k$ and

$$\begin{aligned}\mathbf{a} &= \tfrac{1}{2}k\begin{pmatrix}-\sin 2u\\ 0\\ -\cos 2u\end{pmatrix}\\ \mathbf{b} &= \tfrac{1}{2}k\begin{pmatrix}-\sin 2v\\ 0\\ -\cos 2v\end{pmatrix}\\ \mathbf{c} &= \tfrac{1}{2}k\begin{pmatrix}\sin 2u + \sin 2v\\ 0\\ \cos 2u + \cos 2v\\ 0\end{pmatrix}\end{aligned}$$

Gratifyingly, we have

$$\begin{aligned}a + b + c &= 1\\ \mathbf{a} + \mathbf{b} + \mathbf{c} &= \mathbf{0}\end{aligned}$$

as required. The probability table has become

state	#1	#2	#3
$\vert\alpha)$	0	$k\sin^2(u-v)$	$1 - k\sin^2(u-v)$
$\vert\beta)$	$k\sin^2(u-v)$	0	$1 - k\sin^2(u-v)$

which when optimized

$$k \rightarrow k_{\max} = \frac{1}{1+x} \quad \text{with} \quad x = |(\alpha|\beta)| = |\cos(u-v)|$$

becomes

state	#1	#2	#3
$\vert\alpha)$	0	$1 - \vert\cos(u-v)\vert$	$\vert\cos(u-v)\vert$
$\vert\beta)$	$1 - \vert\cos(u-v)\vert$	0	$\vert\cos(u-v)\vert$

giving $\text{Prob}_? = |\cos(u-v)|$, which obligingly vanishes when $u - v = \pm\pi/2$.

It is instructive to examine the Neumarkian implications of the preceding material. The spectral decompostions of $\mathbb{F}_1$ and $\mathbb{F}_2$ can be written

$$\mathbb{F}_1 = |\phi_1)(\phi_1| \quad \text{with} \quad |\phi_1) = \begin{pmatrix} +\sqrt{k}\,\sin u \\ -\sqrt{k}\,\cos u \end{pmatrix}$$
$$\mathbb{F}_2 = |\phi_2)(\phi_2| \quad \text{with} \quad |\phi_2) = \begin{pmatrix} +\sqrt{k}\,\sin v \\ -\sqrt{k}\,\cos v \end{pmatrix}$$

but the decomposition of $\mathbb{F}_3$ is bipartite (and therefore inconsistent with an assumption fundamental to the argument pursued on pages 16–18) unless we set $k = k_{\text{max}}$, thus killing the lesser of $\mathbb{F}_3$'s eigenvalues. At this point the argument bifurcates, partly because *Mathematica*'s `FullSimplify` command is frustrated by the occurance of absolute value bars in expressions it is asked to to process but mainly for a deeper reason. We adopt the following work-around

$$k_{\text{max}} = \frac{1}{1 + |\cos(u-v)|} \longrightarrow \begin{cases} \dfrac{1}{1+\cos(u-v)} & : \quad \cos(u-v) > 0 \\ \dfrac{1}{1-\cos(u-v)} & : \quad \cos(u-v) < 0 \end{cases}$$

and will find that in some instances the two cases must be carefully distinguished, which I will do by introducing subscripts $_p$ and $_m$. It proves useful in this context to notice that

$$\frac{1}{1+\cos(u-v)} = \tfrac{1}{2}\sec^2 \tfrac{1}{2}(u-v)$$
$$\frac{1}{1-\cos(u-v)} = \tfrac{1}{2}\csc^2 \tfrac{1}{2}(u-v)$$

Thus do we obtain the factorization statements

$$\mathbb{F}_{1p} = \tfrac{1}{2}\sec^2 \tfrac{1}{2}(u-v)|\alpha)(\alpha| = |\phi_{1p})(\phi_{1p}|$$
$$\mathbb{F}_{1m} = \tfrac{1}{2}\csc^2 \tfrac{1}{2}(u-v)|\alpha)(\alpha| = |\phi_{1m})(\phi_{1m}|$$

$$\mathbb{F}_{2p} = \tfrac{1}{2}\sec^2 \tfrac{1}{2}(u-v)|\beta)(\beta| = |\phi_{2p})(\phi_{2p}|$$
$$\mathbb{F}_{2m} = \tfrac{1}{2}\csc^2 \tfrac{1}{2}(u-v)|\beta)(\beta| = |\phi_{2m})(\phi_{2m}|$$

$$\mathbb{F}_{3p} = \mathbb{I} - \mathbb{F}_{1p} - \mathbb{F}_{2p} = |\phi_{3p})(\phi_{3p}|$$
$$\mathbb{F}_{3m} = \mathbb{I} - \mathbb{F}_{1m} - \mathbb{F}_{2m} = |\phi_{3p})(\phi_{3p}|$$

where

$$|\phi_{1p}) = \tfrac{1}{\sqrt{2}}\sec \tfrac{1}{2}(u-v)\cdot|\alpha)$$
$$|\phi_{1m}) = \tfrac{1}{\sqrt{2}}\csc \tfrac{1}{2}(u-v)\cdot|\alpha)$$

$$|\phi_{2p}) = \tfrac{1}{\sqrt{2}}\sec \tfrac{1}{2}(u-v)\cdot|\beta)$$
$$|\phi_{2m}) = \tfrac{1}{\sqrt{2}}\csc \tfrac{1}{2}(u-v)\cdot|\beta)$$

are immediate, while some fairly heavy calculatinon supplies results

$$
\begin{aligned}
|\phi_{3p}) &= \sqrt{+\cos(u-v)}\,\sec\tfrac{1}{2}(u-v)\begin{pmatrix}+\cos\tfrac{1}{2}(u+v)\\ +\sin\tfrac{1}{2}(u+v)\end{pmatrix}\\
|\phi_{3m}) &= \sqrt{-\cos(u-v)}\,\csc\tfrac{1}{2}(u-v)\begin{pmatrix}-\sin\tfrac{1}{2}(u+v)\\ +\cos\tfrac{1}{2}(u+v)\end{pmatrix}
\end{aligned}
$$

that are much more easily checked than derived. Feeding that data into

$$
\mathbb{U}_p = \begin{pmatrix}\phi_{1p,1} & \phi_{2p,1} & \phi_{3p,1}\\ \phi_{1p,2} & \phi_{2p,2} & \phi_{3p,21}\\ x_p & y_p & z_p\end{pmatrix},\quad \mathbb{U}_m = \begin{pmatrix}\phi_{1m,1} & \phi_{2m,1} & \phi_{3m,1}\\ \phi_{1m,2} & \phi_{2m,2} & \phi_{3m,21}\\ x_m & y_m & z_m\end{pmatrix}
$$

we discover that to complete the orthonormality of the rows we must (to within a shared phase factor) set

$$
\begin{aligned}
x_p = -y_p &= +\tfrac{1}{\sqrt{2}}\sqrt{+\cos(u-v)}\,\sec\tfrac{1}{2}(u-v)\\
z_p &= \tan\tfrac{1}{2}(u-v)
\end{aligned}
$$

$$
\begin{aligned}
x_m = +y_m &= -\tfrac{1}{\sqrt{2}}\sqrt{-\cos(u-v)}\,\csc\tfrac{1}{2}(u-v)\\
z_m &= \cot\tfrac{1}{2}(u-v)
\end{aligned}
$$

Orthonormal Neumark bases $\{|E_{1p}), |E_{2p}), |E_{3p})\}$ and $\{|E_{1m}), |E_{2m}), |E_{3m})\}$ are read from the columns of the unitary matrices $\mathbb{U}_p$ and $\mathbb{U}_m$. One needs $\{x_p, y_p, z_p\}$ and $\{x_m, y_m, z_m\}$ to lend detailed substance to Neumark's POMV→PVM demonstration, but it is clear even in the absence of that information that the associated projectors $\mathbb{P}_i = |E_i)(E_i|$ possess the structure

$$
\mathbb{P}_i = \begin{pmatrix}\tilde{\mathbb{F}}_i & \bullet\\ \bullet & \bullet\end{pmatrix} \quad : \quad \bullet\text{-terms are } \{x, y, z\}\text{-dependent}
$$

and that is all one needs to obtain

$$
(\psi|\tilde{\mathbb{F}}_i|\psi) = (\Psi|\mathbb{P}_i|\Psi) \quad : \quad |\Psi) = \begin{pmatrix}|\psi)\\ 0\end{pmatrix}
$$

In the preceding section I developed and illustrated a "general theory of POVMs for qubit measurements" which proceeds from specification of 3-vectors that satisfy the closure condition $\mathbf{a} + \mathbf{b} + \cdots + \mathbf{z} = \mathbf{0}$. In the discussion just concluded I have described an alternative formulation of that theory (at least in so far as it relates to 3-element POVMs) which—more appropriately for application to the state-discrimination problem—proceeds from specification of non-orthogonal qubits $|\alpha)$ and $|\beta)$.

If we set $u = \frac{\pi}{2} + \frac{2\pi}{3}$ and $v = \frac{\pi}{2}$ (which entail $\cos(u - v) < 0$) we obtain matrices

$$\mathbb{F}_{1m} = \begin{pmatrix} \frac{1}{6} & -\frac{1}{2\sqrt{3}} \\ -\frac{1}{2\sqrt{3}} & \frac{1}{2} \end{pmatrix}, \quad \mathbb{F}_{2m} = \begin{pmatrix} \frac{2}{3} & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathbb{F}_{3m} = \begin{pmatrix} \frac{1}{6} & \frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} & \frac{1}{2} \end{pmatrix}$$

which were encountered already on page 21, and find that we can use our general formulae to reproduce all the details of Preskill's "equilateral" example.

If we set $u = \frac{3\pi}{2}$ and $v = \frac{3\pi}{2} + \frac{3\pi}{4}$ (which again entail $\cos(u - v) < 0$) we obtain

$$k = k_{\max} = \frac{1}{1 - \cos(u - v)} = \frac{\sqrt{2}}{1 + \sqrt{2}}$$

$$\mathbb{F}_{1m} = \begin{pmatrix} \frac{\sqrt{2}}{1+\sqrt{2}} & 0 \\ 0 & 0 \end{pmatrix} = |\phi_{1m})(\phi_{1m}| \quad \text{with} \quad |\phi_{1m}) = \sqrt{\tfrac{\sqrt{2}}{1+\sqrt{2}}}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\mathbb{F}_{2m} = \begin{pmatrix} \frac{1}{2+\sqrt{2}} & \frac{-1}{2+\sqrt{2}} \\ \frac{-1}{2+\sqrt{2}} & \frac{-1}{2+\sqrt{2}} \end{pmatrix} = |\phi_{2m})(\phi_{2m}| \quad \text{with} \quad |\phi_{2m}) = \sqrt{\tfrac{\sqrt{2}}{2(1+\sqrt{2})}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\mathbb{F}_{3m} = \begin{pmatrix} \frac{3-2\sqrt{2}}{\sqrt{2}} & \frac{\sqrt{2}-1}{\sqrt{2}} \\ \frac{\sqrt{2}-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = |\phi_{3m})(\phi_{3m}| \quad \text{with} \quad |\phi_{3m}) = \tfrac{1}{\sqrt{\sqrt{2}}}\begin{pmatrix} \sqrt{2} - 1 \\ 1 \end{pmatrix}$$

and find the Pauli coordinates of the $\mathbb{F}$-matrices to be given in this instance by

$$\{a, b, c\} = \left\{\tfrac{1}{2}k, \tfrac{1}{2}k, 1 - k\right\}, \quad \{\mathbf{a}, \mathbf{b}, \mathbf{c}\} = \tfrac{1}{2}k\left\{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\right\}$$

Again, $a + b + c = 1$ and $\mathbf{a} + \mathbf{b} + \mathbf{c} = \mathbf{0}$. The vectors in this instance describe a right isosceles triangle. We have here demonstrated that our general formulae do efficiently reproduce the essentials of the example that Mario Flory details in §2.4 of the essay previously cited.[21] It was, by the way, that example which inspired the preceding discussion.

The question arises: How might Bob undertake to distinguish (optimally) amongst three or more known qubits? This difficult question is addressed (inconclusively) on Barnett's page 103.[22] The theory of optimal state discrimination was pioneered by C. W. Helstrom (1976) and A. S. Holevo (1982).

General theory of density matrix transformations: Kraus operations. System-meter interactions (except those involving prompt repetition with an ideal meter) induce non-trivial state tansformations $|\psi)_{\text{before}} \longrightarrow |\psi)_{\text{after}}$, whence

[21] It can be found at https://wiki.physik.uni-muenchen.de/TMP.images/8/87/POVMs.pdf.

[22] For more detailed (and exceptionally lucid) discussion see §3.2 in A. Chefles, "Quantum state discrimination," Contemporary Physics **41**,401-424 (2000.)

adjustments in the construction of ensembles of states

$$\rho_{\text{before}} \longrightarrow \rho_{\text{after}}$$

So also (at least in the Schrödinger picture) does unobserved quantum motion, which for isolated systems $\mathbb{S}$ is unitary

$$\begin{aligned} |\psi)_0 &\longrightarrow |\psi)_t = \mathbb{U}(t)|\psi)_0 \\ \rho_0 &\longrightarrow \ \rho_t \ = \mathbb{U}(t)\,\rho_0\mathbb{U}^+(t) \end{aligned}$$

but for open/dissipative systems $\mathbb{S}\otimes\mathbb{S}_{\text{environment}}$ turns out to be non-unitary. One is led therefore to ask "What is the *most general* property-preserving transformation to which a density matrix can be subjected?"[23]

Recognizing that quantum theory is an exercise in linear algebra, one looks to $\rho \longrightarrow \mathbb{A}\,\rho\,\mathbb{B}$. Hermiticity-preservation requires $\mathbb{B}=\mathbb{A}^+$, and we have

$$\rho \longrightarrow \mathbb{A}\,\rho\,\mathbb{A}^+$$

Positivity-preservation is then automatic, since

$$(\alpha|\rho|\alpha) \geqq 0 \text{ all } |\alpha) \quad\Longrightarrow\quad (\beta|\rho|\beta) \geqq 0 \text{ all } |\beta) = \mathbb{A}^+|\alpha)$$

Trace-preservation forces $\mathbb{A}$ to be unitary

$$\text{tr}\mathbb{A}\,\rho\,\mathbb{A}^+ = \text{tr}\mathbb{A}^+\mathbb{A}\,\rho = \text{tr}\rho \ \text{(all } \rho) \quad\Longrightarrow\quad \mathbb{A}^+\mathbb{A} = \mathbb{I}$$

But hermiticity and positivity-preservation are undamaged if, more generally, we write

$$\rho \longrightarrow \textstyle\sum_i \mathbb{A}_i\,\rho\,\mathbb{A}_i^+ \tag{10.1}$$

whereupon the former unitarity condition assumes the form

$$\textstyle\sum_i \mathbb{A}_i^+\mathbb{A}_i = \mathbb{I} \tag{10.2}$$

Transformations of the form (10) were first called called "operations" by Kraus[24] and the matrices $\mathbb{A}_i$ have come to be called "Kraus matrices" (or "Kraus operators"; Kraus himself called them "effects"). We have been led by general considerations to precisely the material that was used to assemble the theory of PVOM measurements.

To illustrate the *dynamical* utility of (10) Barnett borrows from quantum dissipation theory the Kraus matrices

$$\mathbb{A}_{\text{up}} = \begin{pmatrix} e^{-\Gamma t} & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbb{A}_{\text{dn}} = \begin{pmatrix} 0 & 0 \\ \sqrt{1-e^{-2\Gamma t}} & 0 \end{pmatrix} \quad : \quad t \geq 0$$

---

[23] My principal source here has been Barnett's §4.5 and Appendix J.

[24] K. Kraus, *States, Effects and Operations* (1983). The formalism developed by Kraus was in fact first sketched by E. C. G. Sudarshan, P. M. Mathews and Jayaseetha Rau, "Stochastic dynamics of quantum-mechanical systems," Phys. Rev. **121**, 920–924 (1961).

where Γ is a positive constant. From

$$\begin{aligned} \mathbb{A}^+_{\text{up}}\mathbb{A}_{\text{up}} &= \begin{pmatrix} e^{-2\Gamma t} & 0 \\ 0 & 1 \end{pmatrix} \\ \mathbb{A}^+_{\text{dn}}\mathbb{A}_{\text{dn}} &= \begin{pmatrix} 1 - e^{-2\Gamma t} & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

we see that indeed $\mathbb{A}^+_{\text{up}}\mathbb{A}_{\text{up}} + \mathbb{A}^+_{\text{dn}}\mathbb{A}_{\text{dn}} = \mathbb{I}$, as required. Write

$$\rho(t) = \mathbb{A}_{\text{up}}\,\rho(0)\,\mathbb{A}^+_{\text{up}} + \mathbb{A}_{\text{dn}}\,\rho(0)\,\mathbb{A}^+_{\text{dn}}$$

and assume that the ensemble of qubits is initially pure "up":

$$\rho(0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Then

$$\begin{aligned} \rho(t) &= \begin{pmatrix} e^{-2\Gamma t} & 0 \\ 0 & 1 - e^{-2\Gamma t} \end{pmatrix} \\ &\downarrow \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{as } t \to \infty \end{aligned}$$

Writing $|\text{up}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|\text{dn}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ we have

$$\left.\begin{aligned} \text{“up” probability at time } t \geq 0 &= (\text{up}|\rho(t)|\text{up}) = e^{-2\Gamma t} \\ \text{“dn” probability at time } t \geq 0 &= (\text{dn}|\rho(t)|\text{dn}) = 1 - e^{-2\Gamma t} \end{aligned}\right\} \tag{11}$$

Note that the dissipative evolution described above is *not unitary*, for Hamiltonian-generated unitary motion

$$\mathbb{U}(t)\,\rho_0\mathbb{U}^+(t) \text{ is } \textit{oscillatory} \text{ in all cases}$$

Moreover, the probabilities (11) become meaningless when $t < 0$, for this deep reason: time-reversal $t \to -t$ sends

$$\mathbb{A}_{\text{up}} = \mathbb{A}^+_{\text{up}} \longrightarrow \begin{pmatrix} e^{\Gamma t} & 0 \\ 0 & 1 \end{pmatrix}$$

$$\mathbb{A}_{\text{dn}} \longrightarrow \begin{pmatrix} 0 & 0 \\ i\sqrt{e^{2\Gamma t} - 1} & 0 \end{pmatrix}, \quad \mathbb{A}^+_{\text{dn}} \longrightarrow \begin{pmatrix} 0 & -i\sqrt{e^{2\Gamma t} - 1} \\ 0 & 0 \end{pmatrix}$$

whence

$$\mathbb{A}^+_{\text{up}}\mathbb{A}_{\text{up}} + \mathbb{A}^+_{\text{dn}}\mathbb{A}_{\text{dn}} \longrightarrow \begin{pmatrix} 2e^{2\Gamma t} - 1 & 0 \\ 0 & 1 \end{pmatrix} \neq \mathbb{I}$$

So the reversed transformation cannot be trace-preserving: quantum dissipation is *irreversible* (as also are quantum measurement processes), whereas unitary quantum dynamical motion is invariably reversible.

The claim is that transformations $\rho \longrightarrow \rho'$ are quantum-mechanically conceivable if and only if they can be achieved by operations. To illustrate the force of that principle, Barnett looks to what he calls "Chefles' state separation process."[25] Let $|\alpha)$ and $|\beta)$ be unknown non-orthogonal states: $|(\alpha|\beta)| \neq 0$. A "separation process" $\mathcal{P}$ is a state-transformation

$$\mathcal{P}: \{|\alpha) \to |\alpha'),\ |\beta) \to |\beta')\} \quad \text{such that} \quad |(\alpha'|\beta')| < |(\alpha|\beta)|$$

Such processes are distinct from (but, as will emerge, related to) the "optimal discrimination processes" considered previously. If it were possible by repeated $\mathcal{P}$-processes to achieve $|(\alpha'|\beta')| = 0$ then Bob would be in position ultimately to discriminate $|\alpha)$ from $|\beta)$ by simple PVM measurement. There must exist a least-possible value $|(\alpha'|\beta')|_{\text{min}} > 0$ of $|(\alpha'|\beta')|$, which it has become our business to calculate. To that end...

[25] Anthony Chefles studies the optical applications of quantum information theory, and during the 1990s co-authored many papers with Barnett (of whom I suspect he was a student). They were, in fact, co-authors of the first papers dealing with the process here in question. A detailed survey of the subject can be found in §5.3 of the paper cited in Note [22].