Simultaneous Measurement

of

Noncommuting Quantum Observables

Nicholas Wheeler
October 2012

Introduction. In 1965, E. Arthurs and J. L. Kelly, Jr.—who were employed as engineers at Bell Lab—published in BSTJ Briefs a short paper[1] bearing a title "On the simultaneous measurement of conjugate observables" which most quantum physicists could be expected to find perplexing. Within a few months C. Y. She and H. Heffner devised an alternative approach[2] to the theory of simultaneous measurement that reproduced the results first obtained by Arthurs and Kelly, and by about 1980 the subject—stimulated mainly by the practical needs of quantum opticians and the development of quantum information theory—had begun to generate wide interest.[3]

That early work took John von Neumann's idealized theory of quantum measurement as its point of departure, but more recently the theory of generalized (non-ideal) quantum measurement has been brought into play. It is from that point of view that S. M. Barnett approaches the subject,[4] and it is Barnett's brief survey (intended to illustrate the utility of the positive operator-valued measure (POVM) concept) that has motivated the following discussion.

---

[1] Bell Systems Technical Journal **44**, 725–729 (1965). This was a journal seldom consulted by most physicists (it ceased publication in 1983), though it was the journal in which the results of the Davisson-Germer electron diffraction experiment were first reported (1928), the journal in which Claude Shannon published his "A mathematical theory of communication" (1948), the journal in which W. Boyle and G. E. Smith announced their invention of the charge-coupled device (1970) and in which many other important developments were first reported.

[2] "Simultaneous measurement of noncommuting observables," Phys. Rev. **152**, 1103–1110 (1966).

[3] For major references see the bibiography in Ingrid Olson, "Simultaneous measurement of conjugate observables" (Reed College Thesis, 2006).

[4] *Quantum Information* (2009), pages 97–98.

Schrödinger's inequality. From the detailed history presented in Chapter 7 (§7.1. "The Uncertainty Relations") of Jammer's *The Conceptual Development of Quantum Mechanics* (1966) we learn that development of the Heisenberg uncertainty principle was not at all the straightforward exercise that textbooks commonly respresent it to have been, and that many physicists (Bohr, Pauli—the usual suspects—plus also Weyl, Kennard, Ruark, Condon, Robertson) contributed to the final sharp formulation of Heisenberg's initial insight. Dirac and (independently) Jordan had observed already in 1926 that—in view of the central place which they assigned to the commultation relation $\mathsf{x}\,\mathsf{p}-\mathsf{p}\,\mathsf{x} = i\hbar\mathsf{I}$—it became impossible to assign sharp values simultaneously to the position and momentum of a quantum particle. Heisenberg undertook to quantify the statistical relationship between x-values and p-values, and arrived in 1927 at the statement $\Delta x\Delta p = \hbar$, which—thus introducing a persistent element of confusion into this story—he attempted to account for physically as an inevitable "observer effect" (think of "Heisenberg's microscope").

In the spring of 1930[5] Schödinger was studying teh problem of how to distribute, in an optimal simultaneous measurement of p and x at time t_0, the unavoidable uncertainty $\frac{1}{2}\hbar$ between two variables in such a way that at a given later instant t the uncertainty Δx in position will be minimal. Sommerfeld drew his attention to recent papers by Condon and Robertson (1929) which Schrödinger instantly saw could be improved upon:[6]

Many copies of $|\psi)$ are presented to an A-meter, respresented by the self-adjoint operator A. The expected mean of the A-meter readings is

$$\langle \mathsf{A}\rangle = (\psi|\mathsf{A}|\psi)$$

Presentation of many copies of $|\psi)$ to a B-meter supplies

$$\langle \mathsf{B}\rangle = (\psi|\mathsf{B}|\psi)$$

Use that $|\psi)$-dependent data to construct "centered" operators

$$\mathsf{a} = \mathsf{A} - \langle \mathsf{A}\rangle\mathsf{I} \quad\text{and}\quad \mathsf{b} = \mathsf{B} - \langle \mathsf{B}\rangle\mathsf{I}$$

The "centered 2^{nd} moments" (or "variance" σ^2 = "squared standard deviation" ="squared uncertainty") of the A/B data can then be described

$$(\Delta A)^2 = (\psi|\mathsf{a}^2|\psi) \quad\text{and}\quad (\Delta B)^2 = (\psi|\mathsf{b}^2|\psi)$$

[5] I quote here from Jammer's page 336.

[6] E. Schrödinger, "Zur Heisenbergschen Unshärfeprinzip," Berliner Berichte, 296–303 (1930). I have taken my argument from my *Quantum Mechanics* notes (1967/68), Chapter 3, pages 55–56. For a somewhat truncated version of the same argument see §3.5.1 in D. J. Griffiths, *Introduction to Quantum Mechanics* (2^{nd} edition, 2005). Quite good also—from many points of view—is the Wikipedia article "Uncertainty principle," which reproduces the same line of argument.

Define

$$|a) = \mathsf{a}|\psi) \quad \text{and} \quad |b) = \mathsf{b}|\psi)$$

Then by the Cauchy-Schwarz

$$\begin{aligned}(\Delta A)^2(\Delta B)^2 &= (a|a)(b|b)\\ &\geq (a|b)(b|a) = |(a|b)|^2 \quad \text{with equality iff } |a) \sim |b)\\ &= |(\psi|\,\mathsf{a}\,\mathsf{b}\,|\psi)|^2\end{aligned}$$

Write

$$\mathsf{a}\,\mathsf{b} = \frac{\mathsf{a}\,\mathsf{b} + \mathsf{b}\,\mathsf{a}}{2} + i\,\frac{\mathsf{a}\,\mathsf{b} - \mathsf{b}\,\mathsf{a}}{2i}$$

and notice that the self-aqjointness of a and b implies that of both $\frac{1}{2}[\mathsf{a},\mathsf{b}]_+$ and $\frac{1}{2i}[\mathsf{a},\mathsf{b}]_-$. We therefore have

$$\begin{aligned}(\Delta A)^2(\Delta B)^2 &\geq \left|\left\langle\frac{\mathsf{a}\,\mathsf{b} + \mathsf{b}\,\mathsf{a}}{2}\right\rangle + i\left\langle\frac{\mathsf{a}\,\mathsf{b} - \mathsf{b}\,\mathsf{a}}{2i}\right\rangle\right|^2\\ &= \left\langle\frac{\mathsf{a}\,\mathsf{b} + \mathsf{b}\,\mathsf{a}}{2}\right\rangle^2 + \left\langle\frac{\mathsf{a}\,\mathsf{b} - \mathsf{b}\,\mathsf{a}}{2i}\right\rangle^2\end{aligned}$$

By quick calculation

$$\mathsf{a}\,\mathsf{b} \pm \mathsf{b}\,\mathsf{a} = \begin{cases}\mathsf{A}\,\mathsf{B} + \mathsf{B}\,\mathsf{A} - 2\mathsf{A}\langle\mathsf{B}\rangle - 2\mathsf{B}\langle\mathsf{A}\rangle + 2\langle\mathsf{A}\rangle\langle\mathsf{B}\rangle\\ \mathsf{A}\,\mathsf{B} - \mathsf{B}\,\mathsf{A}\end{cases}$$

so

$$\langle\mathsf{a}\,\mathsf{b} \pm \mathsf{b}\,\mathsf{a}\rangle = \begin{cases}\langle\mathsf{A}\,\mathsf{B} + \mathsf{B}\,\mathsf{A}\rangle - 2\langle\mathsf{A}\rangle\langle\mathsf{B}\rangle\\ \langle\mathsf{A}\,\mathsf{B} - \mathsf{B}\,\mathsf{A}\rangle\end{cases}$$

which gives Schrödinger's inequality

$$(\Delta A)^2(\Delta B)^2 \geq \left\langle\frac{\mathsf{A}\,\mathsf{B} - \mathsf{B}\,\mathsf{A}}{2i}\right\rangle^2 + \left[\left\langle\frac{\mathsf{A}\,\mathsf{B} + \mathsf{B}\,\mathsf{A}}{2}\right\rangle - \langle\mathsf{A}\rangle\langle\mathsf{B}\rangle\right]^2 \tag{1.1}$$

$$\geq \text{greater of } \left\{\left\langle\frac{\mathsf{A}\,\mathsf{B} - \mathsf{B}\,\mathsf{A}}{2i}\right\rangle^2,\ \left[\left\langle\frac{\mathsf{A}\,\mathsf{B} + \mathsf{B}\,\mathsf{A}}{2}\right\rangle - \langle\mathsf{A}\rangle\langle\mathsf{B}\rangle\right]^2\right\} \tag{1.2}$$

In the most familiar instance we therefore have

$$\begin{aligned}(\Delta x)^2(\Delta p)^2 &\geq \left\langle\frac{\mathsf{x}\,\mathsf{p} - \mathsf{p}\,\mathsf{x}}{2i}\right\rangle^2 + \left[\left\langle\frac{\mathsf{x}\,\mathsf{p} + \mathsf{p}\,\mathsf{x}}{2}\right\rangle - \langle\mathsf{x}\rangle\langle\mathsf{p}\rangle\right]^2\\ &\geq \left\langle\frac{\mathsf{x}\,\mathsf{p} - \mathsf{p}\,\mathsf{x}}{2i}\right\rangle^2 = \left\langle\frac{i\hbar\mathsf{I}}{2i}\right\rangle^2 = (\hbar/2)^2\\ &\Downarrow\\ \Delta x\,\Delta p &\geq \tfrac{1}{2}\hbar\end{aligned}$$

In classical statistics, if x and y are random variables then one has (for all m and n)

$$\langle x^m y^n\rangle = \langle x^m\rangle\langle y^n\rangle \quad \text{iff } x \text{ ane } y \text{ are statistically independent}$$

The number $\langle xy\rangle - \langle x\rangle\langle y\rangle$ provides therefore a leading indicator of the extent to which x and y are statistically dependent or correlated. On the right side of (1) we encounter just such a construction

$$C_{\mathsf{AB}}[|\psi)] \equiv \Big\langle \frac{\mathsf{a\,b} + \mathsf{b\,a}}{2} \Big\rangle = \Big\langle \frac{\mathsf{A\,B} + \mathsf{B\,A}}{2} \Big\rangle - \langle \mathsf{A} \rangle\langle \mathsf{B} \rangle \tag{2}$$

which it becomes natural in this light to call the "quantum correlation coefficient."[7]

If A and B commute then (1.2) supplies

$$(\Delta A)^2(\Delta B)^2 \geq \big[\langle \mathsf{A\,B} \rangle - \langle \mathsf{A} \rangle\langle \mathsf{B} \rangle\big]^2$$

The eigenvectors (but not the eigenvalues) of A and B are in this case shared. If $|\psi)$ is such a shared eigenvector ($\mathsf{A}|\psi) = \alpha|\psi)$ and $\mathsf{B}|\psi) = \beta|\psi)$) then

$$\begin{aligned}(\Delta A)^2(\Delta B)^2 &\geq [\alpha\beta - \alpha\beta]^2 = 0\\ &= 0 \quad \text{because} \quad \Delta A = \Delta B = 0\end{aligned}$$

But linear combinations of such (orthogonal) eigenvectors give $\Delta A\,\Delta B > 0$. Suppose, for example, that $|\psi) = \cos\theta\,|\psi_1) + \sin\theta\,|\psi_2)$. Then

$$\begin{aligned}C_{\mathsf{AB}}[|\psi)] &= (\alpha_1\beta_1\cos^2\theta + \alpha_2\beta_2\sin^2\theta)\\ &\qquad\qquad - (\alpha_1\cos^2\theta + \alpha_2\sin^2\theta)(\beta_1\cos^2\theta + \beta_2\sin^2\theta)\\ &= (\alpha_1 - \alpha_2)(\beta_1 - \beta_2)\cos^2\theta\sin^2\theta\end{aligned}$$

which give back the preceding result as a degenerate special case (set $\theta = n\pi/2$ with $n = 0, \pm1, \pm2, \ldots$).

More interesting are results that follow from the assumption that A and B are *conjugate*:

$$\mathsf{A\,B} - \mathsf{B\,A} = i\,\mathsf{I}$$

Introduce operators

$$\mathsf{W} = \tfrac{1}{\sqrt{2}}(\mathsf{A} + i\,\mathsf{B}) \quad \text{and} \quad \mathsf{W}^+ = \tfrac{1}{\sqrt{2}}(\mathsf{A} - i\,\mathsf{B})$$

which—since not self-adjoint—do not represent observables, but are the key to all that follows. From

$$\begin{aligned}\mathsf{W\,W}^+ &= \tfrac{1}{2}(\mathsf{A\,A} - i\mathsf{A\,B} + i\mathsf{B\,A} + \mathsf{B\,B}) = \tfrac{1}{2}(\mathsf{A\,A} + \mathsf{B\,B} + \mathsf{I})\\ \mathsf{W}^+\mathsf{W} &= \tfrac{1}{2}(\mathsf{A\,A} + i\mathsf{A\,B} - i\mathsf{B\,A} + \mathsf{B\,B}) = \tfrac{1}{2}(\mathsf{A\,A} + \mathsf{B\,B} - \mathsf{I})\end{aligned}$$

obtain

$$\mathsf{W\,W}^+ - \mathsf{W}^+\mathsf{W} = \mathsf{I} \quad \Longrightarrow \quad \begin{cases}\mathsf{W\,W}^+ = \mathsf{W}^+\mathsf{W} + \mathsf{I}\\ \mathsf{W}^+\mathsf{W} = \mathsf{W\,W}^+ - \mathsf{I}\end{cases}$$

[7] See page 202 in D. Bohm, *Quantum Mechanics* (1951). Generally $\mathsf{A\,B} \neq \mathsf{B\,A}$. In (2) we are told to "split the difference."

The operators $\mathsf{W}\,\mathsf{W}^+$ and $\mathsf{W}^+\mathsf{W}$ are manifestly self-adjoint (eigenvalues therefore real, and eigenvectors orthogonal) and positive semi-definite $(\psi|\mathsf{W}\,\mathsf{W}^+|\psi)\geq 0$ and $(\psi|\mathsf{W}^+\mathsf{W}|\psi)\geq 0$ for all $|\psi)$).[8] Suppose it to be the case that

$$\mathsf{W}^+\mathsf{W}|\alpha)=\lambda\,|\alpha)\quad\iff\quad\mathsf{W}\,\mathsf{W}^+|\alpha)=(\lambda+1)\,|\alpha)$$

Multiplication by W^+ supplies $\mathsf{W}^+\mathsf{W}\,\mathsf{W}^+|\alpha)=(\mathsf{W}\,\mathsf{W}^+-\mathsf{I})|\alpha)=\lambda\mathsf{W}^+|\alpha)$ whence $\mathsf{W}^+\mathsf{W}\cdot\mathsf{W}^+|\alpha)=(\lambda+1)\cdot\mathsf{W}^+|\alpha)$ and similarly $\mathsf{W}^+\mathsf{W}\cdot\mathsf{W}|\alpha)=(\lambda-1)\cdot\mathsf{W}|\alpha)$. So ascending powers of W^+ produce eigenvectors $|\alpha_{+n})=(\mathsf{W}^+)^n|\alpha)$ with eigenvalues $\{\lambda,\lambda+1,\lambda+2,\lambda+3,\ldots\}$ while ascending powers of W produce eigenvectors $|\alpha_{-n})=(\mathsf{W})^n|\alpha)$ with eigenvalues $\{\lambda,\lambda-1,\lambda-2,\lambda-3,\ldots\}$. The latter sequence must, however, truncate to avoid violation of the positivity condition: there must exist a (normalized) state $|0)$ with the property that

$$\mathsf{W}\,|0)=0$$

Building on that foundation, we construct

$$\begin{aligned}|1)&=c_0\,\mathsf{W}^+|0)\\|2)&=c_1\,\mathsf{W}^+|1)\\|3)&=c_2\,\mathsf{W}^+|2)\\&\;\vdots\\|n+1)&=c_n\,\mathsf{W}^+|n)\\&\;\vdots\end{aligned}$$

To evaluate the constants c_n (which can without loss of generality be assumed to be real) we proceed

$$\left.\begin{aligned}(n|\mathsf{W}\,\mathsf{W}^+|n)&=(n+1)(n|n)=n+1\\&=c_n^{-2}(n+1|n+1)=c_n^{-2}\end{aligned}\right\}\Longrightarrow c_n=\tfrac{1}{\sqrt{n+1}}$$

It now follows that

$$\begin{aligned}|n)&=\tfrac{1}{\sqrt{n}}\mathsf{W}^+|n-1)\\&=\tfrac{1}{\sqrt{n(n-1)}}(\mathsf{W}^+)^2|n-2)\\&\;\vdots\\&=\tfrac{1}{\sqrt{n!}}(\mathsf{W}^+)^n|0)\end{aligned}$$

For the purposes at hand these results are most conveniently written

$$\mathsf{W}|n)=g_n|n-1),\quad\mathsf{W}^+|n)=g_{n+1}|n+1)\quad\text{with}\quad g_n=\sqrt{n}$$

from which we recover $\mathsf{W}^+\mathsf{W}|n)=g_n\,\mathsf{W}^+|n-1)=g_ng_n|n)=n|n)$.

---

[8] The standard notation $\mathsf{a}\,\mathsf{a}^+$ is not available because the symbol a has been preempted. The W-notation is intended to draw attention to the circumstance that $\mathsf{W}\,\mathsf{W}^+$ and $\mathsf{W}^+\mathsf{W}$ possess "Wishart structure."

Variants of the preceding algebra are encountered in many quantum mechanical contexts, all of which derive from Dirac's approach to the harmonic oscillator problem.[9] It is central to quantum optics (quantized oscillatory modes of the radiation field),[10] And it provides the formal model upon which Witten's "supersymmetric quantum mechanics" is based.[11] But the immediate point of the exercise emeerges when we look back again to Schrödinger's inequality (1). If the state presented repeatedly to the A-meter on Monday—and to the B-meter on Tuesday—is $|n)$, and if A and B are conjugate ($[\mathsf{A}, \mathsf{B}] = i\,\mathsf{I}$) then

$$(\Delta A)^2(\Delta B)^2 \geq \tfrac{1}{4} + \big\{C_{\mathsf{AB}}[|n)]\big\}^2$$

where

$$C_{\mathsf{AB}}[|n)] = \tfrac{1}{2}(n|\mathsf{A}\,\mathsf{B} + \mathsf{B}\,\mathsf{A}|n) - (n|\mathsf{A}|n)(n|\mathsf{B}|n)$$

But from

$$\mathsf{A} = \tfrac{1}{\sqrt{2}}(\mathsf{W}^+ + \mathsf{W}) \quad \text{and} \quad \mathsf{B} = i\tfrac{1}{\sqrt{2}}(\mathsf{W}^+ - \mathsf{W})$$

we obtain

$$\mathsf{A}\,\mathsf{B} + \mathsf{B}\,\mathsf{A} = i(\mathsf{W}^+\mathsf{W}^+ - \mathsf{W}\,\mathsf{W})$$

so[12]

$$\begin{aligned}
C_{\mathsf{AB}}[|n)] &= i\tfrac{1}{2}\Big\{(n|\mathsf{W}^+\mathsf{W}^+ - \mathsf{W}\,\mathsf{W}|n) - (n|\mathsf{W}^+ + \mathsf{W}|n)(n|\mathsf{W}^+ - \mathsf{W}|n)\Big\}\\
&= i\tfrac{1}{2}\Big\{g_{n+1}g_{n+2}(n|n+2) - g_{n-1}g_n(n|n-2)\\
&\qquad - \big[g_{n+1}(n|n+1) + g_n(n|n-1)\big]\big[g_{n+1}(n|n+1) - g_n(n|n-1)\big]\Big\}\\
&= 0 \quad \text{by} \quad (n|m) = \delta_{nm}
\end{aligned} \tag{3}$$

For such states we therefore have

$$(n|\mathsf{A}|n) = (n|\mathsf{B}|n) = 0 \quad \text{and} \quad \Delta A\,\Delta B \geq \tfrac{1}{2}$$

The inequality can, however, be sharpened; from

$$\begin{aligned}
(n|\mathsf{A}^2|n) = (n|\mathsf{B}^2|n) &= \tfrac{1}{2}(n|\mathsf{W}^+\mathsf{W} + \mathsf{W}\,\mathsf{W}^+|n) + \text{two terms that vanish}\\
&= \tfrac{1}{2}\big[g_ng_n + g_{n+1}g_{n+1}\big](n|n)\\
&= \tfrac{1}{2}(2n+1)
\end{aligned}$$

[9] §34, *Principles of Quantum Mechanics* (3^{rd} edition, 1947). For an account of some elegant elaborations of the method due to Schwinger see Chapter 0, pages 40–42 in my *Advanced Quantum Topics* (2000).

[10] See, for example, §3.1.1 in Yoshihisa Yamamoto & Ataç İmamoḡlu, *Mesoscopic Quantum Optics* (1999).

[11] See Christopher Lee, "Supersymmetric quantum mechanics" (Reed College Thesis, 1999), which provides an elaborate bibliography.

[12] If $n = 0$ or 1 some of the terms in the following expression—namely g_{-1}, $|-1)$ and $|-2)$—are undefined, but those formal artifacts all vanish, essentially because $\mathsf{W}^p|0) = 0 : p = 1, 2, \ldots$.

we obtain

$$\Delta A\,\Delta B = n + \tfrac{1}{2} \geq \tfrac{1}{2}$$

It is no accident that those numbers are proportional to the energy eigenvalues $E_n = \hbar\omega(n + \tfrac{1}{2})$ of a quantum oscillator.

The states $|n)$ acquire importance partly (as in oscillator theory) from the circumstance that they are eigenstates of $\mathsf{W}^+\mathsf{W}$ and $\mathsf{W}\,\mathsf{W}^+$, but more generally from (3); they are *minimal uncertainty* states that in quantum mechanics engender wavepackets of "minimal dispersion"(see Griffiths[6], §3.5.2) and in quantum optics[13] are called "coherent states."

The simplest possible non-commutation relation $[\mathsf{A}, \mathsf{B}] = i\,\mathsf{I}$ (from which the preceding discussion proceeded) does not admit of finite-dimensional realization (compare the traces of the left and right sides of $[\mathbb{A}, \mathbb{B}] = i\,\mathbb{I}$). But finite-dimensional quantum mechanics presents many contexts in which Schrödinger's inequality proves valuable. Most commonly those arise when one has in hand either a trace-wise orthonormal basis $\{\mathbb{E}_1, \mathbb{E}_2, \dots, \mathbb{E}_{N^2}\}$ in the space of $N \times N$ hermitian matrices

$$\mathbb{A} = \sum_{j=1}^{N^2} a_j \mathbb{E}_j \quad \text{with} \quad a_k = \tfrac{1}{N}\text{tr}\,\mathbb{A}\mathbb{E}_k \quad \text{by} \quad \tfrac{1}{N}\text{tr}\,\mathbb{E}_j\mathbb{E}_k = \delta_{jk}$$

or a set $\{\mathbb{F}_1, \mathbb{F}_2, \dots, \mathbb{F}_n\}$ of hermitian matrices that is closed under commutation (in short, a Lie algebra):

$$[\mathbb{F}_i, \mathbb{F}_j] = \sum_k c_i{}^k{}_j \mathbb{F}_k$$

Look, for example, to the Pauli matrices

$$\boldsymbol{\sigma}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \boldsymbol{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \boldsymbol{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \boldsymbol{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

which possess *both* of the aforementioned properties: they are trace-wise orthonormal

$$\tfrac{1}{2}\text{tr}\,\boldsymbol{\sigma}_m\boldsymbol{\sigma}_n = \delta_{mn}$$

and since multiplicatively closed

$$\begin{aligned} \boldsymbol{\sigma}_0\boldsymbol{\sigma}_n &= \boldsymbol{\sigma}_n &&: \quad n = 0, 1, 2, 3 \\ \boldsymbol{\sigma}_j\boldsymbol{\sigma}_k &= \delta_{jk}\boldsymbol{\sigma}_0 + i\epsilon_{jkl}\boldsymbol{\sigma}_l &&: \quad \{j, k, l\} = 1, 2, 3 \end{aligned}$$

are closed also under commutation: $[\boldsymbol{\sigma}_j, \boldsymbol{\sigma}_k] = 2i\epsilon_{jkl}\boldsymbol{\sigma}_l$. Suppose

$$\mathbb{A} = a_1\boldsymbol{\sigma}_1 + a_2\boldsymbol{\sigma}_2 + a_3\boldsymbol{\sigma}_3 = \boldsymbol{a}\boldsymbol{\cdot}\boldsymbol{\sigma} \quad \text{and} \quad \mathbb{B} = \boldsymbol{b}\boldsymbol{\cdot}\boldsymbol{\sigma}$$

---

[13] See C. C. Gerry & P. L. Knight, *Introductory Quantum Optics* (2005), Chapter 3.

Then $\mathbb{AB} = (\boldsymbol{a}\cdot\boldsymbol{b})\,\boldsymbol{\sigma}_0 + i(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{\sigma}$ supplies

$$\mathbb{AA} = (\boldsymbol{a}\cdot\boldsymbol{a})\,\boldsymbol{\sigma}_0 \quad\text{and}\quad \mathbb{BB} = (\boldsymbol{b}\cdot\boldsymbol{b})\,\boldsymbol{\sigma}_0$$

$$\begin{aligned}\mathbb{AB}+\mathbb{BA} &= 2(\boldsymbol{a}\cdot\boldsymbol{b})\,\boldsymbol{\sigma}_0\\ \mathbb{AB}-\mathbb{BA} &= 2i(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{\sigma}\end{aligned}$$

and we have $\langle\mathbb{A}\rangle = \sum a_k\langle\boldsymbol{\sigma}_k\rangle = \langle\boldsymbol{a}\cdot\boldsymbol{\sigma}\rangle$ and $\langle\mathbb{B}\rangle = \langle\boldsymbol{b}\cdot\boldsymbol{\sigma}\rangle$ whence

$$(\Delta A)^2(\Delta B)^2 = \big[(\boldsymbol{a}\cdot\boldsymbol{a}) - \langle\boldsymbol{a}\cdot\boldsymbol{\sigma}\rangle^2\big]\big[(\boldsymbol{b}\cdot\boldsymbol{b}) - \langle\boldsymbol{b}\cdot\boldsymbol{\sigma}\rangle^2\big] \tag{4.1}$$

while the Schrödinger inequality supplies a statement with quite a different appearance:

$$(\Delta A)^2(\Delta B)^2 \geq \big\langle(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{\sigma}\big\rangle^2 + \big[(\boldsymbol{a}\cdot\boldsymbol{b}) - \langle\boldsymbol{a}\cdot\boldsymbol{\sigma}\rangle\langle\boldsymbol{b}\cdot\boldsymbol{\sigma}\rangle\big]^2 \tag{4.2}$$

But when (with *Mathematica*'s assistance) I used randomly selected real 3-vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ to construct hermitian matrices $\mathbb{A}$ and $\mathbb{B}$ I was surprised to find that for every the normalized complex 2-vector $|\psi)$ the expressions on the right sides of (4.1) and (4.2) are *identical*; we have stumbled upon a curious identity

$$\begin{aligned}(\boldsymbol{a}\cdot\boldsymbol{a})(\boldsymbol{b}\cdot\boldsymbol{b}) - (\boldsymbol{a}\cdot\boldsymbol{b})^2 = (\boldsymbol{b}\cdot\boldsymbol{b})\langle\boldsymbol{a}\cdot\boldsymbol{\sigma}\rangle^2 + (\boldsymbol{a}\cdot\boldsymbol{a})\langle\boldsymbol{b}\cdot\boldsymbol{\sigma}\rangle^2 + \big\langle(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{\sigma}\big\rangle^2\\ - 2(\boldsymbol{a}\cdot\boldsymbol{b})\langle\boldsymbol{a}\cdot\boldsymbol{\sigma}\rangle\langle\boldsymbol{b}\cdot\boldsymbol{\sigma}\rangle\end{aligned} \tag{5.1}$$

In the case $\mathbb{A} = \boldsymbol{\sigma}_1$, $\mathbb{B} = \boldsymbol{\sigma}_2$ the preceding identity assumes the (strange but) suggestively simple form

$$1 = (\psi|\boldsymbol{\sigma}_1|\psi)^2 + (\psi|\boldsymbol{\sigma}_2|\psi)^2 + (\psi|\boldsymbol{\sigma}_3|\psi)^2 \quad : \quad \text{all } |\psi) \tag{5.2}$$

I am satisfied on the basis of exhaustive numerical evidence that the identities (5) are both correct, which means that the $\geq$ in (4.2) should always be read as equality…for, as it happens, a very simple reason. The $\geq$ in question was inherited from Cauchy-Schwarz, and reduces to $=$ if and only if

$$|\alpha) \equiv \mathbb{A}|\psi) - (\psi|\mathbb{A}|\psi)\cdot|\psi) \;\sim\; |\beta) \equiv \mathbb{B}|\psi) - (\psi|\mathbb{B}|\psi)\cdot|\psi)$$

From $(\psi|\alpha) = (\psi|\beta) = 0$ we learn that both of those vectors are orthogonal to $|\psi)$, which in 2-space means that they are proportional: $|\alpha) \sim |\beta)$. We are brought thus to the striking conclusion that when a randomly selected qubit $|\psi)$ is presented repeatedly first to an arbitrarily designed $\mathbb{A}$-meter and then—in a separate run—to an arbitrarily designed $\mathbb{B}$-meter, analysis of the data thus generated invariably shows the product $\Delta A\Delta B$ to be minimal.[14] In higher-dimensional contexts automatic minimality is lost, for a reason now evident.

[14] When we wrote $\mathbb{A} = \boldsymbol{a}\cdot\boldsymbol{\sigma}$ and $\mathbb{B} = \boldsymbol{b}\cdot\boldsymbol{\sigma}$ we tacitly assumed the matrices $\mathbb{A}$ and $\mathbb{B}$ to be traceless, but it is now clear that invariable minimality persists even in the absence of tracelessness.

Suppose the states presented to our meters are not identical, but are drawn from the mixed ensemble described by the density operator $\boldsymbol{\rho}$. Review of its derivation shows that Schrödinger's inequality (1) remains in force, provided

$$\langle \boldsymbol{X}\rangle = (\psi|\boldsymbol{X}|\psi) \quad \text{is reinterpreted to mean} \quad \mathrm{tr}(\boldsymbol{\rho}\boldsymbol{X})$$

Note in this regard that the effect state superposition $|\psi) \longrightarrow c_1|\psi_1) + c_2|\psi_2)$ —which sends pure states to pure states—is non-linear

$$(\psi|\boldsymbol{X}|\psi) \to \bar{c}_1 c_1(\psi_1|\boldsymbol{X}|\psi_1) + \bar{c}_1 c_2(\psi_1|\boldsymbol{X}|\psi_2) + \bar{c}_2 c_1(\psi_2|\boldsymbol{X}|\psi_1) + \bar{c}_2 c_2(\psi_2|\boldsymbol{X}|\psi_2)$$

while the effect of mixing $\boldsymbol{\rho} \longrightarrow p_1\boldsymbol{\rho}_1 + p_2\boldsymbol{\rho}_2$ is linear

$$\mathrm{tr}(\boldsymbol{\rho}\boldsymbol{X}) \longrightarrow p_1\mathrm{tr}(\boldsymbol{\rho}_1\boldsymbol{X}) + p_2\mathrm{tr}(\boldsymbol{\rho}_2\boldsymbol{X})$$

Non-linear effects do, however, enter into the description of $C_{\mathsf{AB}}[p_1\boldsymbol{\rho}_1 + p_2\boldsymbol{\rho}_2]$ *via* the $\langle \mathsf{A}\rangle\langle \mathsf{B}\rangle$ term; "mixtures of minimal states"[15] are not minimal.

What does Schrödinger's inequality signify? Present copies of $|\psi)$ (else states drawn from the mixed ensemble $\boldsymbol{\rho}$) many times to an A-meter and from the meter readings $\{a_1, a_2, \ldots, a_{\text{many}}\}$ compute the emperical mean $\bar{a}$ and the centered moments

$$\overline{(a-\bar{a})^p} \quad : \quad p = 2, 3, 4, \ldots$$

of which $\langle \mathsf{A}\rangle$ and

$$\big\langle (\mathsf{A} - \langle \mathsf{A}\rangle)^p\big\rangle = (\psi|\big(\mathsf{A} - (\psi|\mathsf{A}|\psi)\big)^p|\psi) \quad \text{else} \quad \mathrm{tr}\big(\boldsymbol{\rho}(\mathsf{A} - \mathrm{tr}\,\boldsymbol{\rho}\mathsf{A})^p\big)$$

by the Born Rule provide theoretical estimates. Do the same—in a separate experimental run—with a B-meter. The Schrödinger inequality (1) describes an inevitable relationship among the lowest-order moments $\{\langle \mathsf{A}^p\rangle, \langle \mathsf{B}^p\rangle\} : p = 1, 2,$ the statement of which requires however that one have access also to data

$$\langle \mathsf{C}\rangle,\ \langle \mathsf{D}\rangle \quad \text{with} \quad \mathsf{C} = \tfrac{1}{2i}(\mathsf{A}\,\mathsf{B} - \mathsf{B}\,\mathsf{A}),\ \mathsf{D} = \tfrac{1}{2}(\mathsf{A}\,\mathsf{B} + \mathsf{B}\,\mathsf{A})$$

acquired from two additional experiments.

There is, of course, no end to the list $\{\mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D}, \mathsf{E}, \mathsf{F}, \mathsf{G}, \ldots\}$ of observables of which one could construct moments of all orders, and indeed; quantum mechanics can in its entirety be portrayed as a "theory of interactive moments."[16] Schrödinger's "binary preoccupation"—his interest in a universal relationship among the lowest-order moments of a *pair* of observables—would in this light seem arbitrarily restrictive but for the clarity of its conceptual roots.

---

[15] I place the phrase between quotation marks because actually it does not make unambiguous sense to speak of the states from which a quantum mixtures has been assembled.

[16] See my *Advanced Quantum Topics*(2000), Chapter 2, pages 51–60.

The Hamiltonian formulation of *classical* mechanics is erected upon the notion that dynamical variables occur in *conjugate pairs* $\{q, p\}$—a notion which leads naturally (*via* the Poisson bracket) to the more general concept of conjugate observables. Heisenberg's early efforts led Born (1926) to the realization that in quantum theory the role of the classical variables $\{q, p\}$ is taken over by objects $\{\mathbf{q}, \mathbf{p}\}$ that fail to commute, and that the statement $\mathbf{q}\,\mathbf{p} - \mathbf{p}\,\mathbf{q} = i\hbar\mathbf{I}$ must lie at the foundation of any mature quantum theory, from which Dirac (and independently Jordan) promptly drew the qualitative conclusion (1926) that "one cannot answer any question on the quantum theory which refers [simultaneously] to numerical values of both q and p." Heisenberg undertook (1927) to quantify that assertion, and by a Fourier-analytic argument[17] was led to a statement $\Delta q\Delta p = \hbar$ for which he then considered himself obliged to provide a physical explanation. This led Heisenberg and others (Ruark, Kennard) to inquire closely into the physics of measurement (and to attempts to design experiments that would achieve $\Delta q\Delta p < \hbar$)—an effort from which we inherit the "Heisenberg microscope." Meanwhile, Condon and Robertson were looking more closely to the purely mathematical ramifications[18] of $[\mathbf{q}, \mathbf{p}] = i\hbar\mathbf{I}$ and, more generally, of $[\mathbf{A}, \mathbf{B}] = i\hbar\mathbf{I}$. Robertson—who by 1929 had $\psi(x)$ and the rest of the Schrödinger formalism at his disposal—obtained

$$
\begin{aligned}
\Big[\int \psi^*(\mathbf{A}-\mathbf{A}_0)^2\psi\,dx\Big]^{\frac{1}{2}}\Big[\int \psi^*(\mathbf{B}-\mathbf{B}_0)^2\psi\,dx\Big]^{\frac{1}{2}} &\geq \tfrac{1}{2i}\int \psi^*[\mathbf{A},\mathbf{B}]\psi\,dx\\
&\Downarrow\\
\Delta x\Delta p &\geq \tfrac{1}{2}\hbar
\end{aligned}
$$

and it was from Robertson's argument that Schrödinger took the clues that led to (1).

While arguments involving devices like Heisenberg's microscope do allude —if in a contingent, phenomenological way—to the *simultaneous* measurement of x and p, Schrödinger's does not, except in this sense: it alludes to properties "*simultaneously latent*" in $|\psi)$, and placed him in position to *describe* the states —solutions of $C_{\mathbf{xp}}[|\psi)] = 0$—which, when subjected to the multi-measurement protocol described previously, can be expected to yield results Δx and Δp for which $\Delta x\Delta p = \frac{1}{2}\hbar$ is realized. The individual measurements contemplated in that protocol are idealized von Neumann (projective) measurements, each of which prepares one or another of the eigenstates of $\mathbf{x}$ else $\mathbf{p}$ (more generally $\mathbf{A}$ else $\mathbf{B}$), but none of which prepares $|\psi_{\text{min}})$.

Schrödinger did not contemplate a simultaneous measurement of x and p, so had nothing to about either how such a measurement might be undertaken or what might in principle be its optimal result. The first to do so were Arthurs and Kelly.[1] Their paper—partly because of its terse obscurtity—inspired other

[17] See Jammer, page 327.

[18] It was unclear at the time whether Heisenberg had touched upon a fundamental feature of quantum physics or merely an artifact of the quantum formalism that was struggling to take shape.

authors to devise alternative approaches[19] to solution of the simultaneous measurement problem, all of which involve "generalized measurements" of one form or another; *i.e.*, relaxation of von Neumann's projection postulate: accepting that conjugate observables do not admit of simultaneous measurement with ideal devices, one undertakes to do the best that can be done with imperfect/noisy devices. I sketch several of those approaches to the solution of that problem in the next few sections of this paper.[20]

Simultaneous measurement according to Arthurs & Kelly. The system $\mathcal{S}$ under observation and a pair of detectors $\mathcal{D}_1$ and $\mathcal{D}_2$ comprise a composite system

$$\mathfrak{S} = \mathcal{S} \otimes \mathcal{D}_1 \otimes \mathcal{D}_2$$

We might, for concreteness, suppose $\mathcal{S}$ to be an oscillator; more critically, we consider $\mathcal{D}_1$ and $\mathcal{D}_2$ to be free-particle-like, except that "position" refers now not to the position of a particle but to the position of a "pointer." The initial state of the composite system is assumed to have the disentangled structure

$$|\Psi)_{\text{before}} = |\psi) \otimes |\varphi_1) \otimes |\varphi_2) \equiv |\psi)|\varphi_1)|\varphi_2)$$

which in the space/space/space representation becomes

$$(q, x, y|\Psi)_{\text{before}} = \psi(q) \otimes \varphi_1(x) \otimes \varphi_2(y) \equiv \psi(q)\varphi_1(x)\varphi_2(y)$$

Measurement is accomplished by brief (time-reversible) unitary evolution

$$|\Psi)_{\text{before}} \longrightarrow |\Psi)_t = \mathsf{U}(t)|\Psi)_{\text{before}}$$

where

$$U(t) = e^{-i\mathsf{H}t} \quad \text{with} \quad \mathsf{H} = \tfrac{1}{\tau}\big\{-\lambda_1(\mathsf{q} \otimes \mathsf{p}_1 \otimes \mathsf{I}) + \lambda_2(\mathsf{p} \otimes \mathsf{I} \otimes \mathsf{p}_2)\big\}$$

---

[19] See, for example, C. Y. She & H. Heffner[2]; S. L. Braunstein, C. M. Caves & G. J. Milburn, "Interpretation for a positive P representation," Phys. Rev A **43**, 1153-1159 (1991); Stig Stenholm, "Simultaneous measurement of conjugate variables," Annals of Physics **218**, 223-254 (1992); M. G. Raymer, "Uncertainty principle for joint measurement of noncommuting variables," AJP **62**, 986-993 (1994); U. Leonhardt, *Measuring the Quantum State of Light* (1997), Chapter 6; Yoshihisa Yamamoto & Ataç İmamoḡlu,[10] §1.4.

[20] A notational remark: I have (following Schrödinger) previously written A and B to emphasize that the operators in question are general; *i.e.*, that they may or may not be conjugate. I will henceforth write X and P when I want to emphasize that the operators in question *are* assumed to be conjugate, though they may or may not (but in physical applications usually will) signify "position" and "momentum." Traditionally I have reserved double-struck characters $\mathbb{A}$, $\mathbb{B}$, *etc.* for use when I wanted to emphasize that the objects in question were matrices. But I am at risk of running out of symbols, so abandon that convention.

Here $[\lambda_1] = (\text{length}\times\text{momentum})^{-1}$, $[\lambda_2] = (\text{momentum})^{-2}$ and τ, which controls the strength of the interaction (assumed to be so brief that Hamiltonian terms that in the absence of interaction would generate the dynamics of the system and detectors can be neglected), has dimension $[\tau] = (\text{time})^{-1}$, and I have set $\hbar = 1$. The detector momenta $\mathbf{p}_1$ and $\mathbf{p}_2$ will serve to generate spatial translations of thte pointers. Turning off the interaction at time $t = \tau$, we have

$$|\Psi)_{\text{before}} \longrightarrow |\Psi)_{\text{after}} = \mathbf{U}(\tau)|\Psi)_{\text{before}}$$

but while the measurement is in progress we (in the Schrödinger picture) have

$$\begin{aligned}\tfrac{\partial}{\partial t}\Psi(q,x,y,t) &= -i\tfrac{1}{\tau}(q,x,y|\mathbf{H}|\Psi_t)\\ &= -\tfrac{1}{\tau}\Big\{\lambda_1 q\tfrac{\partial}{\partial x} - i\lambda_2\tfrac{\partial}{\partial q}\tfrac{\partial}{\partial y}\Big\}\Psi(q,x,y,t)\\ &\Downarrow\end{aligned}$$

$$\Big\{\tfrac{\partial}{\partial t} + \tfrac{1}{\tau}\lambda_1 q\tfrac{\partial}{\partial x} - i\tfrac{1}{\tau}\lambda_2\tfrac{\partial}{\partial q}\tfrac{\partial}{\partial y}\Big\}\Psi(q,x,y,t) = 0 \tag{6}$$

subject to the initial condition $\Psi(q,x,y,0) = \psi(q)\,\varphi_1(x)\,\varphi_2(y)$ (where a couple of $\otimes$s have been surpressed). Fourier transforming with respect to y

$$\Psi(q,x,y,t) = \tfrac{1}{\sqrt{2\pi}}\int \tilde{\Psi}(q,x,k,t)\,e^{iky}dk$$

we have

$$\tfrac{1}{\sqrt{2\pi}}\int\Big\{\tfrac{\partial}{\partial t} + \tfrac{1}{\tau}\lambda_1 q\tfrac{\partial}{\partial x} + \tfrac{1}{\tau}\lambda_2 k\tfrac{\partial}{\partial q}\Big\}\tilde{\Psi}(q,x,k,t)\,e^{iky}dk = 0$$

We are informed by *Mathematica* that solutions of the first-order partial differential equation

$$\Big\{\tfrac{\partial}{\partial t} + \tfrac{1}{\tau}\lambda_1 q\tfrac{\partial}{\partial x} + \tfrac{1}{\tau}\lambda_2 k\tfrac{\partial}{\partial q}\Big\}F(q,x,t) = 0$$

are of the form

$$F(q,x,t) = \mathcal{F}\big(a(q,t), b(q,x)\big) \quad \text{where} \quad \begin{cases} a(q,t) = q - k\lambda_2\,t/\tau \\ b(q,x) = x - \lambda_1 q^2/2k\lambda_2 \end{cases}$$

We notice that $a(q,0) = q$ and to achieve x at time $t = 0$ construct

$$\begin{aligned}c(q,x,t) &= b(q,x) + (\lambda_1/2k\lambda_2)a^2(q,t)\\ &= x - \lambda_1 q\,t/\tau + \tfrac{1}{2}k\lambda_1\lambda_2(t/\tau)^2\end{aligned}$$

Initially we have $\tilde{\Psi}(q,x,k,0) = \psi(q)\,\varphi_1(x)\,\tilde{\varphi}_2(k)$ so at time $t = \tau$

$$\begin{aligned}\Psi(q,x,y,\tau) &= \tfrac{1}{\sqrt{2\pi}}\int \psi\big(a(q,\tau)\big)\varphi_1\big(c(q,x,\tau)\big)\tilde{\varphi}_2(k)\,e^{iky}dk\\ &= \tfrac{1}{\sqrt{2\pi}}\int \psi\big(q - k\lambda_2\big)\,\varphi_1\big(x - \lambda_1 q + \tfrac{1}{2}k\lambda_1\lambda_2\big)\,\tilde{\varphi}_2(k)\,e^{iky}dk \qquad (7)\end{aligned}$$

which Arthurs & Kelly are content to present without comment, though it lies at the heart of their paper.

Arthurs & Kelly assume plausibly that the initial states of the detectors are centered-Gaussian:

$$\varphi_1(x) = \Big[\tfrac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{1}{2}(x/\sigma_1)^2}\Big]^{\frac{1}{2}} \quad\text{and}\quad \varphi_2(y) = \Big[\tfrac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{1}{2}(y/\sigma_2)^2}\Big]^{\frac{1}{2}}$$

To this they bring the *ad hoc* assumption—which will acquire motivation in the course of their argument—that the Gaussians are "balanced" in the sense that $\sigma_1\sigma_2 = \frac{1}{4}$. Writing $\sigma_1 = (4\sigma_2)^{-1} = \frac{1}{2}\sqrt{b}$ (Arthurs and Kelly call b the "balance parameter") we have

$$\begin{aligned}\varphi_1(x) = \Big(\frac{2}{\pi b}\Big)^{\frac{1}{4}}e^{-x^2/b} \quad\text{and}\quad \varphi_2(y) &= \Big(\frac{2b}{\pi}\Big)^{\frac{1}{4}}e^{-by^2}\\ &= \tfrac{1}{\sqrt{2\pi}}\int\tilde{\varphi}_2(k)\,e^{iky}dk\\ \tilde{\varphi}_2(k) &= \Big(\frac{1}{2\pi b}\Big)^{\frac{1}{4}}e^{-k^2/4b}\end{aligned}$$

giving

$$\begin{aligned}\Psi(q,x,y,\tau) &= \tfrac{1}{\sqrt{2\pi}}\int\psi\big(a(q,\tau)\big)\varphi_1\big(c(q,x,\tau)\big)\tilde{\varphi}_2(k)\,e^{iky}dk\\ &= (1/8\pi^3b)^{\frac{1}{4}}\cdot\int\psi\big(q-k\lambda_2\big)\,\varphi_1\big(x-\lambda_1q+\tfrac{1}{2}k\lambda_1\lambda_2\big)e^{-k^2/4b}\,e^{iky}dk\end{aligned}$$

The λ-parameters were introduced for dimensional reasons, but to simplify the notation we henceforth assume the numerical values of both to be unity; then

$$= C_1(b)\cdot\int\psi\big(q-k\big)\,\varphi_1\big(x-q+\tfrac{1}{2}k\big)e^{-k^2/4b}\,e^{iky}dk \tag{8}$$

with $C_1(b) = (1/8\pi^3b)^{\frac{1}{4}}$.[21] Write $k \to \ell = q-k$ to introduce an alternative variable of integration, get

$$\begin{aligned}&= C_1\cdot\int\psi(\ell)\,\varphi_1\big(x-\tfrac{1}{2}(q+\ell)\big)e^{-(q-\ell)^2/4b}e^{i(q-\ell)y}d\ell\\ &\downarrow\\ |\Psi(q,x,y,\tau)| &= C_1\cdot\Big|\int\psi(\ell)\,\varphi_1\big(x-\tfrac{1}{2}(q+\ell)\big)e^{-(q-\ell)^2/4b}e^{-i\,\ell\,y}d\ell\Big|\end{aligned}$$

Drawing upon the assumed Gaussian structure of $\varphi_1(x)$ we obtain

$$\begin{aligned}&= C_1\Big(\frac{2}{\pi b}\Big)^{\frac{1}{4}}\Big|\int\psi(\ell)\exp\Big\{-\frac{(x-q)^2+(\ell-x)^2}{2b}\Big\}e^{-i\,\ell\,y}d\ell\Big|\\ &= C_2\exp\Big\{-\frac{(x-q)^2}{2b}\Big\}\cdot\Big|\int\psi(\ell)\exp\Big\{-\frac{(\ell-x)^2}{2b}\Big\}e^{-i\,\ell\,y}d\ell\Big|\end{aligned}$$

with $C_2 = C_1\cdot(2/\pi b)^{\frac{1}{4}} = (2/\pi^2b)^{\frac{1}{2}}$.

---

[21] I am indebted to Ray Mayer for the following line of argument (note taped to my door, 31 October 2012).

The position/momentum operators of the detector system $\mathcal{D}_1$ commute with those of $\mathcal{D}_1$, so projective measurements of the pointer positions x and y are compatable (can be performed simultaneously) and it makes sense to speak of the joint distribution $P(x,y)$, which is itself a conditional distribution; we have

$$\begin{aligned}P(x,y) &\equiv \int |\Psi(q,x,y,\tau)|^2 dq\\ &= C_2^2\int \exp\Big\{-\frac{(x-q)^2}{b}\Big\}dq\cdot\Big|\int \psi(\ell)\exp\Big\{-\frac{(\ell-x)^2}{2b}\Big\}e^{-i\,\ell\,y}d\ell\Big|^2\\ &= C_3\cdot\Big|\int \psi(q)\exp\Big\{-\frac{(q-x)^2}{2b}\Big\}e^{-i\,q\,y}dq\Big|^2 \qquad (9.1)\end{aligned}$$

with $C_3 = C_2^2\sqrt{\pi b} = \sqrt{1/4\pi^3 b}$ and where in the final equation I have adjusted the name $\ell \to q$ of the integration variable.

Had we (so far as $\mathcal{S}$ is concerned) elected to work not in the q-representation but in the p-representation—writing $\varPhi(p,x,y,t)$ to describe the state of the composite system—the Schrödinger equation (again set $\hbar = \lambda_1 = \lambda_2 = 1$) would have read

$$\Big\{\tfrac{\partial}{\partial t} - i\,\tfrac{1}{\tau}\tfrac{\partial}{\partial p}\tfrac{\partial}{\partial x} - \tfrac{1}{\tau}p\,\tfrac{\partial}{\partial y}\Big\}\varPhi(p,x,y,t) = 0$$

Fourier transforming with respect now to x

$$\varPhi(q,x,y,t) = \tfrac{1}{\sqrt{2\pi}}\int \tilde{\varPhi}(p,k,y,t)\,e^{ikx}dk$$

we have

$$\begin{aligned}&\Big\{\tfrac{\partial}{\partial t} + \tfrac{1}{\tau}k\tfrac{\partial}{\partial p} - \tfrac{1}{\tau}p\,\tfrac{\partial}{\partial y}\Big\}\tilde{\varPhi}(p,k,y,t) = 0\\ &\qquad\qquad\qquad\tilde{\varPhi}(p,k,y,0) = \phi(p)\,\tilde{\varphi}_1(k)\,\varphi_2(y)\end{aligned}$$

giving (again with *Mathematica*'s assistance)

$$\begin{aligned}\varPhi(p,x,y,t) &= \tfrac{1}{\sqrt{2\pi}}\int \phi(p-\tfrac{1}{\tau}kt)\,\tilde{\varphi}_1(k)\,\varphi_2(y+\tfrac{1}{\tau}pt-\tfrac{1}{2\tau^2}kt^2)\,e^{ikx}dk\\ &\Downarrow\\ &= \tfrac{1}{\sqrt{2\pi}}\int \phi(p-k)\,\tilde{\varphi}_1(k)\,\varphi_2(y+p-\tfrac{1}{2}k)\,e^{ikx}dk \quad \text{at } t=\tau\end{aligned}$$

Again invoke the assumption that the intitial detector states are Gaussian

$$\tilde{\varphi}_1(k) = \Big(\frac{b}{2\pi}\Big)^{\frac{1}{4}}e^{-\frac{bk^2}{4}} \quad\text{and}\quad \varphi_2(y) = \Big(\frac{2b}{\pi}\Big)^{\frac{1}{4}}e^{-by^2}$$

and obtain (after a change of variables $k \to p-\ell$ and some simplification)

$$\begin{aligned}|\varPhi(p,x,y,\tau)| &= C_4\Big|\int \phi(\ell)\exp\Big\{-\frac{b(y+q)^2+b(\ell+y)^2}{2}\Big\}e^{-i\,\ell\,x}d\ell\Big|\\ &= C_4\exp\Big\{-\frac{b(y+p)^2}{2}\Big\}\cdot\Big|\int \psi(\ell)\exp\Big\{-\frac{b(\ell+y)^2}{2}\Big\}e^{-i\,\ell\,x}d\ell\Big|\end{aligned}$$

where $C_4 = b/2\pi^2$. Therefore

$$P(x,y) = \int |\Phi(p,x,y,\tau)|^2 dp$$
$$= C_5 \cdot \left| \int \phi(p) \exp\left\{ -\frac{b(p+y)^2}{2} \right\} e^{-ipx} dp \right|^2 \tag{9.2}$$

where

$$C_5 = C_4^2 \int \exp\left\{ -b(y+p)^2 \right\} dp = \sqrt{b/4\pi^3}$$

and in (9.2) I have again adjusted the name $\ell \to p$ of the integration variable.[22] Equations (9) say the same thing in different ways.

Assume by way of example[23] that initially

$$\psi(q) = \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{1}{4}} e^{-q^2/4\sigma^2} = \sqrt{\text{Gaussian}}$$
$$\Updownarrow$$
$$\phi(p) = \left(\frac{2\sigma^2}{\pi} \right)^{\frac{1}{4}} e^{-p^2\sigma^2}$$

Looking to $|\psi(q)|^2$ and $|\phi(p)|^2$ we see that the data generated when q and p are subjected to independent projective measurements are expected to have variances

$$\sigma_q^2 = \sigma^2 \quad \text{and} \quad \sigma_p^2 = 1/4\sigma^2$$

We expect by the Heisenberg uncertainty principle to have $\sigma_q \sigma_p \geq \frac{1}{2}$ (recall $\hbar = 1$) but in the present instance have $\sigma_q \sigma_p = \frac{1}{2}$ since $\psi(q)$ is a minimally dispersive state. Whether we work from (9.1) or (9.2), we find

$$P(x,y) = \frac{\sigma\sqrt{2b}}{\pi(b+2\sigma^2)} \exp\left\{ -\frac{x^2 + 2b\sigma^2 y^2}{b+2\sigma^2} \right\}$$

and verify that $\iint P(x,y)dxdy = 1$. The associated marginal distributions are

$$Q(x) = \int P(x,y)dy = \frac{1}{\sqrt{\pi(b+2\sigma^2)}} \exp\left\{ -\frac{x^2}{b+2\sigma^2} \right\}$$
$$P(y) = \int P(x,y)dx = \frac{\sqrt{2b\sigma^2}}{\sqrt{\pi(b+2\sigma^2)}} \exp\left\{ -2b\sigma^2 \frac{y^2}{b+2\sigma^2} \right\}$$

---

[22] To obtain the Schrödinger equation (6)—which agrees with Arthurs & Kelly—I have been obliged to introduce a minus sign into the interaction Hamiltonian H which is absent from A & K. And because in the momentum representation q becomes $-i\partial_p$ (A & K appear to have overlooked the minus sign) I at (9.2) have $\exp\{-b(p+y)^2/2\}$ instead of their $\exp\{-b(p-y)^2/2\}$.

[23] This example has been selected because it leads to integrals that can be done in closed form.

which are in this instance seen to be Gaussian:

$$Q(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-x^2/2\sigma_x^2} \quad \text{with} \quad \sigma_x^2 = \sigma^2 + \tfrac{1}{2}b$$

$$P(y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-y^2/2\sigma_y^2} \quad \text{with} \quad \sigma_y^2 = \frac{\sigma^2 + \frac{1}{2}b}{2b\sigma^2} = \tfrac{1}{4}\big[(\sigma^2)^{-1} + (\tfrac{1}{2}b)^{-1}\big]$$

So when $\{\mathsf{q}, \mathsf{p}\}$ are subjected to a series of simultaneous A&K-measurements and the resulting detector pointer-positions $\{x, y\}$ measured projectively, the expected variances of the latter data are

$$\left.\begin{aligned} \sigma_x^2 &= \sigma_q^2 + \tfrac{1}{2}b \\ \sigma_y^2 &= \sigma_p^2 + \tfrac{1}{2}b^{-1} \end{aligned}\right\} \tag{10}$$

From

$$\frac{d(\sigma_x^2\sigma_y^2)}{db} = \frac{b^2 - 4\sigma^2}{8b^2\sigma^2} = 0 \quad \Longrightarrow \quad \tfrac{1}{2}b = \sigma^2$$

we discover (compare $\sigma_q\sigma_p \geq \frac{1}{2}$) that

$$\sigma_x\sigma_y \geq 1 \tag{11}$$

with equality if and only if the balance parameter b and ψ-structure stand in the tuned relationship $b = 2\sigma^2$.

The product structure that was assumed to pertain initially to $\Psi(q, x, y, t)$ was seen at (8) to have been lost during the course of the dynamical interaction of the system and detectors; the states of $\mathcal{S}$ and $\{\mathcal{D}_1, \mathcal{D}_1\}$ have become entangled. At the completion of the A & K measurement that caused the detector pointers to register $\{x_m, y_m\}$ the post-measurement state of $\mathcal{S}$ is

$$\psi_{\text{after}}(q; x_m, y_m) = N^{-1}\Psi(q, x_m, y_m, \tau) \quad \text{with} \quad N = \left[\int |\Psi(q, x_m, y_m, \tau)|^2 dq\right]^{\frac{1}{2}}$$

where by (8)

$$\begin{aligned} \Psi(q, x_m, y_m, \tau) = C_2 \exp\left\{-\frac{(x_m - q)^2}{2b} + iqy_m\right\} \\ \cdot\int \psi(\ell) \exp\left\{-\frac{(\ell - x_m)^2}{2b}\right\} e^{-i\ell y_m} d\ell \end{aligned}$$

The integral is a complex number: call it $Ae^{i\alpha}$. We now have

$$\Psi(q, x_m, y_m, \tau) = C_2 \exp\left\{-\frac{(x_m - q)^2}{2b} + iqy_m\right\} Ae^{i\alpha}$$

and the normalization factor becomes

$$N = C_2\left[\int \exp\left\{-\frac{(x_m - q)^2}{b}\right\} dq\right]^{\frac{1}{2}} A = C_2(\pi b)^{\frac{1}{4}} A$$

giving

$$\psi_{\text{after}}(q;x_m,y_m) = \Big(\frac{1}{\pi b}\Big)^{\frac{1}{4}} \exp\Big\{-\frac{(q-x_m)^2}{2b} + iqy_m\Big\}\cdot e^{-i\alpha} \tag{12.1}$$

where the final phase factor is unphysical and can be discarded. Fourier transforming to the momentum representation, we get

$$\phi_{\text{after}}(p;x_m,y_m) = \Big(\frac{b}{\pi}\Big)^{\frac{1}{4}} \exp\Big\{-\frac{b(p-y_m)^2}{2} - ipx_m\Big\}\cdot e^{i(x_m y_m-\alpha)} \tag{12.2}$$

where again the phase factor can be discarded. The states (12) are readily seen to be normalized: $\int|\psi_{\text{after}}(q;x_m,y_m)|^2dq = \int|\phi_{\text{after}}(p;x_m,y_m)|^2dp = 1$. The associated probability densities are Gaussian

$$|\psi_{\text{after}}(q;x_m,y_m)|^2 = \Big(\frac{1}{\pi b}\Big)^{\frac{1}{2}} \exp\Big\{-\frac{(q-x_m)^2}{b}\Big\} \tag{13.1}$$

$$|\phi_{\text{after}}(p;x_m,y_m)|^2 = \Big(\frac{b}{\pi}\Big)^{\frac{1}{2}} \exp\Big\{-b(p-y_m)^2\Big\} \tag{13.2}$$

with variances $\boldsymbol{\sigma}_q^2 = \frac{1}{2}b$ and $\boldsymbol{\sigma}_p^2 = \frac{1}{2}b^{-1}$ that (for a familiar Fourier-analytic reason, nothing more profound) satisfy $\boldsymbol{\sigma}_q\boldsymbol{\sigma}_p = \frac{1}{2}$.

The idealized projective action of an A-meter can, as we have seen, be described

$$|\psi)_{\text{before}} \longrightarrow |\psi)_{\text{after}} = \text{some normalized eigenvector } |a) \text{ of } \mathsf{A}$$

The initial state $|\psi)_{\text{before}}$ is destroyed by the measurement process, and the prepared state $|\psi)_{\text{after}}$ conveys no indication of what $|\psi)_{\text{before}}$ might have been, conceals no "memory" of $|\psi)_{\text{before}}$.[24] Note that the states (12) prepared by the A & K procedure are similar in that regard: they contain no reference to the pre-measurement $\mathcal{S}$-state $\psi(q)$. Note also that (13) supplies

$$\lim_{b\downarrow 0}|\psi_{\text{after}}(q;x_m,y_m)|^2 = \delta(q-x_m)$$
$$\lim_{b\downarrow 0}|\phi_{\text{after}}(p;x_m,y_m)|^2 = 0$$

of which the former can be interpreted to refer to the result of a projective

[24] From the statistical structure of indefinitely many such measurements one can constuct estimates of the real numbers $|(a|\psi)_{\text{before}}|$, but from that information it is still not possible in the absence of all complex phase data to reconstruct

$$|\psi)_{\text{before}} = \int |a)da(a|\psi)_{\text{before}}$$

Efforts to "measure the quantum state of $\mathcal{S}$" would appear therefore to be fundamentally misguided unless one is prepared to bring into play ideas and methods that lie beyond the reach of the von Neumann formalism. In Ulf Leonhardt's *Measuring the Quantum State of Light* (1997) the method is tomographic.

q-measurement, and the latter to provide indication that after such a measurement the expected value of **p** is indeterminate. The situation is reversed in the limit $b \uparrow \infty$.

The fact that x_m and y_m enter in distinctive ways into (12.1) is easily understood. At $y_m = 0$ we have

$$\begin{aligned}\psi_{\text{after}}(q\,; x_m, 0) &= \left(\frac{1}{\pi b}\right)^{\frac{1}{4}} \exp\left\{-\frac{(q-x_m)^2}{2b}\right\}\\ &= \text{Gaussian wavepacket in } q\text{-space}\end{aligned}$$

Introduction of the factor $\exp\{i\,q\,y_m\}$ serves to "launch" the wavepacket.[25]

The circumstance that **q** and **p** enter jointly into the A & K formalism brings to mind Wigner's "phase space formulation of quantum mechanics," wherein to every $\psi(q)$ is associated a "Wigner quasi-distribution"[26]

$$P_\psi(q,p) = (\pi\hbar)^{-1}\int \psi^*(q+\xi)\, e^{2\frac{i}{\hbar}p\xi}\, \psi(q-\xi)\, d\xi$$

Feeding (10.1) into the integral (with $\hbar$ again set to unity) we compute

$$\begin{aligned} &= \pi^{-1}\exp\left\{-\frac{(q-x_m)^2}{b}\right\}\int \exp\left\{-\frac{\xi^2 - i\,2\,b\,(p-y_m)\xi}{b}\right\}d\xi\\ &= \exp\left\{-\frac{(q-x_m)^2}{b} - b(p-y_m)^2\right\} \end{aligned} \tag{14}$$

which could hardly be prettier. We verify that $\iint P_\psi(q,p)dqdp = 1$ and observe that the Wigner distribution (14) is in fact a *proper* distribution: it becomes nowhere negative. The distribution (14) is encountered at (30) on page 14 of the notes just cited[26] and at (3.35) on page 48 of Leonhardt's monograph.[24]

The Arthurs/Kelly paper provides a solitary reference to the literature, that being to von Neumann's *Mathematical Foundations of Quantum Mechanics*. Quoting from their introductory paragraph, "Just as von Neumann uses an ideal measurement together with an interaction to explain an indirect observation, we use ideal measurements together with interactions to explain the simultaneous measurement of an observable and its conjugate." The argument to which they allude appears on the final three pages of von Neumann's classic monograph, at the end of his Chapter VI: "The Measuring Process." I provide now an account of von Neumann's argument, phrased so as to facilitatte comparison with the argument devised by Arthurs & Kelly.

Write $\Psi(q,x,t)$ to describe the devolving state of the composite system that consists of $\mathcal{S}$ (initial state $\psi(q)$) and a solitary detector $\mathcal{D}$ (initial state $\phi(x_1)$). The dynamical interaction is driven by a Hamiltonian of the form

---

[25] See (25) page 9 in my "Gaussian wavepackets" (1998).
[26] See *Advanced Quantum Topics* (2000), Chapter 2, page 10.

$$\mathsf{H} = -\tfrac{1}{\tau}\,\mathsf{q}\otimes\mathsf{p}_1$$

The Schrödinger equation[27] reads $i\,\partial_t\Psi = -\frac{1}{\tau}q(i\partial_x)\Psi$ or

$$\big\{\partial_t + \tfrac{1}{\tau}q\,\partial_x\big\}\Psi = 0$$

solutions of which are of the form $\mathcal{F}\big(q, x-\frac{1}{\tau}q\,t\big)$, which at $t=0$ becomes $\mathcal{F}\big(q,x\big)$. So

$$\begin{aligned}\Psi(q,x,t) &= \psi(q)\,\phi\big(x-\tfrac{1}{\tau}q\,t\big)\\ &\downarrow\\ &= \psi(q)\,\phi(x-q)\quad\text{at } t=\tau\end{aligned}$$

in which the $\mathcal{S}$ and $\mathcal{D}$ variables have become entangled. Assume that the initial (pre-measurement) states $\psi(q)$ and $\phi(x)$ normalized. Then

$$|\Psi(q,x,0)|^2 = 1 \xrightarrow[\text{evolution is unitary}]{} |\Psi(q,x,t)|^2 = |\psi(q)|^2\cdot|\phi(x-q)|^2 = 1$$

and the marginal q-probability is

$$P(q) = |\psi(q)|^2\int|\phi(x-q)|^2dx = |\psi(q)|^2$$

von Neumann remarks parenthetically that because q and x are continuous they can be measured "with arbitrary but not with absolute precision." Suppose $\phi(x)$ is—like (say) a narrow Gaussian—non-zero only in the immediate neighborhood of the origin $(-\delta < x < +\delta)$. Suppose, moreover, that a projective inspection of the detector at time τ shows the pointer to be at x_m. We can conclude that the post-measurement probability density $|\psi_{\text{after}}(q)|^2$ is localized at $q\approx x_m$

$$\int_{x_m-\delta}^{x_m+\delta}|\psi_{\text{after}}(q)|^2\,dq \approx 1$$

and that the prepared state $\psi_{\text{after}}(q)$ of $\mathcal{S}$ is of the form

$$\psi_{\text{after}}(q) = f(q)^{\,i\,\alpha(q)}$$

where concening the localized function we know only that

$$\int_{x_m-\delta}^{x_m+\delta} f^2(q)\,dq \approx \int_{-\infty}^{+\infty} f^2(q)\,dq = 1$$

while $\alpha(q)$ remains entirely undetermined.

---

[27] I write p_1 to distinguish the momentum of $\mathcal{D}$ from that of $\mathcal{S}$, but will drop the pedanatic subscript from x_1. As has been our custom, we set $\hbar = 1$, though in the present instance the $\hbar$-factors would cancel anyway.

Abstract essentials of the Arthurs/Kelly formalism. Familiarly, one cannot assign *simultaneously precise* values to the position and momentum of a quantum state because $[\mathbf{q},\mathbf{p}] \neq \mathbf{0}$.[28] Arthurs & Kelly, building upon the framework erected by von Neumann, looked therefore to the construction of states to which *simultaneously imprecise* position/momentum values can be assigned in a best-possible way.[29] In the final paragraphs of their paper they sketch an abstract generalization of their simultaneous measurement procedure. Leonhardt, in §6.1.1 of Chapter 6 ("Simultaneoous measurment of position and momentum") in the monograph previously cited,[24] presents a clarified paraphrase of the A/K argument upon which I base the following remarks.

Introduce near-variants of $\mathbf{q}$ and $\mathbf{p}$

$$\begin{aligned}\mathbf{Q} &= \mathbf{q} + \mathbf{A}\\ \mathbf{P} &= \mathbf{p} + \mathbf{B}\end{aligned}$$

where the "noise terms" $\mathbf{A}$ and $\mathbf{B}$ model the imprecision which we propose to introduce into the measurement process. Impose upon $\{\mathbf{Q},\mathbf{P}\}$ the simultaneous measureability requirement $[\mathbf{Q},\mathbf{P}] = [\mathbf{q},\mathbf{p}]+[\mathbf{q},\mathbf{B}]+[\mathbf{A},\mathbf{p}]+[\mathbf{A},\mathbf{B}] = \mathbf{0}$, which (recall $\hbar = 1$) we write

$$[\mathbf{A},\mathbf{B}] = -i\,\mathbf{I} - [\mathbf{q},\mathbf{B}] - [\mathbf{A},\mathbf{p}] \tag{15}$$

Look now to the noisy analog $\Delta^2Q\Delta^2P$ of $\Delta^2q\Delta^2p \geq \frac{1}{4}$. We have

$$\begin{aligned}\Delta^2Q &= \big\langle(\mathbf{q}+\mathbf{A})^2\big\rangle - \big\langle\mathbf{q}+\mathbf{A}\big\rangle^2\\ &= \big\langle\mathbf{q}^2\big\rangle - \big\langle\mathbf{q}\big\rangle^2 + \big\langle\mathbf{q}\,\mathbf{A}\big\rangle - \big\langle\mathbf{q}\big\rangle\big\langle\mathbf{A}\big\rangle + \big\langle\mathbf{A}\,\mathbf{q}\big\rangle - \big\langle\mathbf{A}\big\rangle\big\langle\mathbf{q}\big\rangle + \big\langle\mathbf{A}^2\big\rangle - \big\langle\mathbf{A}\big\rangle^2\end{aligned}$$

and if we impose upon the noise operator $\mathbf{A}$ the (plausible?) assumption that for all states $\langle\mathbf{A}\rangle = \langle\mathbf{q}\,\mathbf{A}\rangle = \langle\mathbf{A}\,\mathbf{q}\rangle = 0$ obtain (compare (10))

$$\Delta^2Q = \Delta^2q + \big\langle\mathbf{A}^2\big\rangle \tag{16.1}$$

Similarly

$$\Delta^2P = \Delta^2p + \big\langle\mathbf{B}^2\big\rangle \tag{16.2}$$

Therefore

$$\Delta^2Q\,\Delta^2P = \Delta^2q\,\Delta^2p + 2\,\frac{\Delta^2q\big\langle\mathbf{B}^2\big\rangle + \Delta^2p\big\langle\mathbf{A}^2\big\rangle}{2} + \big\langle\mathbf{A}^2\big\rangle\big\langle\mathbf{B}^2\big\rangle$$

Leonhardt draws at this point upon the fact the if a and b are non-negative real numbers then

$$\frac{a+b}{2} \geq \sqrt{ab} \quad : \quad \text{“arithmetic mean dominates geometric mean”}$$

with equality if and only if $a = b$ (this is the simplest instance of a large class

[28] That statement pertains, of course, to *any* pair (or expanded set) of non-commutative observables, whether or not they happen to be "conjugate" in the sense $[\mathbf{A},\mathbf{B}] = i\,\mathbf{I}$.

[29] Look back, in this light, to (12).

of lovely inequalities of which the Wikipedia article "Inequality of arithmetic and geometric means" provides a good account)... to write

$$\begin{aligned}\Delta^2 Q\,\Delta^2 P &\geq \Delta^2 q\,\Delta^2 p + 2\,\Delta q\,\Delta p\sqrt{\big\langle \mathsf{A}^2\big\rangle\big\langle \mathsf{B}^2\big\rangle} + \big\langle \mathsf{A}^2\big\rangle\big\langle \mathsf{B}^2\big\rangle\\ &= \Big[\Delta q\,\Delta p + \sqrt{\big\langle \mathsf{A}^2\big\rangle\big\langle \mathsf{B}^2\big\rangle}\,\Big]^2\\ &\Downarrow\\ \Delta Q\,\Delta P &\geq \tfrac{1}{2} + \sqrt{\big\langle \mathsf{A}^2\big\rangle\big\langle \mathsf{B}^2\big\rangle}\end{aligned}$$

with equality if and only if $\Delta^2 q\big\langle \mathsf{B}^2\big\rangle = \Delta^2 p\big\langle \mathsf{A}^2\big\rangle$. Retaining the presumption that $\langle \mathsf{A}\rangle = \langle \mathsf{B}\rangle = 0$ we have $\Delta^2 A = \big\langle \mathsf{A}^2\big\rangle$ and $\Delta^2 B = \big\langle \mathsf{B}^2\big\rangle$, so by Schrödinger's inequality (1.1)

$$\big\langle \mathsf{A}^2\big\rangle\big\langle \mathsf{B}^2\big\rangle \geq \Big\langle \frac{[\mathsf{A},\mathsf{B}]}{2i}\Big\rangle^2$$

which by (15) becomes

$$\begin{aligned}&= \bigg[\frac{-\big\langle i\mathsf{I}\big\rangle - \big\langle[\mathsf{q},\mathsf{B}]\big\rangle + \big\langle[\mathsf{p},\mathsf{A}]\big\rangle}{2i}\bigg]^2\\ &= \Big[-\tfrac{1}{2} - 0 + 0\Big]^2 = \tfrac{1}{4}\end{aligned}\tag{17}$$

giving

$$\begin{aligned}\Delta Q\,\Delta P &\geq \Delta q\,\Delta p + \tfrac{1}{2}\\ &\geq 1\end{aligned}\tag{18}$$

We have here reproduced—by a relatively more abstract/general line of argument—a fundamental result that on page 16 was phrased $\sigma_x\sigma_y \geq 1$. It serves to quantify the price one necessarily pays for simultaneous measurement.

The quantum states (12) that result from application of the A/K model of a simultaneous measurement process are "minimal dispersion" states, in the sense mentioned on page 7. They give rise to Wigner distributions (14) that serve *via*

$$b^{-1}(x - x_m)^2 + b(p - y_m)^2 = \text{constant}$$

to inscribe $\{x_m, y_m\}$-centered curves on phase space—curves that are circular if $b = 1$ and otherwise elliptical ("squeezed"). Noting the structural similarity of (10) and (16), we observe the the product of the dangling terms in (10) is $(\frac{1}{2}b)(\frac{1}{2}b^{-1}) = \frac{1}{4}$ while the associated product in (16) is (in the optimal case: see (17)) $\big\langle \mathsf{A}^2\big\rangle\big\langle \mathsf{B}^2\big\rangle = \frac{1}{4}$. Since the shape of the Wigner ellipse is controlled by $b = \big[(\frac{1}{2}b)/(\frac{1}{2}b^{-1})\big]^{1/2}$ it becomes natural to assign similar significance to the ratio $\big[\big\langle \mathsf{A}^2\big\rangle/\big\langle \mathsf{B}^2\big\rangle\big]^{1/2}$, which by Leonhardt's optimality condition $\Delta^2 q\big\langle \mathsf{B}^2\big\rangle = \Delta^2 p\big\langle \mathsf{A}^2\big\rangle$ becomes $\big[\Delta q/\Delta p\big]^{1/4}$. And indeed, what Leonhardt and others call the "sqeezing parameter" is defined

$$\begin{aligned}\zeta &= \tfrac{1}{4}\log(\Delta p/\Delta q) = \tfrac{1}{4}\log\big(\big\langle \mathsf{B}^2\big\rangle/\big\langle \mathsf{A}^2\big\rangle\big)\\ &= 0 \text{ in the circular case (no sqeezing)}\end{aligned}$$

Simultaneous measurements of conjugate observables are necessarily imperfect measurements. The question arises: Can the formal devices employed by Arthurs and Kelly be used to describe imperfect measurements of a *single* observable—a *noisy* von Neumann process? Write

$$\begin{gathered}\mathbf{Q} = \mathbf{q} + \mathbf{A}\\ \Downarrow\\ \mathbf{Q}^2 = \mathbf{q}^2 + \mathbf{q}\,\mathbf{A} + \mathbf{A}\,\mathbf{q} + \mathbf{A}^2\end{gathered}$$

Assume (as Arthurs/Kelly did) that $\langle\mathbf{A}\rangle = \langle\mathbf{q}\,\mathbf{A}\rangle = \langle\mathbf{A}\,\mathbf{q}\rangle = 0$ for all $|\psi)$. Then

$$\Delta^2 Q = \Delta^2 q + \Delta^2 A$$

Such a theory is too impoverished to set a natural bound on $\Delta^2 A$. And it exposes with stark clarity a problem that bedevils also Arthurs/Kelly's formal theory of simultaneous measurement: If (as was implicitly assumed at every step, as when we appealed to Schrödinger's inequality) $\mathbf{A}$ refers to a self-adjoint operator

$$\mathbf{A} = \int |a)a\,da(a|$$

then how is it possible to achieve $\langle\mathbf{A}\rangle = 0$ for all $|\psi)$? That would require that every $|\psi)$ lies in the null space of $\mathbf{A}$ (effectively: $\mathbf{A} = \mathbf{0}$), and would entail

$$\Delta^2 A = \langle\mathbf{A}^2\rangle = 0$$

I am inclined, therefore, to dismiss Arthurs/Kelly's formal theory as a *suggestive hoax*, a provocative idea that stands in need of more careful development. Curiously, the fundamental defect to which I have drawn attention does not appear to offend Leonhardt, whose frequent reference to "fluctuations" seems intended to forgive all sins.

Arthurs/Kelly's dynamical model is susceptible to criticisms of a different sort. Since presented as the analysis of an idealized "gedanken experiment," we can dismiss as irrelevant the circumstance that A/K provide no indication of how their interaction Hamiltonian might be realized physically, or of how the interaction is to be switched off at time $t = \tau$.[30] More significantly, their *ad hoc* assumption that the initial detector states are balanced Gaussians remains—though central to the analytical details of their paper—quite unmotivated.

Construction of conjugate pairs & the number-phase problem. Can Arthurs/Kelly's gedanken experiment be realized as a physical experiment? In a pedagogical paper—valuable not least for its extensive bibliography, and in which no actual experimental results are reported—Michael Raymer[31] shows

[30] The latter criticism can be made also of von Neumann's model, and of many theoretical contributions to the quantum computation literature.

[31] "Uncertainty principle for joint measurement of noncommuting variables," AJP **62**, 986-993 (1994).

how a simple single slit set-up might be used to accomplish a simultaneous measurement of position and momentum. Raymer emphasizes that such measurements are necessarily *imprecise* measurements. He proceeds without explicit reference to the Hamiltonian-generated interaction of system (in this instance a projected particle) and detectors by means of a formalism that is, as it happens, quite similar in its essentials to the formalism I once had occasion to sketch,[32] but which he works out in much greater detail than I attempted, and carries to a point where he becomes able establish contact with Arthurs & Kelly.

Quantum optics provides the conceptual and experimental apparatus that is used most commonly to probe the foundations of quantum theory.[33] It is, therefore, not surprising that Leonhardt devotes the greater part of his final Chapter 6 ("Simultaneous measurement of position and momentum") to discussion of some relevant quantum optical issues. Central to quantum optical theory are non-hermitian operators $\{\mathsf{a}_k, \mathsf{a}_k^+\}$ that permit one to move around in the Foch space of quantum field modes. Those operators satisfy commutations relations of a form $[\mathsf{a}, \mathsf{a}^+] = \mathsf{I}$ identical to the $[\mathsf{W}, \mathsf{W}^+] = \mathsf{I}$ satisfied by operators that were encountered already on page 4. Running in reverse the remark that motivated the introduction of W, we observe that the hermitian operators

$$\mathsf{q} = \tfrac{1}{\sqrt{2}}(\mathsf{W}^+ + \mathsf{W}) \quad \text{and} \quad \mathsf{p} = i\tfrac{1}{\sqrt{2}}(\mathsf{W}^+ - \mathsf{W})$$

are conjugate

$$[\mathsf{q}, \mathsf{p}] = i\,\mathsf{I}$$

and (when decorated with suitably-dimensioned factors) can be interpreted to comprise "position" and "momentum" operators in whatever context they arise. In quantum optics one writes

$$\mathsf{X}_1 = \tfrac{1}{\sqrt{2}}(\mathsf{a}^+ + \mathsf{a}) \quad \text{and} \quad \mathsf{X}_2 = i\tfrac{1}{\sqrt{2}}(\mathsf{a}^+ - \mathsf{a})$$

and calls $\{\mathsf{X}_1, \mathsf{X}_2\}$ "quadrature operators."[34]

---

[32] "Quantum measurement with imperfect devices," notes for a Reed College Physics Seminar presented 16 February 2000. See also my "First steps toward a theory of imperfect meters" on pages 10-11 in "Generalized Quantum Measurement: Imperfect Meters and POVMs" (September 2012).

[33] Raymer—prolific founding director of the Oregon Center for Optics—is a leading experimentalist in the field, which may account for his interest in the work of Arthurs & Kelly; shortly before the 1996 AJP paper was written he and his group had completed some trail-blazing in a closely related area: see M. Beck, D. T. Smithey, J. Cooper & M. G. Raymer, "Experimental determination of number-phase uncertainty relations," Opt. Lett **18**, 1259 (1993); "Measurement of number-phase uncertainty relations of optical fields," Phys. Rev. A **48**, 3159 (1993).

[34] See, for example, C. Gerry & P. Knight, *Introductory Quantum Optics* (2005), page 17. The X_i are, in effect, conjugate coordinates of the "field oscillators" (oscillatory field modes).

Experimentalists—whether they propose to proceed quantum optically of mechanically—might look to established uncertainty relations

position - momentum
time - energy
angle - angular momentum
phase - number

for contexts within which to work. Relatedly, they might notice that conjugate operators q and p can be used to construct conjugate pairs

$$\mathsf{Q} = Q(\mathsf{q},\mathsf{p}) \quad \text{and} \quad \mathsf{P} = P(\mathsf{q},\mathsf{p})$$

in a seemingly infinitely many ways, of which some will be more useful than others. For example,

$$[\mathsf{q},\mathsf{p}] = i\,\mathsf{I} \quad \Longrightarrow \quad \begin{cases} [\mathsf{Q},\mathsf{P}] = i\,\mathsf{I} & \text{with} \quad \mathsf{Q} = \mathsf{q} + f(\mathsf{p}),\ \mathsf{P} = \mathsf{p} \\ [\mathsf{Q},\mathsf{P}] = i\,\mathsf{I} & \text{with} \quad \mathsf{Q} = \mathsf{q},\ \mathsf{P} = \mathsf{p} + g(\mathsf{q}) \end{cases}$$

But Weyl showed long ago that it all realizations of the fundamental commutator are unitarily equivalent: it is always possible to write

$$\mathsf{Q} = \mathsf{U}\,\mathsf{q}\,\mathsf{U}^{-1}, \quad \mathsf{P} = \mathsf{U}\,\mathsf{p}\,\mathsf{U}^{-1} \quad \text{with } \mathsf{U} \text{ unitary}$$

(good news, since otherwise quantum mechanics would split into disjoint fragments). If, for example, we set $\mathsf{U} = e^{iG(\mathsf{q})}$ we get $\mathsf{Q} = \mathsf{q}$ and $\mathsf{P} = \mathsf{p} + G'(\mathsf{q})$. There are, however, aspects of the $\{\mathsf{q},\mathsf{p}\} \longrightarrow \{\mathsf{Q},\mathsf{P}\}$ process that merit closer scrutiny. Here, since commutators become Poisson brackets in the classical limit, the classical theory of canonical transformations provides useful guidance. Look, for example, to the Hamiltonian dynamics of a free particle: $H(q,p) = \frac{1}{2m}p^2$. Energy E is conserved, $H(q,p) = E$ inscribes a curve (actually a straight line of constant $p = \sqrt{2mE}$) on the phase plane, along which the state point moves with constant velocity $v = \sqrt{2E/m}$, so

$$\begin{aligned} T(q,p) &= \text{transit time } \{0,p\} \longrightarrow \{q,p\} \\ &= \frac{q}{\sqrt{2E/m}} = mq/p \end{aligned}$$

Looking to the "time-energy bracket" (analog of the "position-momentum bracket[35])

$$[T(q,p), H(q,p)] = \frac{\partial T}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial H}{\partial q}\frac{\partial T}{\partial p}$$

we find by computation that $T(q,p)$ and $H(q,p)$ are *conjugate* observables:

$$[T(q,p), H(q,p)] = 1$$

[35] Note that dimensionally [time · energy] = [length · momentum] = [action].

We are motivated by this result to introduce quantum observables[36]

$$\mathbf{H} = \tfrac{1}{2m}\mathbf{p}^2 \quad \text{and} \quad \mathbf{T} = m\tfrac{1}{2}(\mathbf{q}\,\mathbf{p}^{-1} + \mathbf{p}^{-1}\mathbf{q})$$

where we have used the simplest means to "hermitianize" $\mathbf{T}$. Drawing formally upon the identity $[\mathbf{A}, \mathbf{B}\,\mathbf{C}] = [\mathbf{A}, \mathbf{B}]\mathbf{C} + \mathbf{B}[\mathbf{A}, \mathbf{C}]$ we compute

$$[\mathbf{T}, \mathbf{H}] = [\mathbf{q}, \mathbf{p}] = i\,\mathbf{I}$$

This attractive result is, however, subject to several serious criticisms. In the first place, we have too casually assumed that $[\mathbf{p}^2, \mathbf{p}^n] = \mathbf{0}$ remains valid even when $n < 0$. In the Schrödinger representation we expect to have $\mathbf{p}^{-1} = \int^q$, or actually $\mathbf{p}^{-1} = \int_a^q$ since indefinite integrals are defined only to within an additive constant. But

$$\partial_q \int_a^q f(q)dq - \int_a^q \partial_q f(q)dq = f(a)$$

so $[\mathbf{p}, \mathbf{p}^{-1}]f(q) = 0$ pertains only to functions that vanish at the fiducial point: $f(a) = 0$. More generally, $[\mathbf{p}^n, \mathbf{p}^{-1}]f(q) = 0$ iff $f^{(n-1)}(a) = 0$. And we confront also a second problem: for $\mathbf{H}$ and $\mathbf{T}$ to be of any quantum mechanical utility they must be self-adjoint. From

$$\partial(\psi\,\partial\phi) - \partial(\phi\,\partial\psi) = \psi\partial^2\phi - \phi\partial^2\psi$$

we see that $\mathbf{H}$ is self-adjoint $(\psi|\mathbf{H}\phi) = \overline{(\phi|\mathbf{H}\psi)}$ only with respect to functions that satisfy boundary conditions that insure

$$\int_a^b \big\{\partial(\bar{\psi}\,\partial\phi) - \partial(\phi\,\partial\bar{\psi})\big\}dq = \big\{\bar{\psi}\,\partial\phi - \phi\,\partial\bar{\psi}\big\}\Big|_a^b = 0$$

which is achieved most simply (and most commonly) by imposition of periodic or box boundary conditions. The self-adjointness of the time operator $\mathbf{T}$ imposes, however, a condition

$$\int_a^b \bar{\psi}(q)\bigg[q\int_a^q \phi(x)dx + \int_a^q x\,\phi(x)dx\bigg]dq = \int_a^b \phi(q)\bigg[q\int_a^q \bar{\psi}(x)dx + \int_a^q x\,\bar{\psi}(x)dx\bigg]dq$$

the implications of which are much more obscure, but which appears on its face to be much more restrictive. Moreover, $\mathbf{T}|\tau) = \lambda|\tau)$ in the q-representation reads

$$\bigg[q\int_a^q \tau(x)dx + \int_a^q x\,\tau(x)dx\bigg] = \lambda\tau(q)$$

$$\Downarrow$$

$$2q\tau'(q) + 3\tau(q) = \lambda\tau''(q)$$

[36] It is my dim recollection that the following $\mathbf{T}$ operator was considered long ago by Bohm and Aharanov.

Mathematica reports that solutions are linear combinations

$$\tau(q) = c_1\text{HermiteH}[-\tfrac{3}{2},\ q/\sqrt{\lambda}] + c_2\text{Hypergeometric1F1}[\tfrac{3}{4},\ \tfrac{1}{2},\ q^2/\lambda]$$

of Hermite "polynomials" of a negative fractional order and of a certain confluent hypergeometric function. But both are seen when plotted to be too unruly to make sense of. And anyway, what meaning could one attach to a temporal eigenvalue, or a temporal eigenfunction? Pretty clearly, the conditions which insure the self-adjointness of $\mathbf{H}$ do not—even in this simplest of cases—imply the automatic self-adjointness of $\mathbf{T}$. We therefore *cannot* conclude from Schrödinger's inequality that $\Delta H \Delta T \geq \frac{1}{2}\hbar$.

Look more generally to the classical system

$$H(q,p) = \tfrac{1}{2m}p^2 + U(q)$$

Energy conservation $E = \frac{1}{2}m\dot{q}^2 + U(q)$ leads to the transit-time construction[37]

$$\int dt = \sqrt{\tfrac{m}{2}}\int \frac{1}{\sqrt{E - U(x)}}\,dx$$

which motivates the definition

$$T(q,p) = \sqrt{\tfrac{m}{2}}\int^q \frac{1}{\sqrt{H(q,p) - U(x)}}\,dx$$

Computation now as before supplies $[T(q,p), H(q,p)] = 1$. In the associated quantum theory we have $\mathbf{H} = H(\mathbf{q}, \mathbf{p}) = \frac{1}{2m}\mathbf{p}^2 + U(\mathbf{q})$, but the specific meaning of

$$\mathbf{T} = \sqrt{\tfrac{m}{2}}\int^{\mathbf{q}} \frac{1}{\sqrt{H(\mathbf{q},\mathbf{p}) - U(x)\mathbf{I}}}\,dx$$

is not obvious,[38] nor is it clear how one would undertake to demonstrate the conjugacy of $\{\mathbf{T}, \mathbf{H}\}$. It is, however, already clear that we cannot expect $\mathbf{T}$ to be self-adjoint with respect to the functions that render $\mathbf{H}$ self-adjoint.

Look now more particularly to the classical oscillator: $U(q) = \frac{1}{2}m\omega^2q^2$. From $E = \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2q^2$ we see that at the turning points (where $p = 0$ and the energy is entirely potential) $E = \frac{1}{2}m\omega^2a^2$ where a is the amplitude of the oscillation. The transit time $0 \to q \leq a$ is

$$\begin{aligned} T(q,p) &= \sqrt{\tfrac{m}{2}}\int_0^q \frac{1}{\sqrt{H - \frac{1}{2}m\omega^2x^2}}\,dx\bigg|_{H = \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2q^2} \\ &= \tfrac{1}{\omega}\arctan(m\omega q/p) \end{aligned}$$

[37] See my *Classical Mechanics* (1983/84), page 269.

[38] The Weyl transform (see *Advanced Quantum Topics* (2000), Chapter 2, pages 4-7) supplies one possible way to proceed $T(q,p) \to \mathbf{T}$, but while attractive in principal it looks to be unworkable in the present general context.

Set $m = \omega = 1$ to reduce notational clutter, get

$$\begin{aligned} H(q,p) &= \tfrac{1}{2}(q^2 + p^2) \\ T(q,p) &= \arctan(q/p) \end{aligned}$$

and verify that the Poisson bracket $[T, H\,] = 1$. The equations

$$H(q,p) = \text{constant} \quad \text{and} \quad T(q,p) = \text{constant}$$

inscribe circles/rays on the phase plane, so the canonical transformation

$$\{q, p\} \longrightarrow \{T(q,p), H(q,p)\}$$

is in effect a transformation from Cartesian to polar coordinates. Oscillators are "clocks," in the sense that the moving phase point $\{q(t), p(t)\}$ traces a circular (generally elliptical) orbit with constant (energy-independent) angular velocity; Hamilton's canonical equations $\frac{d}{dt}A = [A, H\,]$ supply

$$\tfrac{d}{dt}T = [T, H\,] = 1 \quad \Longrightarrow \quad T_t = T_0 + t$$

so in this context "time" and "angle" are equivalent notions.[39] Useful insight into the origin of the conjugacy statement $[T, H\,] = 1$ follows from the observation that if $\xi(q,p) = q/p$ then

$$[\xi, H\,] = 1 + \xi^2$$

More generally, we have—either by direct calculation or by appeal to the general identity $[A, g(B)] = [A, B]\,g'(B)$

$$[\xi^n, H\,] = n(\xi^{n-1} + \xi^{n+1}) \quad : \quad n = 1, 2, 3, \ldots$$

We have $\arctan \xi = \xi - \frac{1}{3}\xi^3 + \frac{1}{5}\xi^5 - \frac{1}{7}\xi^7 + \cdots = \sum_{n=0}^{\infty}(-)^n \frac{1}{2n+1}\xi^{2n+1}$ so

$$\big[\textstyle\sum_{n=0}^{N}(-)^n \frac{1}{2n+1}\xi^{2n+1}, H\big] = 1 + (-)^N \xi^{2N+2}$$

While $\lim\limits_{N\to\infty}$(expression on the left) $= [\arctan \xi, H]$, it would be difficult to argue that the expression on the right $\to 1$. This appears to comprise yet another indication that time/angle observables are pathological beasts.

Turning now to the associated quantum theory,[40] from

$$\left.\begin{aligned} \mathsf{q} &= \tfrac{1}{\sqrt{2}}(\mathsf{a}^+ + \mathsf{a}) \\ \mathsf{p} &= i\tfrac{1}{\sqrt{2}}(\mathsf{a}^+ - \mathsf{a}) \end{aligned}\right\} \Longleftrightarrow \left\{\begin{aligned} \mathsf{a} &= \tfrac{1}{\sqrt{2}}(\mathsf{q} + i\,\mathsf{p}) \\ \mathsf{a}^+ &= \tfrac{1}{\sqrt{2}}(\mathsf{q} - i\,\mathsf{p}) \end{aligned}\right.$$

---

[39] When we reinstate the dimensioned physical parameters $\{m, \omega\}$ we get $T_t = T_0 + \omega t$; "time" and "angle" become dimensionally distinct, but retain their proportionality.

[40] I borrow material from pages 4–7, but with notation adjusted

$$\{\mathsf{A}, \mathsf{B}, \mathsf{W}, \mathsf{W}^+\} \to \{\mathsf{q}, \mathsf{p}, \mathsf{a}, \mathsf{a}^+\}$$

to conform to quantum optical convention.

we have

$$\mathsf{N} = \mathsf{a}^+\mathsf{a} = \tfrac{1}{2}(\mathsf{q}^2 + \mathsf{p}^2 - \mathsf{I}) = \mathsf{H} - \tfrac{1}{2}\mathsf{I} = \mathsf{a}\,\mathsf{a}^+ - \mathsf{I}$$

which differs from H only by an additive (zero-point energy) term, and in quantum optical (bosonic quantum field-theoretic) contexts is called the "number operator." It's eigenvalues $n = \{0, 1, 2, \ldots\}$ indicate how many field quanta occupy the mode in question. The conjugate operator $\boldsymbol{\varPhi}$ is interpreted to refer not to "time" but to angular "phase." One might expect to be able to write

$$[\boldsymbol{\varPhi}, \mathsf{N}] = i\,\mathsf{I} \quad \text{with} \quad \boldsymbol{\varPhi} = \arctan\boldsymbol{\xi}$$
$$\boldsymbol{\xi} = \tfrac{1}{2}(\mathsf{q}\,\mathsf{p}^{-1} + \mathsf{p}^{-1}\mathsf{q})$$

but several serious problems immediately arise: (i) The meaning of

$$\mathsf{p}^{-1} = -i\sqrt{2}\,(\mathsf{a}^+ - \mathsf{a})^{-1}$$

is unclear. (ii) To lend meaning to the symbol $\arctan\boldsymbol{\xi}$ we might write

$$\arctan\boldsymbol{\xi} = \lim_{N\to\infty}\sum_{n=0}^{N}(-)^n\tfrac{1}{2n+1}\,\boldsymbol{\xi}^{2n+1}$$

but then confront

$$[\boldsymbol{\varPhi}, \mathsf{H}] = \lim_{N\to\infty}\sum_{n=0}^{N}(-)^n\tfrac{1}{2n+1}\,[\boldsymbol{\xi}^{2n+1}, \mathsf{H}\,]$$

Iteration of the general identity $[\mathsf{A}, \mathsf{B}\,\mathsf{C}] = [\mathsf{A}, \mathsf{B}]\mathsf{C} + \mathsf{B}[\mathsf{A}, \mathsf{C}]$ supplies

$$[\boldsymbol{\xi}^{2n+1}, \mathsf{H}] = \sum_{k=0}^{2n}\boldsymbol{\xi}^k[\boldsymbol{\xi}, \mathsf{H}]\boldsymbol{\xi}^{2n-k}$$

Non-commutivity has produced here such a mess that it appears to be unfeasible to demonstrate even formally (convergence questions aside) that $[\boldsymbol{\varPhi}, \mathsf{H}] = i\,\mathsf{I}$. (iii) If $\boldsymbol{\varPhi}$ is so complicated as it now appears to be then proof (in whatever sense turns out to be meaningful) of the self-adjointness of $\boldsymbol{\varPhi}$ appears to be quite out of the question.

Many attempts to solve the "phase operator problem" have been devised over the years.[41] I mention only one. In the paper[42] in which Dirac reported his first attempt to construct a quantum electrodynamics he assumed that it is possible to write

$$\mathsf{a}^+ = \sqrt{\mathsf{N}}\,e^{i\boldsymbol{\varPhi}} \quad \Longleftrightarrow \quad \mathsf{a} = e^{-i\boldsymbol{\varPhi}}\sqrt{\mathsf{N}} \tag{19}$$

---

[41] For an exhaustive review, see P. Carruthers & M. M. Nieto, "Phase and angle variables in quantum mechanics," Rev. Mod. Phys. **40**, 411-440 (1968). A very nice brief account can be found in §2.7 of Gerry & Knight.[31]

[42] Proc. Roy. Soc. (London) **A114**, 243 (1927).

which presumes $\boldsymbol{\Phi}^+ = \boldsymbol{\Phi}$. Then $\mathsf{a}^+\mathsf{a} = \mathsf{N}$ while $\mathsf{a}\,\mathsf{a}^+ - \mathsf{a}^+\mathsf{a} = \mathsf{I}$ becomes

$$e^{-i\boldsymbol{\Phi}}\mathsf{N}e^{i\boldsymbol{\Phi}} - \mathsf{N} = \mathsf{I}$$

giving $e^{-i\boldsymbol{\Phi}}\mathsf{N} - \mathsf{N}e^{-i\boldsymbol{\Phi}} = e^{-i\boldsymbol{\Phi}}$. Expand the exponentials and get

$$\begin{aligned} -i[\boldsymbol{\Phi},\mathsf{N}] - \tfrac{1}{2!}[\boldsymbol{\Phi}^2,\mathsf{N}] + \boldsymbol{i}\tfrac{1}{3!}[\boldsymbol{\Phi}^3,\mathsf{N}] + \tfrac{1}{4!}[\boldsymbol{\Phi}^4,\mathsf{N}] - \cdots \\ = \mathsf{I} - i\boldsymbol{\Phi} - \tfrac{1}{2!}\boldsymbol{\Phi}^2 + i\tfrac{1}{3!}\boldsymbol{\Phi}^3 + \tfrac{1}{4!}\boldsymbol{\Phi}^4 - \cdots \end{aligned}$$

Identification of the leading terms supplies

$$[\boldsymbol{\Phi},\mathsf{N}] = i\,\mathsf{I}$$

from which it follows that $[\boldsymbol{\Phi}^k,\mathsf{N}] = ik\boldsymbol{\Phi}^{k-1}$, which serve to establish

$$k^{\text{th}}\text{ term on left} = k^{\text{th}}\text{ term on right} \quad : \quad k = 2, 3, 4, \ldots$$

Dirac's argument achieves its elegance by *postulating* the existence of hermitian operators $\sqrt{\mathsf{N}}$ and $\boldsymbol{\Phi}$ such that $\mathsf{a}^+ = \sqrt{\mathsf{N}}\,e^{i\boldsymbol{\Phi}}$, but has nothing to say about their explicit construction and supplies no explicit proof of self-adjointness. And, as was first pointed out by L. Susskin & J. Glogower,[43] Dirac's argument leads to contradictions. The root problem, as I see it, is that $e^{i\boldsymbol{\Phi}} = \mathsf{N}^{-\frac{1}{2}}\mathsf{a}^+$ requires $\mathsf{N}^{-\frac{1}{2}}\mathsf{a}^+$ to be unitary. But

$$\mathsf{a}\,\mathsf{N}^{-1}\mathsf{a}^+ \neq \mathsf{I}$$

because $\mathsf{N} = \sum_0^\infty |n)n(n|$ and $\mathsf{N}|0) = 0$ imply that N is singular: N^{-1} does not exist. Dirac's assumption (19) is untenable.

The short of it is this: In the absence well-defined self-adjoint "time" and "phase" operators T and $\boldsymbol{\Phi}$—conjugate respectively to the Hamiltonian and number operators H and N—it is impossible to construe statements of the forms $\Delta E\,\Delta T \geq \frac{1}{2}\hbar$ and $\Delta N\,\Delta\Phi \geq \frac{1}{2}\hbar$ to be instances of Schrödinger's inequality. And it certainly impossible to contemplate the "simultaneous measurement" of energy/time or number/phase. Indeed, it is impossible to speak of "time measurements" or "phase measurements" in *any* standard quantum mechanical sense. The above uncertainty relations—which (particularly the former) are of undeniably great practical importance—must arise from considerations and procedures entirely separate from those of quantum measurement theory.[33] The problems that bedevil the construction of T and $\boldsymbol{\Phi}$ have been seen to be essentially identical. It is—when one thinks about it—not clear what it would *mean* to "measure T," and is presumably equally unclear what it would mean to "measure $\boldsymbol{\Phi}$": the announcements of detectors are spectrally based, and in the absence of an operator there is no spectrum. For a recent contribution to

---

[43] Physics **1**, 49 (1964). J. S. Bell published his "On the Einstein Podolsky Rosen paradox" on pages 195-200 in the same volume of that now-defunct journal.

to the quantum theory of time operators, see C. M Bender & M. Gianfreda, "Matrix representation of the time operator.[44] Michael Nieto[45] has provided a useful (if amusingly informal) summary of the substance and recent history (through 1992) of work on the quantum phase problem.

One final remark before we take leave of this topic: It is interesting that the classical theory of angle variables (London's "Winklevariablen") is so much more tractable than its quantum counterpart. Look, for example,[37] to the 3-component of angular momentum

$$L_3 = x_1 p_2 - x_2 p_1$$

Introduce angles

$$\begin{aligned}\alpha_3 &= \arctan(x_1/x_2)\\ \beta_3 &= \arctan(p_1/p_2)\end{aligned}$$

and verify by Poisson bracket evaluation that α_3 and β_3 are both conjugate to L_3: $[L_3, \alpha_3] = [L_3, \beta_3] = 1$. Therefore $\phi_3 = \alpha_3 - \beta_3$ commutes with L_3:

$$[L_3, \phi_3] = 0$$

Drawing upon the identity

$$\arctan(a) - \arctan(b) = \arctan\Big(\frac{a-b}{1+ab}\Big)$$

we find

$$\phi_3 = \arctan\Big(\frac{x_1 p_2 - x_2 p_1}{x_1 p_1 + x_2 p_2}\Big)$$

We observe finally that

$$[\alpha_3, \beta_3] = \frac{x_1 p_1 + x_2 p_2}{(x_1^2 + x_2^2)(p_1^2 + p_2^2)}$$

[44] http://arxiv.org/abs/1201.3838. This 13-page paper is dated 18 Jan 2012. Google reports the existence of quite an extensive time operator literature.

[45] "Quantum phase & quantum phase operators: some physics and some history," http://arxiv.org/abs/hep-th/9304036 (8 Apr 1993). Nieto reports that Susskind & Glogower's discovery of the defect in Dirac's argument was anticipated by F. London, "Winkelvariable und kanonische Transformationen in der Undulationsmechanik," Zeitscrhift für Physik 40,193 (1927). Here again, Google reports the existence of an extensive literature.