2

WEYL TRANSFORM
and the
PHASE SPACE FORMALISM

Introduction. It would be consistent with the historical facts to assert that "quantum mechanics is a child of the Hamiltonian formulation of classical mechanics." The original version of the theory proceeded, after all, from a statement

$$\oint p\,dx = nh \quad : \quad \text{Planck-Bohr-Sommerfeld quantization condition}$$

the intent of which was to identify "quantum mechanically allowed" trajectories on classical phase space. Heisenberg's uncertainty principle

$$\Delta x \cdot \Delta p \geqslant \tfrac{1}{2}\hbar$$

refers to a pair of variables which spring as twins from Hamiltonian mechanics. And when Schrödinger wrote

$$\mathsf{H}\psi = i\hbar\partial_t\psi \quad \text{with} \quad \mathsf{H} \equiv H(x,p)\Big|_{p\rightarrow(\hbar/i)\partial_x}$$

he assumed one to be already in possession of the Hamiltonian $H(x,p)$ of the classical system which one proposed to "quantize." True, Planck's "quantum of action" refers to a concept $S = \int L\,dt$ borrowed from Lagrangian mechanics (from which Hamiltonian mechanics itself descended), but for several decades—until brought into the sunshine in the late 1940's by Feynman and Schwinger—Lagrangian notions lived only in dark shadows of quantum theory.

Hamiltonian mechanics, in its simplest form, contemplates the motion of points $\{\boldsymbol{x}(t),\boldsymbol{p}(t)\}$ in $2n$-dimensional phase space, while Schrödinger looked to the motion of a "point" $\psi(\boldsymbol{x},t)$ in the ∞-dimensional space of nice functions defined on n-dimensional configuration space. The formal disjunction is fairly profound. If one has interest in the nest of problems which live at the classical/quantum interface . . . in the *comparative design of, and relationship between* the two theories . . . then it behooves one to try to minimize the element of formal dissimilarity, to get the respective theories "into the same room together." To that end . . .

One might look, on the classical side, to Hamilton-Jacobi theory, where the object is to develop properties of the solutions $S(x,t)$ of

$$H(x,S_x) + S_t = 0$$

The classical/quantum bridge is established by a relation of the form

$$\psi(x,t) = e^{\frac{i}{\hbar}S(x,t)}$$

One holds then, in this hand and that, a pair of partial differential equations, in fields S/ψ which range on the same set of independent variables. Good physics, in rich variety, results when one rubs one against the other . . . but it is not the physics that concerns me here.

Hamiltonian physics invites one to look not to the motion of individual state points $\{x,p\}$ but to *sprinkle many state points onto phase space, and watch the motion of the population* (which is to say: watch the motion of the underlying phase fluid)—to watch not the leaf but the lake, as revealed by many floating leaves. That wholistic view is more abstract, but moderate abstraction is a small price to pay to gain access to the conceptual apparatus of statistical mechanics, chaos theory . . . and, as will emerge, quantum mechanics. How is it accomplished? Let $\{x(t;x_0,p_0),\, p(t;x_0,p_0)\}$ describe the present position of the state point which at $t=0$ resided at $\{x_0,p_0\}$; suppose, in other words, that

$$\begin{matrix} x(0;x_0,p_0) = x_0 \\ p(0;x_0,p_0) = p_0 \end{matrix} \quad \text{and} \quad \begin{matrix} \dot{x} = +\partial H/\partial p \equiv H_p \\ p = -\partial H/\partial x \equiv H_x \end{matrix}$$

We might write $\delta(x-x(t;x_0,p_0))\cdot\delta(p-p(t;x_0,p_0))$ to describe the moving state point as a moving "spike distribution." To describe the dynamical evolution of an *arbitrary* initial distribution $P(x_0,p_0;0)$ we evidently have

$$P(x,p;t) = \iint \delta(x-x(t;x_0,p_0))\cdot\delta(p-p(t;x_0,p_0))P(x_0,p_0;0)\,dx_0dp_0 \tag{1}$$

Immediately

$$\iint P(x,p;t)\,dxdp = \iint P(x_0,p_0;0)\,dx_0dp_0 = 1 \tag{2}$$

which expresses classical "conservation of probability." The rate at which the value of $P(x,p;t)$ is seen to change by an observer riding with the flow is

$$\tfrac{d}{dt}P = P_x\dot{x} + P_p\dot{p} + P_t = \{x\partial_x + p\partial_p + \partial_t\}\cdot\ \text{right side of (1)} = 0 \tag{3.1}$$

so we have

$$\tfrac{\partial}{\partial t} P = [H, P] \tag{3.2}$$

Note the subtle distinction between the result just obtained and the equation

$$\tfrac{d}{dt} A = A_x \dot{x} + A_p \dot{p} = -[H, A] \tag{4}$$

which describes the rate at which a co-flowing observer sees the value of an observable $A(x, p)$ to change.[1] The "Liouville equation" (3) can be interpreted to be an expression of the "phase flow is isovolumetric" (more picturesquely: "phase fluid is incompressible").[2]

The Liouville equation—which when written out in detail reads

$$\frac{\partial P}{\partial t} = \sum_{k=1}^{n} \left\{ \frac{\partial H}{\partial x^k}\frac{\partial P}{\partial p_k} - \frac{\partial P}{\partial x^k}\frac{\partial H}{\partial p_k} \right\} \tag{5.1}$$

and has obviously the form (not of coupled non-linear ordinary differential equations but) of a linear partial differential equation—presents Hamiltonian dynamics in an elegantly compact, statistically predisposed nutshell. Let us, in fact, assign to $P(x, p; t)$ the attributes of a probability distribution:

$$P(x, p; t) \geqslant 0 \quad \text{and} \quad \iint P(x, p; t)\, dxdp = 1 \tag{5.2}$$

It then makes sense to speak of the "expected value" of the observable $A(x, p)$:

$$\langle A \rangle = \iint A(x, p) P(x, p; t)\, dxdp \tag{5.3}$$

Equations (5) are structurally reminiscent of these equations fundamental to

---

[1] Set $A(x, p) = x$ else p to recover the canonical equations of motion. Observe also that from $[H, A] = 0$ it follows that A is conserved.

[2] That, in turn, is an expression of the proposition that the H-generated flow map is symplectic (therefore necessarily unimodular) but can be understood quite simply as follows: Look to

$$\begin{aligned} x &\mapsto \mathsf{x} = x + \tau H_p + \cdots \\ p &\mapsto \mathsf{p} = p - \tau H_x + \cdots \end{aligned}$$

Expansion of the Jacobian

$$J = \begin{vmatrix} \mathsf{x}_x & \mathsf{x}_p \\ \mathsf{p}_x & \mathsf{p}_p \end{vmatrix} = \begin{vmatrix} 1 + \tau H_{px} + \cdots & \tau H_{pp} + \cdots \\ -\tau H_{xx} + \cdots & 1 - \tau H_{xp} + \cdots \end{vmatrix} = 1 + 0\tau + \cdots$$

gives $\tfrac{d}{dt} J = 0$.

quantum mechanics:

$$i\hbar\tfrac{\partial}{\partial t}\boldsymbol{\rho} = [\,\mathsf{H},\boldsymbol{\rho}\,] \tag{6.1}$$

$$\mathrm{tr}\boldsymbol{\rho} = 1 \tag{6.2}$$

$$\langle\,\mathsf{A}\,\rangle = \mathrm{tr}\,\mathsf{A}\boldsymbol{\rho} \tag{6.3}$$

An objective and accomplishment of the "phase space formulation of quantum mechanics" is to render the association (5) ↔ (6) clear and explicit. The bridge, as will emerge, is provided by the fundamental linkage

$$[x,p] = 1 \quad\longleftrightarrow\quad [\,\mathsf{x},\mathsf{p}\,] = i\hbar\,\mathsf{I}$$

as reflected in properties of the "Weyl transform."

Introduction to the Weyl transform. Assume the classical observable $A(x,p)$ to be Fourier transformable

$$A(x,p) = \iint a(\alpha,\beta)\,e^{\frac{i}{\hbar}(\alpha p+\beta x)}\,d\alpha d\beta$$

and, with the aid of $a(\alpha,\beta)$, construct the operator-valued *Weyl transform*

$$\mathsf{A} = \iint a(\alpha,\beta)\,e^{\frac{i}{\hbar}(\alpha\mathsf{p}+\beta\mathsf{x})}\,d\alpha d\beta \tag{7}$$

of $A(x,p)$. Note that

$$A(x,p)\text{ real} \quad\Longleftrightarrow\quad \mathsf{A}\text{ self-adjoint} \tag{8}$$

since both conditions entail $a^*(\alpha,\beta) = a(-\alpha,-\beta)$. So we have in (7) an explicit **rule of correspondence** (see again (0–12)), a rule for associating quantum observables with their classical counterparts.

Many of the attractive properties of the "Weyl correspondence" reflect properties of the operators

$$\mathsf{E}(\alpha,\beta) \equiv e^{\frac{i}{\hbar}(\alpha\mathsf{p}+\beta\mathsf{x})}$$

which I now summarize. Note first that

$$\mathsf{E}(\alpha,\beta)\text{ is unitary: }\ \mathsf{E}^{-1}(\alpha,\beta) = \mathsf{E}^{+}(-\alpha,-\beta) = \mathsf{E}^{+}(\alpha,\beta)$$

And that the simplest fruit of Campbell-Baker-Hausdorff theory[3] supplies

$$\mathsf{E}(\alpha,\beta) = \begin{cases} e^{+\frac{1}{2}\frac{i}{\hbar}\alpha\beta}\cdot e^{\frac{i}{\hbar}\beta\mathsf{x}}e^{\frac{i}{\hbar}\alpha\mathsf{p}} & : \ \mathsf{xp}\text{-ordered form} \\ e^{-\frac{1}{2}\frac{i}{\hbar}\alpha\beta}\cdot e^{\frac{i}{\hbar}\alpha\mathsf{p}}e^{\frac{i}{\hbar}\beta\mathsf{x}} & : \ \mathsf{px}\text{-ordered form} \end{cases} \tag{9.1}$$

[3] See again (⋆73.6) in Chapter 0.

and also these more elaborate identities:

$$\mathsf{E}(\alpha',\beta')\,\mathsf{E}(\alpha'',\beta'') = e^{\frac{1}{2}\frac{i}{\hbar}(\alpha'\beta''-\beta'\alpha'')}\cdot \mathsf{E}(\alpha'+\alpha'',\beta'+\beta'') \tag{9.2}$$

$$= e^{\frac{i}{\hbar}(\alpha'\beta''-\beta'\alpha'')}\cdot \mathsf{E}(\alpha'',\beta'')\,\mathsf{E}(\alpha',\beta') \tag{9.3}$$

$$\mathsf{E}^{-1}(\alpha'',\beta'')\,\mathsf{E}(\alpha',\beta')\,\mathsf{E}(\alpha'',\beta'') = e^{\frac{i}{\hbar}(\alpha'\beta''-\beta'\alpha'')}\cdot \mathsf{E}(\alpha',\beta') \tag{9.4}$$

Much hinges on the circumstance that[4]

$$\mathrm{tr}\,\mathsf{E}(\alpha,\beta) = h\delta(\alpha)\delta(\beta) \tag{10}$$

from which it follows in particular that the operators $\mathsf{E}(\alpha,\beta)$ and $\mathsf{E}(\alpha',\beta')$ are *tracewise orthogonal* in the sense that

$$\mathrm{tr}\big\{\mathsf{E}(\alpha',\beta')\,\mathsf{E}^+(\alpha,\beta)\big\} = h\delta(\alpha'-\alpha)\delta(\beta'-\beta) \tag{11}$$

It follows that if—as at (7)— A is presented in the form

$$\mathsf{A} = \iint a(\alpha',\beta')\,\mathsf{E}(\alpha',\beta')\,d\alpha' d\beta'$$

then

$$a(\alpha,\beta) = \tfrac{1}{h}\,\mathrm{tr}\big\{\mathsf{A}\,\mathsf{E}^+(\alpha,\beta)\big\} \tag{12}$$

We are, in other words, in possession now of an *operator analog of the Fourier integral theorem*

$$\mathsf{A} = \iint \Big\{\tfrac{1}{h}\,\mathrm{tr}\big[\mathsf{A}\,\mathsf{E}^+(\alpha,\beta)\big]\Big\}\,\mathsf{E}(\alpha,\beta)\,d\alpha d\beta \quad : \quad \text{all } \mathsf{A} \tag{13}$$

---

[4] Work from this variant of (9.1):

$$\mathsf{E}(\alpha,\beta) = e^{\frac{1}{2}\frac{i}{\hbar}\alpha\mathsf{p}}\,e^{\frac{i}{\hbar}\beta\mathsf{x}}\,e^{\frac{1}{2}\frac{i}{\hbar}\alpha\mathsf{p}}$$

Pass, as a matter of momentary convenience into the x-representation, writing

$$\mathrm{tr}\,\mathsf{E}(\alpha,\beta) = \sum_n \int \psi_n^*(x) e^{\frac{1}{2}\alpha\frac{d}{dx}}\,e^{\frac{i}{\hbar}\beta x} e^{\frac{1}{2}\alpha\frac{d}{dx}}\psi_n(x)\,dx$$

where $\psi_n(x) = (x|n)$, and $\big\{|n)\big\}$ is complete orthonormal. Then

$$\begin{aligned}
&= \sum_n \int \psi_n^*(x) e^{\frac{1}{2}\alpha\frac{d}{dx}}\,e^{\frac{i}{\hbar}\beta x}\psi_n(x+\tfrac{1}{2}\alpha)\,dx\\
&= \sum_n \int \psi_n^*(x) e^{\frac{i}{\hbar}\beta(x+\frac{1}{2}\alpha)}\psi_n(x+\alpha)\,dx\\
&= \int \Big\{\sum_n \psi_n^*(\xi-\tfrac{1}{2}\alpha)\psi_n(\xi+\tfrac{1}{2}\alpha)\Big\} e^{\frac{i}{\hbar}\beta\xi}\,d\xi\\
&= \delta(\alpha)\cdot h\delta(\beta)
\end{aligned}$$

where in the final step we have used completeness $\sum \psi_n(x)\psi_n^*(y) = \delta(x-y)$ and Fourier's $\int e^{\frac{i}{\hbar}\beta\xi}\,d\xi = h\delta(\beta)$.

and in position to do Fourier analysis on operators. The Weyl correspondence is seen in this light to arise from an identification of the form

$$\begin{array}{r}\text{Fourier analysis on functions } A(x,p)\\ \updownarrow\\ \text{Fourier analysis on operators } \mathbf{A}\end{array}$$

An operator ordering calculus. Returning to (7) with (9.1) we have

$$\begin{aligned}\mathbf{A} &= \iint a(\alpha,\beta)\, e^{+\frac{1}{2}\frac{i}{\hbar}\alpha\beta}\cdot e^{\frac{i}{\hbar}\beta\mathbf{x}}e^{\frac{i}{\hbar}\alpha\mathbf{p}}\,d\alpha d\beta\\ &= {}_{\mathbf{x}}\Big[\iint a(\alpha,\beta)\, e^{+\frac{1}{2}\frac{i}{\hbar}\alpha\beta}\cdot e^{\frac{i}{\hbar}\beta x}e^{\frac{i}{\hbar}\alpha p}\,d\alpha d\beta\Big]_{\mathbf{p}}\\ &= {}_{\mathbf{x}}\Big[A_{xp}(x,p)\Big]_{\mathbf{p}} \text{ with } A_{xp}(x,p) = \exp\Big\{+\tfrac{1}{2}\tfrac{\hbar}{i}\tfrac{\partial^2}{\partial x\partial p}\Big\}A(x,p) \qquad (14.1)\end{aligned}$$

and by a similar argument

$$= {}_{\mathbf{p}}\Big[A_{px}(x,p)\Big]_{\mathbf{x}} \text{ with } A_{px}(x,p) = \exp\Big\{-\tfrac{1}{2}\tfrac{\hbar}{i}\tfrac{\partial^2}{\partial x\partial p}\Big\}A(x,p) \qquad (14.2)$$

The implication is that we can proceed

$$A(x,p) \xrightarrow[\text{Weyl}]{\hspace{3cm}} \mathbf{A}$$

by first constructing $A_{xp}(x,p)$ else $A_{px}(x,p)$ and then making the appropriate ordered substitutions $x \to \mathbf{x}$, $p \to \mathbf{p}$. The procedure requires that we compute no Fourier transform, and can be performed *even when $A(x,p)$ does not possess a Fourier transform*, so serves to extend the reach of the Weyl formalism. Look, for example, to the polynomial

$$A(x,p) \equiv xp \Rightarrow \begin{cases} A_{xp} = xp + \frac{1}{2}\frac{\hbar}{i} \Rightarrow \mathbf{A} = \mathbf{x}\,\mathbf{p} + \frac{1}{2}\frac{\hbar}{i}\mathbf{I}\\ A_{px} = xp - \frac{1}{2}\frac{\hbar}{i} \Rightarrow \mathbf{A} = \mathbf{p}\,\mathbf{x} - \frac{1}{2}\frac{\hbar}{i}\mathbf{I}\end{cases}$$

The alternative descriptions of $\mathbf{A}$ are equivalent by $[\mathbf{x},\mathbf{p}] = i\hbar\mathbf{I}$, and when taken in combination give a result which can be expressed

$$xp \xrightarrow[\text{Weyl}]{\hspace{3cm}} \tfrac{1}{2}(\mathbf{x}\,\mathbf{p} + \mathbf{p}\,\mathbf{x})$$

Generally, (14) supplies

$$A_{xp}(x,p) = A(x,p) + \text{power series in } \hbar$$

but in the example a symmetrization technique has made it possible to eliminate all exposed $\hbar$'s from the description of $\mathbf{A}$; such an objective can, of course, always be accomplished by substitutions $\hbar \mapsto -i(\mathbf{x}\,\mathbf{p} - \mathbf{p}\,\mathbf{x})$.

Equations (14) can also be used to efficiently *reverse the ordering* of ordered expressions, the essential point being that

$$A_{xp}(x,p) = \exp\Big\{+\tfrac{\hbar}{i}\tfrac{\partial^2}{\partial x\partial p}\Big\}A_{px}(x,p) \tag{15}$$

Looking back, by way of illustration, to our recent example, we find

$$\exp\Big\{+\tfrac{\hbar}{i}\tfrac{\partial^2}{\partial x\partial p}\Big\}\big(xp - \tfrac{1}{2}\tfrac{\hbar}{i}\big) = \big(xp + \tfrac{1}{2}\tfrac{\hbar}{i}\big)$$

as required.

Operator products & commutators. Proceeding from

$$\begin{aligned} A(x,p) &\xrightarrow[\text{Weyl}]{} \mathsf{A} = \iint a(\alpha',\beta')\,\mathsf{E}(\alpha',\beta')\,d\alpha'\,d\beta' \\ B(x,p) &\xrightarrow[\text{Weyl}]{} \mathsf{B} = \iint b(\alpha'',\beta'')\,\mathsf{E}(\alpha'',\beta'')\,d\alpha''d\beta'' \end{aligned}$$

we ask: What is the Weyl transform of AB? Straightforward calculation[5] supplies

$$\mathsf{AB} = \iiiint a(\alpha',\beta')b(\alpha'',\beta'')e^{\frac{1}{2}\frac{i}{\hbar}(\alpha'\beta''-\beta'\alpha'')}\,\mathsf{E}(\alpha'+\alpha'',\beta'+\beta'')\,d\alpha'd\beta'd\alpha''d\beta''$$

which is evidently the Weyl transform of

$$\begin{aligned} &\iiiint e^{\frac{1}{2}\frac{i}{\hbar}(\alpha'\beta''-\beta'\alpha'')}\cdot a(\alpha',\beta')b(\alpha'',\beta'')e^{\frac{i}{\hbar}[(\alpha'+\alpha'')p+(\beta'+\beta'')x]}d\alpha'd\beta'd\alpha''d\beta'' \\ &\qquad = \exp\Big\{\tfrac{1}{2}\tfrac{\hbar}{i}\Big[\big(\tfrac{\partial}{\partial p}\big)_{\!A}\big(\tfrac{\partial}{\partial x}\big)_{\!B} - \big(\tfrac{\partial}{\partial x}\big)_{\!A}\big(\tfrac{\partial}{\partial p}\big)_{\!B}\Big]\Big\}A(x,p)B(x,p) \qquad (16) \\ &\qquad = A(x,p)B(x,p) + \text{power series in }\hbar \end{aligned}$$

Since on the one hand

$$\mathsf{AB} = \tfrac{1}{2}(\mathsf{AB}+\mathsf{BA}) + \tfrac{1}{2}(\mathsf{AB}-\mathsf{BA})$$

while on the other

$$\begin{aligned} \exp\Big\{\tfrac{1}{2}\tfrac{\hbar}{i}\Big[\big(\tfrac{\partial}{\partial p}\big)_{\!A}\big(\tfrac{\partial}{\partial x}\big)_{\!B} - \big(\tfrac{\partial}{\partial x}\big)_{\!A}\big(\tfrac{\partial}{\partial p}\big)_{\!B}\Big]\Big\} &= \exp\Big\{i\tfrac{\hbar}{2}\Big[\big(\tfrac{\partial}{\partial x}\big)_{\!A}\big(\tfrac{\partial}{\partial p}\big)_{\!B} - \big(\tfrac{\partial}{\partial x}\big)_{\!B}\big(\tfrac{\partial}{\partial p}\big)_{\!A}\Big]\Big\} \\ &= \cos\Big\{\tfrac{\hbar}{2}[\text{etc.}]\Big\} + i\sin\Big\{\tfrac{\hbar}{2}[\text{etc.}]\Big\} \end{aligned}$$

we are brought to the pretty conclusion that

---

[5] See QUANTUM MECHANICS (1967), Chapter 2, p. 27 for the tedious details.

$$\cos\left\{\tfrac{\hbar}{2}\left[\left(\tfrac{\partial}{\partial x}\right)_A\left(\tfrac{\partial}{\partial p}\right)_B - \left(\tfrac{\partial}{\partial x}\right)_B\left(\tfrac{\partial}{\partial p}\right)_A\right]\right\}AB \xrightarrow[\text{Weyl}]{} \tfrac{1}{2}(\mathsf{AB}+\mathsf{BA}) \tag{17.1}$$

$$i\sin\left\{\tfrac{\hbar}{2}\left[\left(\tfrac{\partial}{\partial x}\right)_A\left(\tfrac{\partial}{\partial p}\right)_B - \left(\tfrac{\partial}{\partial x}\right)_B\left(\tfrac{\partial}{\partial p}\right)_A\right]\right\}AB \xrightarrow[\text{Weyl}]{} \tfrac{1}{2}(\mathsf{AB}-\mathsf{BA}) \tag{17.2}$$

The latter of the preceding equations sets up this relation between quantum mechanical commutators and classical Poisson brackets:

$$\text{commutator } [\mathsf{A},\mathsf{B}] \xrightarrow[\text{Weyl}]{} i\hbar\Big\{\text{Poisson bracket } [A,B] + \text{terms of order } \hbar^2\Big\}$$

Which is satisfying, yet inconsistent with Dirac's stipulation[6] that

$$[\mathsf{A},\mathsf{B}] \longleftrightarrow i\hbar\,[A,B]$$

should be precise (no correction terms). From (17.1) we obtain

$$\tfrac{1}{2}(\mathsf{AB}+\mathsf{BA}) \xrightarrow[\text{Weyl}]{} AB - \tfrac{1}{2!}\left(\tfrac{\hbar}{2}\right)^2\Big\{A_{xx}B_{pp} - 2A_{xp}B_{px} + A_{pp}B_{xx}\Big\} + \cdots$$

which in the case $B = A$ yields a result

$$\mathsf{A}^2 \xrightarrow[\text{Weyl}]{} A^2 - \left(\tfrac{\hbar}{2}\right)^2\Big\{A_{xx}A_{pp} - A_{xp}A_{px}\Big\} + \cdots \neq A^2$$

which is inconsistent with von Neumann's stipulation[7] that

$$A \longrightarrow \mathsf{A} \quad\Longrightarrow\quad f(A) \longrightarrow f(\mathsf{A})$$

But the principles advanced by Dirac and von Neumann are readily shown to be inconsistent, and both are susceptible to the criticism that (except in the simplest cases) the **A**-operator which they assign to $A(x,p)$ is *non-unique*. Many alternatives to (variants of) Weyl's procedure have been proposed,[8] but none offers distinct advantages, except in isolated circumstances. Maybe someday it will become possible to resolve the matter on observational grounds. In the meantime, I base my tentative embrace of Weyl's procedure on the fact that it leads with swift elegance to what seem to me to be some valuable insights. I am content to live in violation of Dirac's/von Neumann's (faulty, and not very deeply motivated) principles on grounds that quantum mechanics is a profoundly strange subject, entitled to its surprising quirks . . . and one cannot reasonably expect to get from the beginning to the end without encountering wrinkles.

---

[6] *Principles of Quantum Mechanics* (4th edition 1958), Chapter 4.

[7] *Mathematical Foundations of Quantum Mechanics* (1955), pp. 313 *et seq*.

[8] For a nice review of the older literature on this subject, see J. R. Shewell, "On the formation of quantum mechanical operators," AJP **27**, 16 (1959). See also "Correspondence rules *via* Feynmanism" in TRANSFORMATIONAL PHYSICS & PHYSICAL GEOMETRY (1971–1983), which contains many references to the more recent literature.

Finally, we bring (10) to the equation that gave (16) and obtain

$$
\begin{aligned}
\text{tr}\big\{\mathbf{AB}\big\} &= h\iiiint a(\alpha',\beta')b(\alpha'',\beta'')e^{\frac{1}{2}\frac{i}{\hbar}(\alpha'\beta''-\beta'\alpha'')} \\
&\qquad\qquad \cdot\,\delta(\alpha'+\alpha'')\delta(\beta'+\beta'')\,d\alpha'd\beta'd\alpha''d\beta'' \\
&= h\iint a(\alpha,\beta)b(-\alpha,-\beta)\,d\alpha d\beta \\
&= h\tfrac{1}{h^4}\int\!\!\!\int\!\!\!\int\!\!\!\int\!\!\!\int\!\!\!\int A(x',p')B(x'',p'')e^{-\frac{i}{\hbar}(\alpha p'+\beta x')} \\
&\qquad\qquad \cdot\,e^{+\frac{i}{\hbar}(\alpha p''+\beta x'')}\,d\alpha d\beta dx'dp'dx''dp'' \\
&= \tfrac{1}{h^3}\iiiint A(x',p')B(x'',p'')h\delta(p''-p')h\delta(x''-x')\,dx'dp'dx''dp'' \\
&= \tfrac{1}{h}\iint A(x,p)B(x,p)\,dxdp
\end{aligned}
\tag{18}
$$

The beauty of this result lies in the circumstance that it permits the quantum mechanical statement (6.3) to be cast in the notation of its classical statistical counterpart (5.3). For suppose

$$
\mathbf{A}\xrightarrow[\text{Weyl}]{} A(x,p)\qquad\text{and}\qquad \boldsymbol{\rho}\xrightarrow[\text{Weyl}]{} hP(x,p)
\tag{19}
$$

Then (18) can be used to write

$$
\langle\mathbf{A}\rangle = \text{tr}\,\mathbf{A}\boldsymbol{\rho} = \iint A(x,p)P(x,p)\,dxdp
\tag{20}
$$

The Wigner distribution. Let us, for the moment, suppose that $\boldsymbol{\rho}$ refers to a pure state: $\boldsymbol{\rho} = |\psi)(\psi|$. Working from a slight variant of (13), we have

$$
\begin{aligned}
|\psi)(\psi| &= \tfrac{1}{h}\iint(\psi|\,\mathbf{E}(\alpha,\beta)|\psi)\,e^{-\frac{i}{\hbar}(\alpha\mathbf{p}+\beta\mathbf{x})}\,d\alpha d\beta \\
&\Big\downarrow \text{Weyl} \\
hP_\psi(x,p) &= \tfrac{1}{h}\iint(\psi|\,\mathbf{E}(\alpha,\beta)|\psi)\,e^{-\frac{i}{\hbar}(\alpha p+\beta x)}\,d\alpha d\beta
\end{aligned}
\tag{21}
$$

Pass to the $\mathbf{x}$-representation and argue as we did[4] in the derivation of (10), to obtain

$$
\begin{aligned}
(\psi|\,\mathbf{E}(\alpha,\beta)|\psi) &= \int\psi^*(y)\,e^{\frac{i}{\hbar}\beta(y+\frac{1}{2}\alpha)}\psi(y+\alpha)\,dy \\
&= \int\psi^*(z-\tfrac{1}{2}\alpha)\,e^{\frac{i}{\hbar}\beta z}\psi(z+\tfrac{1}{2}\alpha)\,dz
\end{aligned}
$$

whence

$$
\begin{aligned}
hP_\psi(x,p) &= \tfrac{1}{h}\iiint\psi^*(z-\tfrac{1}{2}\alpha)\,e^{-\frac{i}{\hbar}\alpha p}\,\psi(z+\tfrac{1}{2}\alpha)\,e^{\frac{i}{\hbar}\beta(z-x)}\,d\beta dz d\alpha \\
&= \tfrac{1}{h}\iiint\psi^*(z-\tfrac{1}{2}\alpha)\,e^{-\frac{i}{\hbar}\alpha p}\,\psi(z+\tfrac{1}{2}\alpha)\,h\delta(z-x)\,dz d\alpha \\
&= \int\psi^*(x-\tfrac{1}{2}\alpha)\,e^{-\frac{i}{\hbar}\alpha p}\,\psi(x+\tfrac{1}{2}\alpha)\,d\alpha
\end{aligned}
$$

which by simple change of variable $\alpha \mapsto \xi = -\frac{1}{2}\alpha$ becomes

$$P_\psi(x,p) = \tfrac{2}{h}\int \psi^*(x+\xi)\, e^{2\frac{i}{\hbar}p\,\xi}\, \psi(x-\xi)\, d\xi \tag{22.1}$$

Had we elected to work in the $\mathbf{p}$-representation we would have been led by the same argument to

$$= \tfrac{2}{h}\int \Psi^*(p-\zeta)\, e^{2\frac{i}{\hbar}x\,\zeta}\, \Psi(p+\zeta)\, d\zeta \tag{22.2}$$

where $\Psi(p) = (p|\psi)$; alternatively, we might have worked from (22.1) with the assistance of (0–82): $(x|\psi) = (1/\sqrt{h})\int e^{(i/\hbar)px}(p|\psi)\,dp$.

At (22) we have obtained the famous "Wigner distribution function," which Wigner, in the course of some early work concerned with the relation of quantum to classical statistical mechanics,[9] was content simply to pluck from his hat,[10] but which we see now is intimately associated with Weyl's procedure.[11] I turn now to an account of some of the striking general properties of Wigner's unpromising-looking construction.

The density matrix $\boldsymbol{\rho}$ is self-adjoint, so its Weyl transform is necessarily *real*, and in fact the reality of $P_\psi(x,p)$, as described at (22), is manifest.

In (20) set $\mathbf{A} = \mathbf{I}$ to obtain

$$(\psi|\psi) = \mathrm{tr}\boldsymbol{\rho} = \iint P_\psi(x,p)\, dxdp = 1 \tag{23}$$

From (22) we are led to "marginal distributions"

$$\int P_\psi(x,p)\, dp = |\psi(x)|^2 \tag{24.1}$$

$$\int P_\psi(x,p)\, dx = |\Psi(p)|^2 \tag{24.2}$$

which could not be more satisfactory, and from which (23) can be recovered as a corollary.

---

[9] "On the quantum correction for thermodynamic equilibrium," Phys. Rev. **40**, 749 (1932).

[10] Or perhaps from Leo Szilard's hat. In a footnote, Wigner reports that "This expression was found by L. Szilard and the present author some years ago for another purpose," but gives no indication of what that "other purpose" might have been, and cites no reference.

[11] Recognition of the Wigner-Weyl connection is usually attributed to J. E. Moyal, "Quantum mechanics as a statistical theory," Proc. Camb. Phil. Soc. **45**, 92 (1949), but clear anticipations of many of Moyal's results can be found in J. H. Groenwold, "On the principles of elementary quantum mechanics," Physica **12**, 405 (1946).

Concerning that most recent use of quotation marks: Equations (24) would describe literal marginal distributions if $P_\psi(x,p)$ were itself a distribution, but in fact it is not. Wigner's function enters into the equations of the phase space formalism *as though* it were a classical distribution function, but can—and typically does—assume negative values.[12] For this reason $P_\psi(x,p)$ is sometimes called a "quasi-distribution." The point is most readily established by example. The ground state and first two excited states of a harmonic oscillator can be described[13]

$$\begin{aligned}\psi_0(x) &= \tfrac{1}{\sqrt{a\sqrt{2\pi}}}\,e^{-\frac{1}{4}\varkappa^2}\\ \psi_1(x) &= \tfrac{1}{\sqrt{a\sqrt{2\pi}}}\,e^{-\frac{1}{4}\varkappa^2}\cdot\tfrac{1}{\sqrt{1!}}\,\varkappa\\ \psi_2(x) &= \tfrac{1}{\sqrt{a\sqrt{2\pi}}}\,e^{-\frac{1}{4}\varkappa^2}\cdot\tfrac{1}{\sqrt{2!}}(\varkappa^2-1)\end{aligned}$$

with $a \equiv \sqrt{\hbar/2m\omega}$. Working from (22.1) with the assistance of *Mathematica*—later we will develop analytical means to obtain such results—we find that the Wigner transforms of those wave functions can be described

$$\left.\begin{aligned}P_0(x,p) &= +\tfrac{2}{h}e^{-\frac{1}{2}\mathcal{E}}\\ P_1(x,p) &= -\tfrac{2}{h}e^{-\frac{1}{2}\mathcal{E}}(1-\mathcal{E})\\ P_2(x,p) &= +\tfrac{2}{h}e^{-\frac{1}{2}\mathcal{E}}(1-2\mathcal{E}+\tfrac{1}{2}\mathcal{E}^2)\end{aligned}\right\} \tag{25}$$

where $\mathcal{E} \equiv \varkappa^2 + \wp^2$ is "dimensionless energy," interpretable as squared radius on the $(\varkappa,\wp)$-coordinatized phase plane. The functions (25) are plotted in Figure 1, where regions of negativity are evident.

The essence of the Heisenberg uncertainty principle is very neatly conveyed by a property of $P_\psi(x,p)$—known to Wigner, but first reported (without proof) by Takabayasi[14]—which is much easier to state

$$|P_\psi(x,p)| \leqslant \tfrac{2}{h} \tag{26}$$

---

[12] Recall my earlier remark about quantum mechanics being "a profoundly strange subject, entitled to its surprising quirks . . . "

[13] I find it notationally convenient in the present context to write

$$H = \tfrac{1}{4}\hbar\omega(\varkappa^2+\wp^2)$$

with

$$\begin{aligned}\varkappa &\equiv \sqrt{2m\omega/\hbar}\cdot x \quad &&:\quad \text{dimensionless length}\\ \wp &\equiv \sqrt{2/m\omega\hbar}\cdot p \quad &&:\quad \text{dimensionless momentum}\end{aligned}$$

See QUANTUM MECHANICS (1967), Chapter 2, p. 58 for discussion of the relation of these to some other conventions.

[14] T. Takabayasi, "The formulation of quantum mechanics in terms of ensembles in phase space," Prog. Theo. Phys. **11**, 341 (1954). See especially §7 in that important paper.

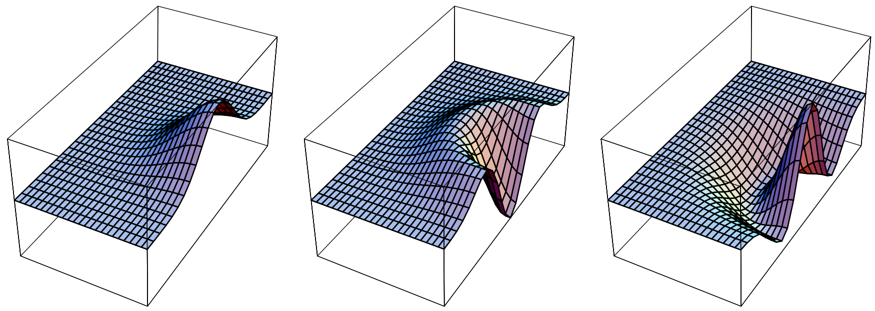

FIGURE 1: *Wigner functions (25) for the three lowest-lying energy eigenstates of a harmonic oscillator. The excited states display "regions of negativity." It is on account of the typical occurance of such regions that $P_\psi(x,p)$ is sometimes called a "quasi-distribution."*

than to prove, but the proof is highly instructive. To the observation that (21) can be written

$$\begin{aligned} P_\psi(x,p) &= \tfrac{2}{h}(\psi|\mathsf{W}(x,p)|\psi) \\ \mathsf{W}(x,p) &\equiv \tfrac{1}{2h}\iint \mathsf{E}(\alpha,\beta)e^{-\frac{i}{\hbar}(\alpha p+\beta x)}\,d\alpha d\beta \end{aligned} \tag{27}$$

bring the observations that (almost obviously)

$$\mathsf{W}^+(x,p) = \mathsf{W}(x,p) \quad : \quad \mathsf{W}(x,p) \text{ is self-adjoint}$$

and (not at all obviously, though the tedious proof is elementary[15])

$$\mathsf{W}^{-1}(x,p) = \mathsf{W}(x,p) \quad : \quad \mathsf{W}(x,p) \text{ is also unitary}$$

The operators $\mathsf{W}(x,p)$ are, in other words, "self-adjoint square roots of unity:"

$$\mathsf{W}^2(x,p) = \mathsf{I} \quad : \quad \text{all } x \text{ and } p$$

We now have

$$\begin{aligned} |\tfrac{h}{2}P_\psi|^2 &= (\psi|\Omega)(\Omega|\psi) \quad \text{with} \quad |\Omega) \equiv \mathsf{W}|\psi) \\ &\leqslant \underbrace{(\psi|\psi)(\Omega|\Omega)}_{1 \quad \text{because } (\Omega|\Omega) = (\psi|\mathsf{W}^2|\psi) = (\psi|\psi)} \quad \text{by the Schwarz inequality} \end{aligned}$$

which completes the pretty argument. The uncertainty principle arises now as an expression of the proposition (see Figure 2) that—since $P_\psi(x,p)$ lives "under a ceiling," yet is obliged to satisfy the normalization condition (23)—there exists a least-allowed area of the phase domain on which $P_\psi(x,p)$ is non-zero:

$$\text{"footprint"} \geqslant \tfrac{h}{2}$$

[15] For the details see QUANTUM MECHANICS (1967), Chapter 3, p. 127.

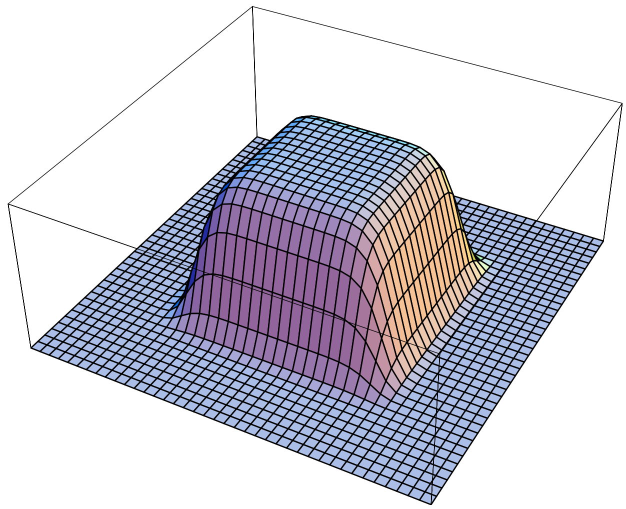

FIGURE 2: *Wigner functions $P_\psi(x,p)$, since bounded by (26) and normed by (23), possess "footprints" of area not less than $\frac{1}{2}h$, which is the upshot of the Heisenberg uncertainty principle. The figure has been adapted from my "A mathematical note: Gaussians of square cross-section," which some readers may find to be of independent interest.*

Formally, as $\hbar \downarrow 0$ the least-possible footprint becomes progressively smaller; in the classical limit—but *only* in the classical limit—does it become possible to contemplate distributions of the form

$$P(x,p) \sim \delta(x-x_0)\cdot\delta(p-p_0)$$

The "Wigner transform" (22.1) sends

$$\psi(x) \xrightarrow[\text{Wigner}]{} P_\psi(x,p)$$

It seems natural to ask: Can one, if given $P_\psi(x,p)$, recover $\psi(x)$? But my occasional attempts to resolve the matter had been fruitless, so I was quite surprised when Mark Beck, a young colleague whom I had invited to speak with my students about some applications of the phase space formalism to quantum optics, referred casually to the "inverse Wigner transform." When I asked "How is it accomplished?" he proceeded to show me. Mark does not claim to have invented the trick in question, but can cite no source, and it was, so far as I am aware, unknown to the founding fathers of the field; I call it "Beck's trick." It runs

$$\psi(x) \xleftarrow[\text{Beck}]{} P_\psi(x,p)$$

and proceeds as follows: By Fourier transformation of (22.1) obtain

$$\int P_\psi(x,p)\,e^{-2\frac{i}{\hbar}p\zeta}\,dp = \int \psi^*(x+\xi)\,\delta(\xi-\zeta)\psi(x-\xi)\,d\xi$$
$$= \psi^*(x+\zeta)\,\psi(x-\zeta)$$

Select a point a at which $\int P_\psi(a,p)\,dp = \psi^*(a)\,\psi(a) \neq 0$.[16] Set $\zeta = a - x$ to obtain

$$\int P_\psi(x,p)e^{-2\frac{i}{\hbar}p(a-x)}\,dp = \psi^*(a)\,\psi(2x-a)$$

and by notational adjustment $2x - a \mapsto x$ obtain

$$\begin{aligned}\psi(x) &= [\psi^*(a)]^{-1}\cdot\int P_\psi(\tfrac{x+a}{2},p)\,e^{\frac{i}{\hbar}p(x-a)}\,dp\\ &\downarrow\\ &= [\psi^*(0)]^{-1}\cdot\int P_\psi(\tfrac{x}{2},p)\,e^{\frac{i}{\hbar}px}\,dp \quad \text{in the special case } a=0 \end{aligned}\tag{28}$$

where the prefactor is, in effect, a normalization constant, fixed to within an arbitrary phase factor.

To test the efficacy of (28) we look to the oscillator ground state, for which at (25) we obtained

$$P_0(x,p) = \tfrac{2}{\hbar}\exp\Big\{-\tfrac{m\omega}{\hbar}x^2 - \tfrac{1}{m\omega\hbar}p^2\Big\}\tag{29}$$

Mathematica, working from (28), supplies

$$\psi_0^*(0)\cdot\psi_0(x) = \big(\tfrac{m\omega}{\pi\hbar}\big)^{\frac{1}{2}}\exp\Big\{-\tfrac{m\omega}{2\hbar}x^2\Big\}$$

which—since therefore $\psi_0^*(0) = \big(\frac{m\omega}{\pi\hbar}\big)^{\frac{1}{4}}e^{i(\text{arbitrary phase})}$—is exactly correct.

The distribution $P_0(x,p)$ introduced at (29) is of the class

$$P(x,p;a,b) \equiv \tfrac{1}{\pi ab}\exp\big\{-(x/a)^2-(p/b)^2\big\}\tag{30}$$

and belongs more particularly to the subclass $ab = \hbar$. The bivariate Gaussians (30) conform to the normalization condition (23) in all cases, but conform to the boundedness condition (26) if and only if $ab \geqslant \hbar$. Beck's trick, as implemented by *Mathematica*, supplies

$$\psi(x;a,b) = \frac{1}{\sqrt{a\sqrt{\pi}}}\exp\Big\{-\tfrac{1}{4}\big(\tfrac{1}{a^2}+\tfrac{b^2}{\hbar^2}\big)x^2\Big\}$$

in all cases. What goes wrong when the boundedness condition is violated; i.e., when (30) describes an "impossible" Wigner function? The question is answered by the observation that

$$\int |\psi(x;a,b)|^2\,dx = 1 \quad \text{if and } \underline{\text{only}} \text{ if } ab = \hbar$$

[16] Such a point is, by $\int \psi^*(x)\,\psi(x)\,dx = 1$, certain to exist. It is often most convenient (but not always possible) to—with Beck—set $a = 0$.

As he approached the end of his working career, Wigner was several times tempted by circumstance to revisit this site of his youthful invention.[17] In his contribution to a collection of essays in honor of Alfred Landé[18] he develops a short list of conditions which are sufficient to insure that a given $P(x,p)$ can be displayed as an instance of (22.1), but professes dissatisfaction with his final condition, which was that in the absence of forces $P(x,p;t)$ should satisfy the classical equation $\partial_t P = -(p/m)\partial_x P$. A decade later he was able to replace that condition with one that he found more satisfactory.[19] I describe the new condition, as it arises from the theory already in hand. Enlarging upon (21), let us write

$$\begin{aligned}\boldsymbol{\rho}_\psi \equiv |\psi)(\psi| &\xrightarrow[\text{Weyl}]{} hP_\psi(x,p)\\ \boldsymbol{\rho}_\varphi \equiv |\varphi)(\varphi| &\xrightarrow[\text{Weyl}]{} hP_\varphi(x,p)\end{aligned}$$

Then

$$\begin{aligned}|(\psi|\varphi)|^2 &= (\psi|\varphi)(\varphi|\psi)\\ &= \text{tr}\big\{\boldsymbol{\rho}_\psi\boldsymbol{\rho}_\varphi\big\}\\ &= h\iint P_\psi(x,p)P_\varphi(x,p)\,dxdp \quad \text{by (18)}\\ &\leqslant (\psi|\psi)(\varphi|\varphi) = 1\end{aligned} \tag{31}$$

which is the condition embraced by O'Connell & Wigner. As a special case, one has (compare (23))

$$h\iint P_\psi(x,p)P_\psi(x,p)\,dxdp = \text{tr}\big\{\boldsymbol{\rho}_\psi^2 = \boldsymbol{\rho}_\psi\big\} = 1 \tag{32}$$

which, at least from a function-theoretic point of view, is remarkable.

[17] Wigner was thirty when the 1932 paper[9] was published. He was awarded the Nobel Prize in 1963 for "systematically improving and extending the methods of quantum mechanics . . . "

[18] "Quantum-mechanical distribution functions revisited" in *Perspectives in Quantum Theory*, edited by W. Yourgrau & A. van der Merwe (1971). The paper presents a good list of the references that Wigner considered significant.

[19] R. F. O'Connell & E. P. Wigner, "Quantum-mechanical distribution functions: conditions for uniqueness," Physics Letters **83A**, 145 (1981). Shortly later the same two authors published "Some properties of a non-negative quantum-mechanical distribution function," Physics Letters **85A**, 121 (1981), which will concern us later. And a comprehensive review of the entire subject is presented in M. Hillery, R. F. O'Connell, M. O. Scully & E. P. Wigner, "Distribution functions in physics: fundamentals," Physics Reports **106**, 121 (1984). The Weyl–Wigner connection does receive mention in the last of those papers, but seems otherwise not to have interested Wigner.

The Wigner representation of mixed states. Let $\boldsymbol{\rho}$ refer to a mixture, and write

$$\boldsymbol{\rho} = \sum_n \rho_n \boldsymbol{\rho}_n \quad \text{with} \quad \boldsymbol{\rho}_n \equiv |n)(n|$$

to describe its spectral representation. From the linearity of the Weyl transform it follows that if

$$\boldsymbol{\rho}_n \xrightarrow[\text{Weyl}]{} hP_n(x,p)$$

then

$$\boldsymbol{\rho} \xrightarrow[\text{Weyl}]{} hP(x,p) = h\sum_n \rho_n P_n(x,p) \tag{33}$$

Familiar arguments (or slight variations of them) serve to establish that

- $P(x,p)$ is real-valued, whether it refers to a pure state or a mixture;
- $\int P(x,p)dxdp = 1$, whether ... a pure state or a mixture;
- $|P(x,p)| \leqslant \frac{2}{h}$, whether ... a pure state or a mixture.

But we saw at (0–113) that tr$\boldsymbol{\rho}^2 \leqslant$ tr$\boldsymbol{\rho}$, with equality if and only if the state is pure; the implication, by (18), is that

$$h\iint P^2(x,p)\,dxdp \leqslant 1 \tag{34}$$

with equality (see again (32)) if and only if $P(x,p)$ refers to a pure state.

The projective operators $\boldsymbol{\rho}_n$ are tracewise orthogonal

$$\text{tr}\big\{\boldsymbol{\rho}_m \boldsymbol{\rho}_n\big\} = \delta_{mn}$$

and—if known—permit one to write

$$\rho_n = \text{tr}\big\{\boldsymbol{\rho}\boldsymbol{\rho}_n\big\}$$

One can announce that $\boldsymbol{\rho}$ is a density operator (describes the state of some mixture) if it is found that $0 \leqslant \rho_n \leqslant 1$ (all n) and that $\sum_n \rho_n = 1$. In Wigner language we therefore have

$$\begin{aligned} h\iint P_m(x,p)P_n(x,p)\,dxdp &= \delta_{mn} \\ \rho_n = h\iint P(x,p)P_n(x,p)\,dxdp & \end{aligned} \tag{35}$$

and can announce under those same spectral conditions that $P(x,p)$ is a Wigner function. But I know of no way short of such full-blown spectral analysis to distinguish Wigner functions from other functions of the same arguments. In particular, we presently possess no direct way to establish that

$$\text{No function of the form } f(x)\cdot g(p) \text{ can be a Wigner function} \tag{36}$$

though we know on other grounds that that surely must be the case.

Quantum motion of a Wigner distribution. We elect to work in the Schrödinger picture, where observables (at least those with time-independent definitions) remain at rest and the state descriptor $\boldsymbol{\rho}$ moves, as described by

$$i\hbar\partial_t\boldsymbol{\rho} = [\,\mathbf{H}\,,\boldsymbol{\rho}\,] \qquad (6.1) = (37)$$

Drawing upon

$$\mathbf{H} \xrightarrow[\text{Weyl}]{} H(x,p) \quad : \quad \begin{cases}\text{the classical Hamiltonian; inverse of the}\\ \text{transformation that originally gave us } \mathbf{H}\end{cases}$$

$$\boldsymbol{\rho} \xrightarrow[\text{Weyl}]{} hP(x,p)$$

and the rule (17.2) for transforming commutators, we have

$$\frac{\partial}{\partial t}P(x,p;t) = \frac{2}{\hbar}\sin\Big\{\frac{\hbar}{2}\Big[\big(\tfrac{\partial}{\partial x}\big)_H\big(\tfrac{\partial}{\partial p}\big)_P - \big(\tfrac{\partial}{\partial x}\big)_P\big(\tfrac{\partial}{\partial p}\big)_H\Big]\Big\}H(x,p)P(x,p;t) \tag{38}$$

$$= [H,P] + \text{terms of order } \hbar^2 \tag{39}$$

↑ Poisson bracket

which, in view of the fact that (39) → (3.2) as $\hbar \downarrow 0$, serves beautifully our motivating objective: at (39) quantum/classical dynamics have at last and quite explicitly been brought "into the same room together."

Spelling out in more detail the meaning of (38), we have

$$\begin{aligned}P_t &= \big\{H_xP_p - H_pP_x\big\}\\ &\quad - \tfrac{1}{3!}\big(\tfrac{\hbar}{2}\big)^2\big\{H_{xxx}P_{ppp} - 3H_{xxp}P_{ppx} + 3H_{xpp}P_{pxx} - H_{ppp}P_{xxx}\big\} + \cdots\end{aligned}$$

which for systems of the standard design $H = \frac{1}{2m}p^2 + U(x)$ simplifies:

$$P_t = \big\{U_xP_p - (p/m)P_x\big\} - \tfrac{1}{3!}\big(\tfrac{\hbar}{2}\big)^2U_{xxx}P_{ppp} + \tfrac{1}{5!}\big(\tfrac{\hbar}{2}\big)^4U_{xxxxx}P_{ppppp} - \cdots \tag{40}$$

└ becomes exactly classical when the Hamiltonian depends at most quadratically on its arguments

For a free particle we recover the condition $P_t = -(p/m)P_x$ which O'Connell & Wigner sought to replace.

Equation (38) is exactly equivalent to the Schrödinger equation. It presents not a physical alternative to, but simply a reformulation of orthodox quantum mechanics. It possesses some obviously attractive properties, but those are purchased at a price: while the Schrödinger equation is (when the Hamiltonian depends at most quadratically upon momentum) a linear partial differential equation of 2$^{\text{nd}}$ order, (38) is a linear partial differential equation of *infinite* order. The latter circumstance cuts us off from the familiar resources of Sturm-Liouville theory, and suggests that (38) might more naturally be formulated as an integral equation. This can—sometimes usefully—be done:

one obtains[20]

$$\tfrac{\partial}{\partial t}P(x,p;t) = \iint \mathcal{K}(x,p;x_0,p_0)P(x_0,p_0;t)\,dx_0dp_0$$

with

$$\mathcal{K}(x,p;x_0,p_0) = 2\pi\big(\tfrac{2}{\hbar}\big)^3 \iint H(x',p')\sin\left\{\tfrac{2}{\hbar}\det\begin{pmatrix}1 & x & p\\ 1 & x' & p'\\ 1 & x_0 & p_0\end{pmatrix}\right\}dx'dp'$$

but I will not belabor the point.

Often (more often quantum mechanically than classically) one has special interest in those aspects of dynamics where in fact nothing moves. I allude to the practical importance we attach to the *time-independent Schrödinger equation*. I discuss now—first in general terms, then in reference to a familiar example—how that theory fits within the phase space formalism.

Suppose $\mathsf{H}|n) = E_n|n)$. Temporally the eigenfunctions "buzz"

$$|n) \longrightarrow e^{-i\omega_n t}|n) \quad\text{with}\quad \omega_n \equiv E_n/\hbar$$

but the exponential buzz factors cancel when one constructs $\boldsymbol{\rho}_n \equiv |n)(n|$. It is evident that $\mathsf{H}\boldsymbol{\rho}_n = \frac{1}{2}(\mathsf{H}\boldsymbol{\rho}_n + \boldsymbol{\rho}_n\mathsf{H}) + \frac{1}{2}(\mathsf{H}\boldsymbol{\rho}_n - \boldsymbol{\rho}_n\mathsf{H}) = E_n\boldsymbol{\rho}_n$, and evident also that $(\mathsf{H}\boldsymbol{\rho}_n - \boldsymbol{\rho}_n\mathsf{H}) = \mathbf{0}$. So we have

$$\mathsf{H}|n) = E_n|n) \quad\Longleftrightarrow\quad \begin{cases}\frac{1}{2}(\mathsf{H}\boldsymbol{\rho}_n - \boldsymbol{\rho}_n\mathsf{H}) = \mathbf{0}\\ \frac{1}{2}(\mathsf{H}\boldsymbol{\rho}_n + \boldsymbol{\rho}_n\mathsf{H}) = E_n\boldsymbol{\rho}_n\end{cases} \tag{41}$$

In density operator language the time-independent theory hinges on a *pair* of equations, which in the phase space formalism become

$$\begin{aligned}\sin\Big\{\tfrac{\hbar}{2}\Big[\big(\tfrac{\partial}{\partial x}\big)_H\big(\tfrac{\partial}{\partial p}\big)_P - \big(\tfrac{\partial}{\partial x}\big)_P\big(\tfrac{\partial}{\partial p}\big)_H\Big]\Big\}H(x,p)P_n(x,p) &= 0\\ \cos\Big\{\tfrac{\hbar}{2}\Big[\big(\tfrac{\partial}{\partial x}\big)_H\big(\tfrac{\partial}{\partial p}\big)_P - \big(\tfrac{\partial}{\partial x}\big)_P\big(\tfrac{\partial}{\partial p}\big)_H\Big]\Big\}H(x,p)P_n(x,p) &= E_nP_n(x,p)\end{aligned} \tag{42}$$

We look to see what (42) has to say about a couple of examples, of which the first is standard to the phase space literature . . . and the second isn't.

---

[20] See QUANTUM MECHANICS (1967), Chapter 3, p. 110 –114. That discussion owes a little to Wigner but much to Moyal, and culminates in a description of the "phase space propagator"—a function of the form $\mathcal{K}(x,p,t;x_0,p_0,t_0)$ that permits one to write

$$P(x,p;t) = \iint \mathcal{K}(x,p,t;x_0,p_0,t_0)P(x_0,p_0;t_0)\,dx_0dp_0$$

and is therefore an object which would assume high importance if one were to pursue this subject to its depths.

HARMONIC OSCILLATOR This system derives its special simplicity from the circumstance that

$$H(x,p) = \tfrac{1}{2m}p^2 + \tfrac{1}{2}m\omega^2x^2 \tag{43.1}$$

is quadratic in its arguments. The first of equations (42) therefore reads

$$[H, P_n] = 0 \quad : \quad P_n(x,p) \text{ is a classical constant of the motion} \tag{43.2}$$

From the classical theory of conservative systems with one degree of freedom it follows therefore that $P_n(x,p)$ can be described $P_n(x,p) = f_n(H(x,p))$. Returning with this information to the second of equations (42)—which when the Hamiltonian is quadratic reads

$$HP - \tfrac{1}{2}\big(\tfrac{\hbar}{2}\big)^2\big\{H_{xx}P_{pp} - 2H_{xp}P_{px} + H_{pp}P_{xx}\big\} = EP \tag{43.3}$$

—we use

$$P = f(H) \quad : \quad \begin{aligned} P_{pp} &= (H_p)^2 f_{HH} + H_{pp} f_H \\ P_{xx} &= (H_x)^2 f_{HH} + H_{xx} f_H \end{aligned}$$

and

$$\begin{aligned} H_x &= m\omega^2 x \\ H_p &= p/m \end{aligned} \quad : \quad \begin{aligned} H_{xx} &= m\omega^2 \\ H_{xp} &= 0 \\ H_{pp} &= 1/m \end{aligned}$$

to obtain (after simplifications)

$$\Big\{H - \tfrac{1}{4}(\hbar\omega)^2 H \tfrac{d^2}{dH^2} - \tfrac{1}{4}(\hbar\omega)^2 \tfrac{d}{dH}\Big\}f = Ef$$

Multiplication by $(\hbar\omega)^{-1}$ gives

$$\Big\{W - \tfrac{1}{4}W\tfrac{d^2}{dW^2} - \tfrac{1}{4}\tfrac{d}{dW}\Big\}g = \epsilon g$$

where $W \equiv H/\hbar\omega$, $\epsilon \equiv E/\hbar\omega$ are dimensionless variables, and $g(W) \equiv f(H)$. Adjust the dependent variable

$$g(W) \longmapsto k(W) \equiv e^{2W} g(W)$$

and, after some elementary rearrangement, obtain

$$\Big\{\tfrac{1}{4}W\tfrac{d^2}{dW^2} + \tfrac{1-4W}{4}\tfrac{d}{dW} + \big(\epsilon - \tfrac{1}{2}\big)\Big\}k = 0$$

which by a final adjustment becomes

$$\Big\{\mathcal{E}\tfrac{d^2}{d\mathcal{E}^2} + (1-\mathcal{E})\tfrac{d}{d\mathcal{E}} + n\Big\}\ell = 0 \tag{43.4}$$
$$n \equiv \epsilon - \tfrac{1}{2}$$

where $\mathcal{E} \equiv 4W = 4H/\hbar\omega$ and $\ell(\mathcal{E}) \equiv k(W) = e^{2H/\hbar\omega} f(H)$. The point of preceding manipulations has come finally into view, for (43.4) is precisely *Laguerre's differential equation*.[21] The solutions of interest (those that will lead us to normalizable Wigner distributions) become available if and only if $n = 0, 1, 2, \ldots$, and are called "Laguerre polynomials." We are brought thus to the conclusion that

$$E_n = \hbar\omega\epsilon_n = \hbar\omega(n + \tfrac{1}{2}) \quad : \quad n = 0, 1, 2, \ldots \tag{43.5}$$

$$P_n(x, p) = C_n\, e^{-\frac{1}{2}\mathcal{E}} L_n(\mathcal{E}) \tag{43.6}$$

where the constants C_n are to be fixed by the requirement that $\int P_n\, dxdp = 1$ and where

$$\begin{aligned} L_0(z) &= 1 \\ L_1(z) &= 1 - z \\ L_2(z) &= 1 - 2z + \tfrac{1}{2}z^2 \\ L_3(z) &= 1 - 3z + \tfrac{3}{2}z^2 - \tfrac{1}{6}z^3 \\ &\vdots \end{aligned}$$

Looking finally to the explicit evaluation of the normalization constants C_n . . .

In terms of the dimensionless phase coordinates introduced previously[13] we have $dxdp = \frac{\hbar}{2} d\varkappa d\wp$ which, if we use $\varkappa = \sqrt{\mathcal{E}}\cos\vartheta, \wp = \sqrt{\mathcal{E}}\sin\vartheta$ to install polar coordinates on the dimensionless phase plane, becomes $dxdp = \frac{\hbar}{4} d\vartheta\, d\mathcal{E}$. So we have

$$\iint P_n(x, p)\, dxdp = C_n \tfrac{\hbar}{4} 2\pi \int_0^\infty e^{-\frac{1}{2}\mathcal{E}} L_n(\mathcal{E})\, d\mathcal{E}$$

Classical theory supplies the generating function $\frac{1}{1-t} e^{-zt/(1-t)} = \sum_n L_n(z) t^n$ so we have

$$\begin{aligned} \sum_{n=0}^\infty \Big\{ \int_0^\infty e^{-\frac{1}{2}\mathcal{E}} L_n(\mathcal{E})\, d\mathcal{E} \Big\} t^n &= \tfrac{1}{1-t} \int_0^\infty e^{-\frac{1}{2}\mathcal{E} - \mathcal{E}\, t/(1-t)}\, d\mathcal{E} \\ &= \tfrac{1}{1-t} \int_0^\infty e^{-\frac{1}{2}\frac{1+t}{1-t}\mathcal{E}}\, d\mathcal{E} \\ &= \tfrac{2}{1+t} = 2\,(1 - t + t^2 - t^3 + t^4 - \cdots) \end{aligned}$$

from which it follows that $C_n = (-)^n \frac{2}{h}$. Returning with this information to (43.6), we have

$$P_n(x, p) = (-)^n \tfrac{2}{h}\, e^{-\frac{1}{2}\mathcal{E}} L_n(\mathcal{E}) \tag{43.7}$$

—in exact agreement with the results (25) obtained previously by other means.

---

[21] See J. Spanier & K. O. Oldham, *An Atlas of Functions* (1987), 23:3:5. The whole of Chapter 23—to which I refer henceforth without specific attribution—is given over to an excellent account of properties of the Laguerre polynomials.

The preceding exercise serves to demonstrate that

- the phase space formalism can be made the basis (at least in favorable cases) of effective quantum mechanical calculation, but
- lends new patterns to the analytical details.

Carrying this discussion just a little farther: by application of Beck's trick (28) we might expect to have

$$\begin{aligned}\sum_n \psi_n^*(0)\,\psi_n(x)\,t^n &= \sum_n \Big\{\int P_n(\tfrac{x}{2},p)e^{\frac{i}{\hbar}xp}\,dp\Big\}t^n \\ &= \int e^{\frac{i}{\hbar}xp}\Big\{\sum_n P_n(\tfrac{x}{2},p)\,t^n\Big\}\,dp\end{aligned}$$

But $\sum P_n t^n = \frac{2}{h}e^{-\frac{1}{2}\mathcal{E}}\sum L_n(\mathcal{E})(-t)^n = \frac{2}{h}e^{-\frac{1}{2}\mathcal{E}}\frac{1}{1+t}e^{\mathcal{E}t/(1+t)} = \frac{2}{h}\frac{1}{1+t}e^{\frac{1}{2}\frac{t-1}{t+1}\mathcal{E}}$ so

$$\begin{aligned}&= \tfrac{2}{h}\tfrac{1}{1+t}\int \exp\Big\{\tfrac{i}{\hbar}xp - \tfrac{1-t}{1+t}\tfrac{1}{\hbar}\big[\tfrac{1}{m\omega}p^2 + m\omega\big(\tfrac{x}{2}\big)^2\big]\Big\}\,dp \\ &= \sqrt{\tfrac{2m\omega}{h}}\,\tfrac{1}{\sqrt{1-t^2}}\,\exp\Big\{-\tfrac{1+t^2}{1-t^2}\tfrac{m\omega}{2\hbar}x^2\Big\} \\ &= \tfrac{1}{a\sqrt{2\pi}}\,\tfrac{1}{\sqrt{1-t^2}}\,\exp\Big\{-\tfrac{1+t^2}{1-t^2}\tfrac{1}{4}\varkappa^2\Big\} \quad \text{in notation of p. 11} \\ &= \tfrac{1}{a\sqrt{2\pi}}e^{-\frac{1}{4}\varkappa^2}\Big\{1 + \tfrac{1}{2}\big[1-\varkappa^2\big]t^2 \\ &\qquad + \tfrac{1}{8}\big[3 - 6\varkappa^2 + \varkappa^4\big]t^4 \\ &\qquad + \tfrac{1}{48}\big[15 - 45\varkappa^2 + 15\varkappa^4 - \varkappa^6\big]t^6 + \cdots\Big\}\end{aligned}$$

At $x = 0$ (which is to say: at $\varkappa = 0$) we therefore have

$$\sum_n |\psi_n(0)|^2\,t^n = \tfrac{1}{a\sqrt{2\pi}}\Big\{1 + \tfrac{1}{2}t^2 + \tfrac{3}{8}t^4 + \tfrac{15}{48}t^6\cdots\Big\}$$

The polynomials are recognized to be monic Hermitian:

$$\begin{aligned}H_0(z) &= 1 \\ H_1(z) &= z \\ H_2(z) &= z^2 - 1 \\ H_3(z) &= z^3 - 3z \\ H_4(z) &= z^4 - 6z^2 + 3 \\ H_5(z) &= z^5 - 10z^3 + 15z \\ H_6(z) &= z^6 - 15z^4 + 45z^2 - 15 \\ &\;\vdots\end{aligned}$$

We are in position therefore to write

$$\begin{aligned}
\psi_0(x) &= \left[\tfrac{1}{a\sqrt{2\pi}}\right]^{-\frac{1}{2}}\tfrac{1}{a\sqrt{2\pi}}e^{-\frac{1}{4}\varkappa^2}H_0(\varkappa) &&= \frac{1}{\sqrt{a\sqrt{2\pi}}}\frac{1}{\sqrt{1\cdot 1}}e^{-\frac{1}{4}\varkappa^2}H_0(\varkappa)\\
\psi_2(x) &= \left[\tfrac{1}{a\sqrt{2\pi}}\tfrac{1}{2}\right]^{-\frac{1}{2}}\tfrac{1}{a\sqrt{2\pi}}\tfrac{1}{2}e^{-\frac{1}{4}\varkappa^2}H_2(\varkappa) &&= \frac{1}{\sqrt{a\sqrt{2\pi}}}\frac{1}{\sqrt{2\cdot 1}}e^{-\frac{1}{4}\varkappa^2}H_2(\varkappa)\\
\psi_4(x) &= \left[\tfrac{1}{a\sqrt{2\pi}}\tfrac{3}{8}\right]^{-\frac{1}{2}}\tfrac{1}{a\sqrt{2\pi}}\tfrac{1}{8}e^{-\frac{1}{4}\varkappa^2}H_4(\varkappa) &&= \frac{1}{\sqrt{a\sqrt{2\pi}}}\frac{1}{\sqrt{8\cdot 3}}e^{-\frac{1}{4}\varkappa^2}H_4(\varkappa)\\
\psi_6(x) &= \left[\tfrac{1}{a\sqrt{2\pi}}\tfrac{15}{48}\right]^{-\frac{1}{2}}\tfrac{1}{a\sqrt{2\pi}}\tfrac{1}{48}e^{-\frac{1}{4}\varkappa^2}H_6(\varkappa) &&= \frac{1}{\sqrt{a\sqrt{2\pi}}}\frac{1}{\sqrt{48\cdot 15}}e^{-\frac{1}{4}\varkappa^2}H_6(\varkappa)\\
&\vdots
\end{aligned}$$

But $\sqrt{1\cdot 1} = \sqrt{0!}\,,\ \sqrt{2\cdot 1} = \sqrt{2!}\,,\ \sqrt{8\cdot 3} = \sqrt{4!}\,,\ \sqrt{48\cdot 15} = \sqrt{6!}\,,\ \ldots$ so we have obtained normalized oscillator eigenstates which agree precisely with those presented in the text books. We missed the states of odd order because we placed Beck's reference point at the origin . . . where, as it happens, the oscillator states $\psi_{\text{odd}}(x)$ vanish.

PARTICLE IN FREE FALL The Hamiltonian

$$H(x,p) = \tfrac{1}{2m}p^2 + mgx \tag{44.1}$$

again has the property that x and p enter with powers not exceeding two, so the resulting physics exhibits some of the simplicity of oscillator theory, from which in other respects it differs profoundly. It is to introduce some mathematical ideas and notation (and to prepare the ground for a surprising development) that I look first to the seldom-discussed wave mechanics[22] of free fall, and take up the phase space formulation of the problem only after that preparation is complete.

The time-independent Schrödinger equation $-\frac{\hbar^2}{2m}\psi'' + mgx\psi = E\psi$ can be written

$$\psi''(x) = \tfrac{2m^2g}{\hbar^2}\big(x - \tfrac{E}{mg}\big)\psi(x) \tag{44.2}$$

which by change of variable[23]

$$x \longmapsto y_E \equiv \big(\tfrac{2m^2g}{\hbar^2}\big)^{\frac{1}{3}}\big(x - \tfrac{E}{mg}\big) \tag{44.3}$$

becomes

$$\tfrac{d^2}{dy^2}\Psi(y) = y\Psi(y) \tag{44.4}$$

This is *Airy's differential equation*, first encountered in George Airy's "Intensity of light in the neighborhood of a caustic" (1838). The solutions are linear

[22] I use that antique term to distinguish Schrödinger's $\psi(x)$-theory from other formulations of quantum mechanics.

[23] The subscript emphasizes that the *the eigenvalue E has been absorbed into the definition of the independent variable*, and will be omitted when its presence makes no immediate difference.

combinations of the *Airy functions* $Ai(y)$ and $Bi(y)$, which are close relatives of the Bessel functions of orders $\pm\frac{1}{3}$, and of which (since $Bi(y)$ diverges as $y \to \infty$) only the former

$$Ai(y) \equiv \tfrac{1}{\pi}\int_0^\infty \cos\Big(yu + \tfrac{1}{3}u^3\Big)\,du \tag{44.5}$$

will concern us.[24] To gain insight into the origin of Airy's construction, write

$$f(y) = \tfrac{1}{2\pi}\int_{-\infty}^{+\infty} g(u)e^{iyu}\,du$$

and notice that $f'' - yf = 0$ entails

$$\tfrac{1}{2\pi}\int_{-\infty}^{+\infty}\big[-u^2 g(u) + ig(u)\tfrac{d}{du}\big]e^{iyu}\,du = 0$$

Integration by parts gives

$$\tfrac{1}{2\pi}ig(u)e^{iyu}\Big|_{-\infty}^{+\infty} - \tfrac{1}{2\pi}\int_{-\infty}^{+\infty}\big[u^2 g(u) + ig'(u)\big]e^{iyu}\,du = 0$$

The leading term vanishes if we require $g(\pm\infty) = 0$. We are left then with a first-order differential equation $u^2 g(u) + ig'(u) = 0$ of which the general solution is $g(u) = A\cdot e^{i\frac{1}{3}u^3}$. So we have

$$f(y) = A\cdot\tfrac{1}{2\pi}\int_{-\infty}^{+\infty} e^{i\,(yu+\frac{1}{3}u^3)}\,du = A\cdot\tfrac{1}{\pi}\int_0^\infty \cos\Big(yu + \tfrac{1}{3}u^3\Big)\,du$$

It was to achieve

$$\int_{-\infty}^{+\infty} Ai(y)\,dy = 1 \tag{44.6}$$

that Airy assigned the value $A = 1$ to the constant of integration.

Returning with this mathematics to the quantum physics of free fall, we see that solutions of the Schrödinger equation (44.2) can be described

$$\psi_{\mathcal{E}}(x) = N\cdot Ai\big(k(x - a_E)\big) \tag{44.7}$$

where N is a normalization factor (soon to be determined), and where

$$k \equiv \big(\tfrac{2m^2 g}{\hbar^2}\big)^{\frac{1}{3}} = \frac{1}{\text{“natural length” of the quantum free fall problem}}$$

$$a_E \equiv \tfrac{E}{mg} = \text{classical turning point of a particle lofted with energy } E$$

$$\mathcal{E} \equiv ka_E = \frac{E}{mg(\text{natural length})} \equiv \text{dimensionless energy parameter}$$

It is a striking fact—evident in (44.7)—that the eigenfunctions $\psi_E(x)$ *all have the same shape* (i.e., are translates of one another: see Figure 3), and remarkable

---

[24] For a summary of the properties of Airy functions see Chapter 56 in Spanier & Oldham.[21] Those functions are made familiar to students of quantum mechanics by their occurance in the “connection formulæ” of simple WKB approximation theory: see Griffiths' §8.3, or C. M. Bender & S. A. Orszag, *Advanced Mathematical Methods for Scientists & Engineers* (1978), §10.4.

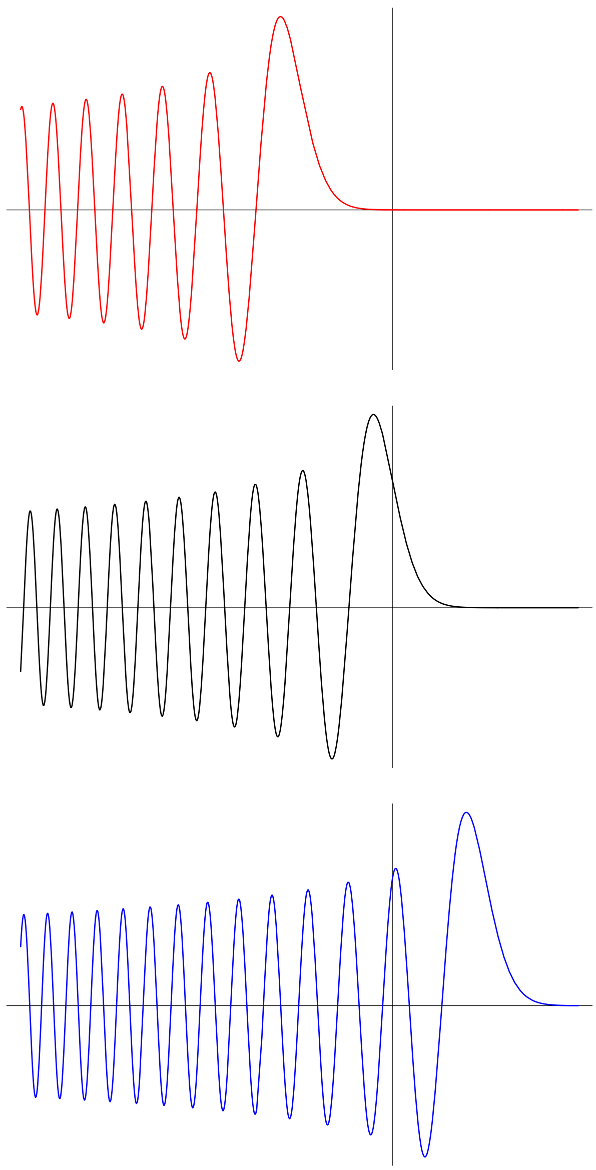

FIGURE 3: *Free fall eigenfunctions* $\psi_E(x)$ *with* $E<0$, $E=0$, $E>0$, *in descending order. The remarkable* <u>*translational similarity*</u> *of the eigenfunctions is perhaps not surprising, in view of the translational similarity of the parabolic graphs of the solutions*

$$x(t) - a_E = -\tfrac{1}{2}g(t-t_0)^2$$

of the classical free fall equation $m\ddot{x} = -mgx$.

also that the *the energy spectrum is continuous, and has no least member*: the system possesses no ground state. One might view this highly unusual circumstance to be a consequence of the notion that *"free fall" is free motion relative to a non-inertial frame*.

The eigenfunctions $\psi_{\mathcal{E}}(x)$ share with the free particle functions $e^{\pm\frac{i}{\hbar}\sqrt{2mE}\,x}$ the property that they are not individually normalizable,[25] but require assembly into "wavepackets." They do, however, comprise a complete orthonormal set, in the sense which I digress now to establish. Let

$$f(y,m)\equiv Ai(y-m)$$

To ask of the m-indexed functions $f(y,m)$

- Are they *orthonormal*: $\int f(y,m)f(y,n)\,dy=\delta(m-n)$?
- Are they *complete*: $\int f(x,m)f(y,m)\,dm=\delta(x-y)$?

is, in fact, to ask the same question twice, for both are notational variants of this question: Does

$$\int_{-\infty}^{+\infty} Ai(y-m)Ai(y-n)\,dy=\delta(m-n)?$$

An affirmative answer (which brings into being a lovely "Airy-flavored Fourier analysis") is obtained as follows:

$$\begin{aligned}
&= \left(\tfrac{1}{2\pi}\right)^2\iiint e^{i\,[(y-m)u+\frac{1}{3}u^3]}e^{i\,[(y-n)v+\frac{1}{3}v^3]}\,dudvdy\\
&= \tfrac{1}{2\pi}\iint e^{i\frac{1}{3}(u^3+v^3)}e^{-i(mu+nv)}\underbrace{\left\{\tfrac{1}{2\pi}\int e^{iy(u+v)}\,dy\right\}}_{\delta(u+v)}dudv\\
&= \tfrac{1}{2\pi}\int \underbrace{e^{i\frac{1}{3}(v^3-v^3)}}_{1}e^{iv(m-n)}\,dv=\delta(m-n)
\end{aligned}$$

So for our free fall wave functions we have the "orthogonality in the sense of Dirac:"

$$\begin{aligned}
\int_{-\infty}^{+\infty}\psi^*_{\mathcal{E}'}(x)\,\psi_{\mathcal{E}''}(x)\,dx &= N^2\int_{-\infty}^{+\infty} Ai\big(k(x-a_{E'}\big)Ai\big(k(x-a_{E''}\big)\,dx\\
&= N^2\tfrac{1}{k}\cdot\delta(\mathcal{E}'-\mathcal{E}'')\\
&\downarrow\\
&= \delta(\mathcal{E}'-\mathcal{E}'')\quad\text{if we set } N=\sqrt{k}
\end{aligned}\tag{44.8}$$

The functions thus normalized are complete in the sense that

$$\int_{-\infty}^{+\infty}\psi^*_{\mathcal{E}}(x')\,\psi_{\mathcal{E}}(x'')\,d\mathcal{E}=\delta(x'-x'')\tag{44.9}$$

---

[25] Asymptotically $Ai^2(y)\sim\frac{1}{\pi\sqrt{|y|}}\sin^2\left(\frac{2}{3}|y|^{\frac{3}{2}}+\frac{\pi}{4}\right)$ dies as $y\downarrow-\infty$, but so slowly that the limit of $\int_y^0 Ai^2(u)\,du$ blows up.

I record one final result which issues from Schrödinger's formulation of the free fall problem. A few lines of fairly straightforward calculation[26] lead to the conclusion that the associated Green's function can be described

$$\begin{aligned} G(x,t;x_0,0) &= \int \psi_{\mathcal{E}}(x)\psi^*_{\mathcal{E}}(x_0)e^{-\frac{i}{\hbar}E(\mathcal{E})t}\,d\mathcal{E} \quad \text{with} \quad E(\mathcal{E}) \equiv (mg/k)\mathcal{E} \\ &= \sqrt{\tfrac{m}{iht}}\exp\Big\{\tfrac{i}{\hbar}\big[\tfrac{m}{2t}(x-x_0)^2 - \tfrac{mg}{2}(x+x_0)t - \tfrac{mg^2}{24}t^3\big]\Big\} \end{aligned} \tag{44.10}$$

We observe that this result can be notated (compare (0–95))

$$= \sqrt{\tfrac{i}{\hbar}\tfrac{\partial^2 S}{\partial x \partial x_0}}\,e^{\frac{i}{\hbar}S(x,t;x_0,0)}$$

and that the $S(x_1,t_1;x_0,t_0)$ thus defined is precisely the *classical action function* associated with the dynamical path

$$(x_0,t_0) \xrightarrow[\text{free fall}]{} (x_1,t_1)$$

The latter fact is non-obvious, but emerges when one introduces

$$x(t) = -\tfrac{1}{2}gt^2 + \tfrac{(x_1-x_0)-\frac{1}{2}g(t_1^2-t_0^2)}{t_1-t_0}\,t + \tfrac{(x_0-\frac{1}{2}gt_0^2)t_1-(x_1-\frac{1}{2}gt_1^2)t_0}{t_1-t_0}$$

(the free fall parabola which links the specified spacetime points) into

$$S[x(t)] = \int_{t_0}^{t_1}\big\{\tfrac{1}{2}m\dot{x}(t)^2 - mgx(t)\big\}\,dt$$

and performs the simple integration.[27] With $G(x,t;x_0,0)$ now in hand we are in position to study the *free fall of lofted wavepackets* . . . but won't; this is done in the notes to which I refer below.

Returning now to the phase space formalism, we introduce $\psi_{\mathcal{E}}(x)$ into (22.1) and undertake to obtain a description of $P_{\mathcal{E}}(x,p)$. We have

$$P_{\mathcal{E}}(x,p) = \tfrac{2}{h}k\int Ai\big(k(x+\xi)-\mathcal{E}\big)e^{2\frac{i}{\hbar}p\xi}Ai\big(k(x-\xi)-\mathcal{E}\big)\,d\xi$$

Introduce dimensionless variables $y \equiv kx$, $\zeta \equiv k\xi$, $q \equiv p/\hbar k$ and obtain

---

[26] Details can be found on p. 32 of my "Classical/quantum mechanics of a bouncing ball" (1994), which provides a fairly exhaustive account of the classical and quantum physics of constrained/unconstrained free fall. I hope to produce an electronic version of that material in the not-too-distant future. I would expect to include material developed at that same time by Richard Crandall. In the meantime, see pp. 101–105 of S. Flüge, *Practical Quantum Mechanics* (1974) for discussion of the rudiments of the bouncing ball problem; I am indebted to Robert Reynolds for this reference.

[27] See QUANTUM MECHANICS (1967), Chapter 1, p. 21 for the details.

$$
\begin{aligned}
P_{\mathcal{E}} &= \tfrac{2}{\hbar}\int Ai\big(y+\zeta-\mathcal{E}\big)e^{i2q\zeta}Ai\big(y-\zeta-\mathcal{E}\big)\,d\zeta\\
&= \tfrac{2}{\hbar}\big(\tfrac{1}{2\pi}\big)^2\iiint e^{i\left[(y+\zeta-\mathcal{E})u+\frac{1}{3}u^3\right]+i2q\zeta+i\left[(y-\zeta-\mathcal{E})v+\frac{1}{3}v^3\right]}\,dudvd\zeta
\end{aligned}
$$

The ζ-integral eats a $\frac{1}{2\pi}$-factor and burps out $\delta(u-v+2q)$. We then get the v-integral for free, and are left with

$$
P_{\mathcal{E}} = \tfrac{2}{\hbar}\tfrac{1}{2\pi}\int e^{i\left[\frac{2}{3}u^3+2qu^2+(4q^2+2\hat{y})u+(\frac{8}{3}q^3+2q\hat{y})\right]}\,du \quad : \quad \hat{y}\equiv y-\mathcal{E} \tag{44.11}
$$

Now, it has been known for millennia that the term of next-to-leading-order in

$$
F(x)\equiv Ax^n+Bx^{n-1}+Cx^{x-2}+\cdots+Px+Q
$$

can be killed by translation; i.e., that it is always possible—and invariably useful—to exhibit a polynomial of the design $f(x)=Ax^n+0+cx^{n-2}+\cdots+px+q$ such that

$$
F(x)=f(x+\tfrac{B}{nA})
$$

In the present instance it is a wonderful fact that the tranlation designed to kill the quadratic term *kills also the constant term*; i.e., that

$$
\begin{aligned}
\tfrac{2}{3}u^3+2qu^2+(4q^2+2\hat{y})u+(\tfrac{8}{3}q^3+2q\hat{y}) &= \tfrac{2}{3}(u+q)^3+2(q^2+\hat{y})(u+q)\\
&= \tfrac{1}{3}w^3+2^{\frac{2}{3}}(q^2+\hat{y})w\\
&\qquad\qquad w\equiv 2^{\frac{1}{3}}(u+q)
\end{aligned}
$$

Returning with this information to (44.11) we have

$$
\begin{aligned}
P_{\mathcal{E}}(x,p) &= \tfrac{2}{\hbar}2^{-\frac{1}{3}}\cdot\tfrac{1}{2\pi}\int e^{i\left[\frac{1}{3}w^3+2^{\frac{2}{3}}(q^2+\hat{y})w\right]}\,dw\\
&= \tfrac{2}{\hbar}2^{-\frac{1}{3}}\cdot Ai\big(2^{\frac{2}{3}}(q^2+y-\mathcal{E})\big) \qquad (44.12)
\end{aligned}
$$

I promised at the outset a "surprising development," and it is this: in the quantum oscillator problem we encountered

$$
\text{Hermite}\xrightarrow[\text{Wigner}]{}\text{Laguerre}
$$

but the problem of quantum mechanical free fall is in this respect elegantly symmetric:

$$
\text{Airy}\xrightarrow[\text{Wigner}]{}\text{Airy}
$$

The construction $q^2+y-\mathcal{E}$ which enters as the argument of the Airy function on the right side of (44.12) can be understood as follows: The system under consideration affords a

$$
\text{natural energy} = mg\cdot(\text{natural length}) = mg/k
$$

and the equations $\frac{1}{2m}p^2+mgx-E=0$ inscribe isoenergetic parabolas on classical phase space. Division by the natural energy yields equations which in

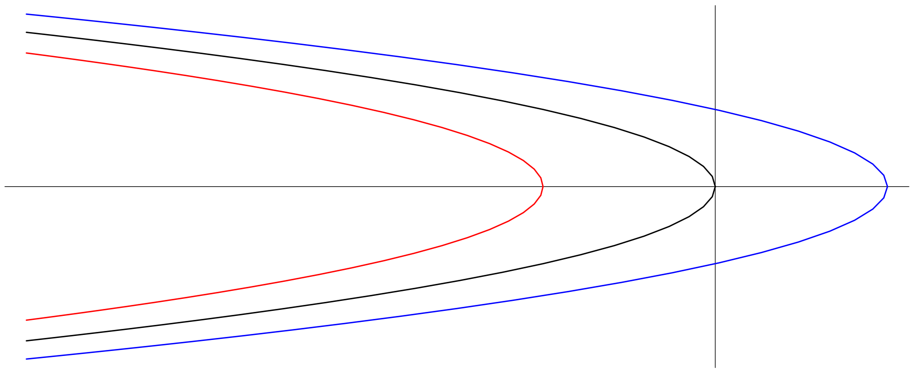

FIGURE 4: *Classical isoenergetic curves on the dimensionless phase plane. The turning point occurs at* $y = \mathcal{E}$. *The curves shown have* $\mathcal{E} < 0$, $\mathcal{E} = 0$, $\mathcal{E} > 0$.

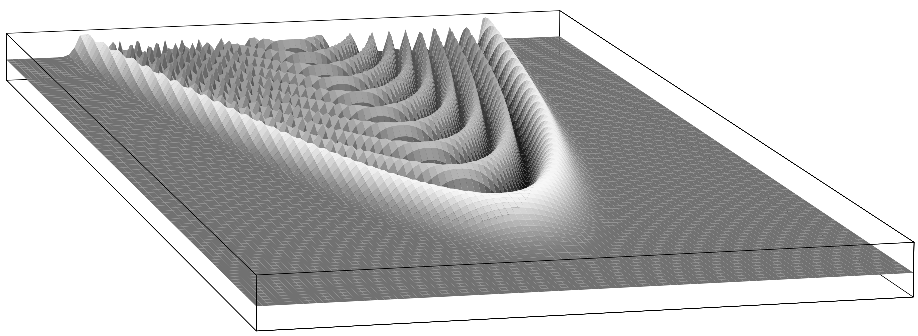

FIGURE 5: *Each of the* $\mathcal{E}$*-indexed Wigner functions* $P_{\mathcal{E}}(x, p)$ *is translationally equivalent to each of the others, and each is constant on each of the curves shown in the preceding figure. The functions assume negative values in each of the troughs: see again Figure 3, which can be read now as a description of* $P_{\mathcal{E}}(x, 0)$.

terms of the "dimensionless momentum/length/energy" variables q, y, and $\mathcal{E}$ read $q^2 + y - \mathcal{E} = 0$. Those equations inscribe $\mathcal{E}$-parameterized coaxial parabolas on the $\{y, q\}$-plane (dimensionless phase plane), as shown in Figure 4. Each of the Wigner functions $P_{\mathcal{E}}(x, p)$ is constant on each of those curves (Figure 5).

The eigenstates $\psi_{\mathcal{E}}(x)$ do not describe possible quantum states of the free fall system for the reason—already remarked—that they are *not normalizable*. The same has now to be said—for the same reason—of the Wigner functions $P_{\mathcal{E}}(x, p)$. The point can be established either by general argument

$$\int P_{\mathcal{E}}(x, p)\, dp = |\psi_{\mathcal{E}}(x)|^2 \quad \text{and} \quad \int |\psi_{\mathcal{E}}(x)|^2\, dx = \infty$$

or more specifically:

$$\iint P_{\mathcal{E}}(x,p)\,dxdp = \iint \tfrac{2}{h}2^{-\frac{1}{3}}\cdot Ai\big(2^{\frac{2}{3}}(q^2+y-\mathcal{E}\big)\,\hbar dydq$$
$$= \tfrac{1}{2\pi}\int dq = \infty$$

I will return in a moment to discussion of the implications of this circumstance.

Suppose we had elected to proceed directly from (42)—for free fall as we did for the oscillator. We would as before be led to the conclusion that

$$P(x,p) = f(H(x,p))$$

and by simple adjustment of our former argument to the conclusion that $f(H)$ must satisfy

$$\Big\{H - \tfrac{1}{8}\hbar^2 mg^2\tfrac{d^2}{dH^2}\Big\}f = Ef$$

which if we write $\mathcal{H}\equiv\frac{k}{mg}H = q^2+y$ ("dimensionless energy") becomes

$$\tfrac{1}{4}\tfrac{d^2}{d\mathcal{H}^2}f = (\mathcal{H}-\mathcal{E})f$$

Thus are we led—with swift economy—back to the statement first encountered at (44.12):

$$P_{\mathcal{E}}(x,p) = (\text{constant})\cdot Ai\big(4^{\frac{1}{3}}(\mathcal{H}-\mathcal{E})\big)$$

But we find ourselves now (as then) unable to use $\iint P_{\mathcal{E}}(x,p)\,dxdp = 1$ to assign enforced value to the numerical prefactor.

Were this discussion to be protracted one would want to consider (among other things) how—in analytical detail—it comes about that

$$\text{free particle theory arises from}\begin{cases}\text{the oscillator as }\omega\downarrow 0\\ \text{free fall as }g\downarrow 0\end{cases}$$

The delicacy of the issue is made less surprising when one considers how different from one another (geometrically/topologically) are the families of isoenergetic curves encountered in the three cases. We will have occasion to review the free particle theory (but not the limiting process) in the next section.

Remarks concerning the distinction between "wavepackets" and "mixtures." Let $\big\{|n)\big\}$ be some orthonormal basis in the space of states, and let

$$|\psi) = \begin{cases}\sum_n c_n|n) & : \quad\text{discrete case}\\[2ex] \int c(n)|n)\,dn & : \quad\text{continuous case}\end{cases}$$

describe some superposition of such states. By vague convention we usually reserve the term "wavepacket" for circumstances in which $(x|\psi)$ is in some relevant sense "semi-localized," but here I find it convenient to abandon that restrictive convention. In the continuous case the orthonormality condition

$$(m|n) = \delta(m-n) \text{ renders } (n|n) = 1 \text{ impossible}$$

so in that case $|n)$ cannot refer literally to a "quantum state," but has rather the status of an *analytical crutch*. That $|\psi)$ refers to such a state is by

$$(\psi|\psi) = \iint (m|c^*(m)c(n)|n)\,dmdn = \int |c(n)|^2\,dn = 1$$

a responsiblity borne by its coordinates $c(n)$.

The projector onto $|\psi)$ can be described

$$|\psi)(\psi| = \begin{cases} \sum_m \sum_n |m)c_m c_n^*(n| & : \quad \text{discrete case} \\[2mm] \int\int |m)c(m)c^*(n)(n|\,dmdn & : \quad \text{continuous case} \end{cases}$$

Both formulæ present us with

- projection operators $|n)(n|$ on the diagonal, but
- non-projectors $|m)(n|$ at off-diagonal positions.

Projectivity is in either case an easy consequence of

$$(m|n) = \delta_{mn} \qquad \text{else} \qquad (m|n) = \delta(m-n)$$

and

$$\sum_n |c_n|^2 = 1 \qquad \text{else} \qquad \int |c(n)|^2\,dn = 1$$

If the off-diagonal terms could on some grounds be expunged[28] then we would be left with operators of the design

$$\boldsymbol{\rho} = \begin{cases} \sum_n |n)p_n(n| & : \quad \text{discrete case} \\[2mm] \int |n)p(n)dn(n| & : \quad \text{continuous case} \end{cases}$$

with $p_n \equiv |c_n|^2$ else $p(n) \equiv |c(n)|^2$. We would, in other words, be left with a density operator—the descriptor not of a wavepacket (superposition of states) but of a mixture of states. We are led to the view that

Mixtures are "incoherent superpositions"

[28] The simplest procedure: Write $c_n = a_n e^{i\phi_n}$ and average over all phases, though of (on what physical grounds?) as independent random variables. This is the *method of random phases*, encountered already once in each of the preceding chapters.

Any given state can be represented

$$\text{state} = \sum \text{component states}$$

in many ways. We assign objective significance to *no* such decomposition, but only to the measurement devices which project onto the elements of one or another of them. We allow ourselves, as a matter of analytical convenience (Fourier analysis provides an example), to write

$$\text{state} = \sum \text{component non-states}$$

even though no measurement device can "prepare (or project onto) a non-state." Similarly . . .

We write

$$\text{density operator } \boldsymbol{\rho}_{\text{mixture}} = \sum \text{weighted projectors } \boldsymbol{\rho}_{\text{pure}}$$

but have learned to assign objective significance to no particular representation of the mixture. The question before us: Is it (not physically but) formally possible/expedient to contemplate admixtures of "non-states"?[29] Is it sense or nonsense to write (as a moment ago we casually did) $\boldsymbol{\rho} = \int |n)p(n)dn(n|$? I assert that to write such a thing would be to write nonsense . . . on grounds that if $\{|a)\}$ is some arbitary basis (whether discrete or continuous: I arbitrarily assume the latter) then

$$\text{tr}\,\boldsymbol{\rho} = \iint (a|n)p(n)dn(n|a)\,da = \iint (n|a)da(a|n)\,p(n)dn = \int (n|n)\,p(n)dn$$

is uninterpretable; maybe infinite, but certainly not unity. Relatedly: while it makes sense to write

$$|\psi) = \int c(n)|n)\,dn \longrightarrow |n_0) \quad \text{when} \quad c(n) \rightarrow \delta(n - n_0)$$

and while $|\psi)(\psi| = \iint |m)c^*(m)c(n)(n|\,dmdn$ is a meaningful construction, it would be meaningless to assert that phase averaging yields $\int |n)|c(n)|^2(n|\,dn$, and doubly meaningless to allow $c(n) \rightarrow \delta(n - n_0)$, absurd to claim that the result can be described $|n_0)(n_0|$.

The immediate point of this discussion: it would be improper to construct

$$P(x,p) \equiv \int P_{\mathcal{E}}(x,p) \cdot w(\mathcal{E})\,d\mathcal{E}$$

and futile to expect to recover $\iint P(x,p)\,dxdp = 1$ from a stipulation of the form$\int w(\mathcal{E})\,d\mathcal{E} = 1$.

[29] By which term I understand continuously-indexed "states" which can be normalized only formally, in the sense of Dirac.

FREE PARTICLE This system—which in so many contexts is (if not "too simple to be interesting") deserving of the description "simplest possible" —is considered only now because it exhibits delicate anomalies of the sort just discussed. Working from $\mathsf{H}|\psi) = E|\psi)$ with $\mathsf{H} = \frac{1}{2m}\mathsf{p}^2$ one is led to energy eigenfunctions which can in the x-representation be described[30]

$$\psi_\wp(x) = \tfrac{1}{\sqrt{h}}e^{\frac{i}{\hbar}\wp x} \quad \text{with} \quad E = \tfrac{1}{2m}\wp^2 \equiv \tfrac{1}{2}mv^2 \tag{45.1}$$

Those continuously-indexed eigenfunctions are orthonormal only in the sense of Dirac

$$\int \psi_\wp^*(x)\,\psi_{\hat\wp}(x)\,dx = \delta(\wp - \hat\wp)$$

so refer not to proper quantum states, but to the formal devices I have called "non-states." When launched into motion they become

$$\begin{aligned}\psi_\wp(x,t) &\equiv \psi_\wp(x)\cdot e^{-\frac{i}{\hbar}E(\wp)t}\\ &= \tfrac{1}{\sqrt{h}}e^{\frac{i}{\hbar}[\wp x - \frac{1}{2m}\wp^2 t]}\end{aligned} \tag{45.2}$$

Returning with this information to (22.1) we easily obtain

$$P_\wp(x,p;t) = \tfrac{1}{h}\delta(p-\wp) \tag{45.3}$$

from which all t-dependence (ditto all x-dependence) has disappeared. That $P_\wp(x,p;t)$ does in fact satisfy the dynamical equation (39) is now almost obvious. But it is obvious also that the pathologies that famously haunt Schrödinger's free particle theory have been inherited by the phase space formalism. The $P_\wp(x,p;t)$ of (45.3) is everywhere non-negative, and at phase points off the isoenergetic line $p = \wp$ conforms to the boundedness condition (26). But on the line $P_\wp(x,p;t)$ becomes singular, and it is clear that $\iint P_\wp(x,p;t)\,dxdp \neq 1$.

In textbook quantum mechanics[31] one remedies the pathology either by placing the free particle in a large (confining = freedom-breaking) box, or by assembling normalized wavepackets. The latter process is, for present purposes, by far the most convenient. Form

$$\psi(x,t) = \int\Big\{\big[\tfrac{1}{\lambda\sqrt{2\pi}}\big]^{\frac{1}{2}}\exp\Big\{-\tfrac{1}{4}\big[\tfrac{\wp-\wp_0}{\lambda}\big]^2\Big\}\Big\}\,\psi_\wp(x,t)\,d\wp$$

and by integration[32] obtain

$$= \big[\tfrac{1}{\sigma[1+i(t/\tau)]\sqrt{2\pi}}\big]^{\frac{1}{2}}\exp\Big\{-\tfrac{x^2}{4\sigma^2[1+i(t/\tau)]} + \tfrac{i}{\hbar}\tfrac{\wp_0 x-(\wp_0^2/2m)}{1+i(t/\tau)}\Big\} \tag{45.4}$$

[30] See again (0–80).
[31] See L. I. Schiff, *Quantum Mechanics* (3rd edition 1968), §§10 & 12.
[32] The details are spelled out on p. 9 of my "Gaussian wavepackets" (1998).

where $\sigma \equiv \hbar/2\lambda$ and $\tau \equiv \hbar m/2\lambda^2 = 2m\sigma^2/\hbar$. Straightforward calculation now gives

$$|\psi(x,t)|^2 = \frac{1}{\sigma(t)\sqrt{2\pi}}\exp\Big\{-\tfrac{1}{2}\big[\tfrac{x-vt}{\sigma(t)}\big]^2\Big\} \tag{45.5}$$

which describes a Gaussian drifting to the right with constant speed $v = \wp_0/m$, growing progressively shorter/fatter[33] as indicated by the hyperbolic rule

$$\sigma^2(t) = \sigma^2[1 + (t/\tau)^2] \tag{45.6}$$

Returning with the normalized "launched Gaussian" (45.4) to Wigner's construction (22.1), we obtain[34]

$$\begin{aligned} P_{\text{gaussian}}(x,p;t) &= \tfrac{2}{h}\exp\Big\{-\big[\tfrac{x-vt}{\sigma} - (t/\tau)\tfrac{p-mv}{\lambda}\big]^2 - \tfrac{1}{2}\big[\tfrac{p-mv}{\lambda}\big]^2\Big\} \\ &= \frac{1}{\sigma\sqrt{2\pi}}\frac{1}{\lambda\sqrt{2\pi}}\exp\Big\{-\tfrac{1}{2}\big[\tfrac{x-(p/m)t}{\sigma}\big]^2 - \tfrac{1}{2}\big[\tfrac{p-mv}{\lambda}\big]^2\Big\} \end{aligned} \tag{45.7}$$

in which not only λ but also σ are *constants*, interrelated by $\sigma\lambda = \frac{1}{2}\hbar$. It is impossible to imagine a lovelier result: the distribution (45.7) is—for all assignments of $\lambda \sim \sigma^{-1}$ and $v \sim \wp_0$

- normalized;
- in compliance with the boundedness condition (26);
- everywhere non-negative.

The distribution is readily seen to be a solution of the dynamical equation (38), and the equation

$$P_{\text{gaussian}}(x,p;t) = \text{constant}$$
$$0 < \text{constant} < \tfrac{2}{h}$$

inscribes on phase space an ellipse, which *moves as though carried along by the classical free particle phase flow*. The resulting shear results in the temporal development of *x-p correlation*: see Figure 6.

In (45.7) we possess a class of distribution functions which (not quite obviously) exhibit time-dependent "dispersion," which at $t = 0$ is "minimal" (meaning "least allowed by the uncertainty principle"):

$$\begin{aligned} \Delta x \cdot \Delta p &= \tfrac{1}{2}\hbar \cdot \sqrt{1 + (t/\tau)^2} \\ &\downarrow \\ &= \tfrac{1}{2}\hbar \quad : \quad \text{minimal dispersion at } t = 0 \end{aligned}$$

Let us for a moment set aside all "free particle" considerations, and look to the bivariate normal distribution

$$P(x - x_0, p - p_0; \sigma, \lambda) \equiv \frac{1}{\sigma\sqrt{2\pi}}\frac{1}{\lambda\sqrt{2\pi}}\exp\Big\{-\tfrac{1}{2}\big[\tfrac{x-x_0}{\sigma}\big]^2 - \tfrac{1}{2}\big[\tfrac{p-p_0}{\lambda}\big]^2\Big\} \tag{46}$$

---

[33] It is interesting to notice that we could in principle have arranged things so that the "launched Gaussian" grows for a while progressively taller/skinnier, before yielding to the inevitable.

[34] See §6 in "Gaussian wavepackets"[32] for the tedious but straightforward computational details. Also the end of §8.

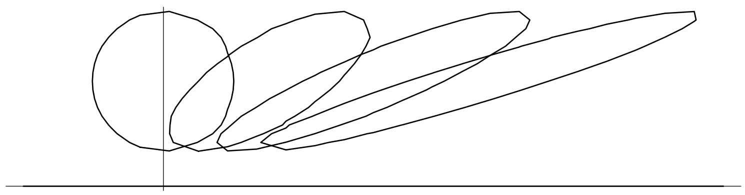

FIGURE 6: *Mechanism responsible for the dynamical development of correlation. The figure derives from (45.7), in which I have set* σ, λ, m *and* v *all equal to unity, and* $t = \{0, 1, 2, 3\}$. *A similar graphic appears on p. 204 of Bohm's text, but is claimed by him to refer only to the* classical *physics of a free particle, and because he works without knowledge of the phase space formalism he is obliged to be vaguely circumspect in drawing his quantum conclusions. We, however, are in position to identify the sense in which (48) pertains as directly and literally to the quantum physics of a free particle as it does to the classical physics. Also implicit in the figure are the statements*

$$\sigma_x(t) = \sigma\sqrt{1 + (t/\tau)^2}$$
$$\sigma_p(t) = \textit{constant}$$

which we associate familiarly with the quantum motion of Gaussian wavepackets, but are seen now to pertain equally well to the classical motion of Gaussian populations of free particles.

as a free-standing mathematical object. Note first that if we set

$$x_0 = p_0 = 0 \quad \text{and} \quad \sigma = \sqrt{\hbar/m\omega 2}\,, \;\; \lambda = \sqrt{\hbar m\omega/2}$$

then (46) gives back the harmonic oscillator groundstate (29), so (46) can be described as a "translated copy" of that state. The oscillator groundstate is the best known instance of a state of minimal dispersion.

It is the boundedness condition (26) which forces $\sigma\lambda \geqslant \frac{1}{2}\hbar$, and which declares "sub-minimal distributions" (those with $\sigma\lambda < \frac{1}{2}\hbar$) to be quantum mechanically disallowed.

It can be shown[35] that the inverse Wigner transform (Beck's trick (28)) leads to a *normalized* state $\psi_0(x)$—effectively: the groundstate of an oscillator—if and only if the minimality condition $\sigma\lambda = \frac{1}{2}\hbar$ is satisfied. To say the same thing another way: minimality is necessary and sufficient to insure that the distribution $P(x - x_0, p - p_0; \sigma, \lambda)$ satisfies the "pure state condition" (32).

All non-minimal instances of $P(x - x_0, p - p_0; \sigma, \lambda)$ must refer to mixtures. It is, we note in passing, very easy to construct representatives of mixtures from

[35] See §9 in "Gaussian wavepackets"[32] for the fairly straightforward details. Or see the discussion pursuant to (30).

the material now at hand; one has only to write

$$P(x,p) = \iiiint P(x-x_0, p-p_0; \sigma, \lambda)\, w(x_0,p_0,\sigma,\lambda)\, dx_0 dp_0 d\sigma d\lambda$$

where is an ordinary distribution on $\{x_0, p_0, \sigma, \lambda\}$-space. It is, perhaps, most natural to impose the minimality condition (so that we are mixing states, rather than mixing mixtures), and a simplification to fix the value of σ (whence also of λ); one then has

$$P(x,p) = \iint P(x-x_0, p-p_0; \sigma, \lambda)\, w(x_0,p_0)\, dx_0 dp_0 \tag{47}$$

In §12 of "Gaussian wavepackets"[32] I examine in particular detail the theory of "centered fat Gaussians"

$$P(x,p;\boldsymbol{\sigma},\boldsymbol{\lambda}) \equiv \tfrac{1}{\boldsymbol{\sigma}\sqrt{2\pi}}\tfrac{1}{\boldsymbol{\lambda}\sqrt{2\pi}} \exp\Big\{-\tfrac{1}{2}\big[\tfrac{x}{\boldsymbol{\sigma}}\big]^2 - \tfrac{1}{2}\big[\tfrac{p}{\boldsymbol{\lambda}}\big]^2\Big\} \tag{48}$$

where $\boldsymbol{\sigma} \equiv b\sigma$, $\boldsymbol{\lambda} \equiv b\lambda$ and $b \geqslant 1$ is the "fatness parameter." It emerges that

$$P(x,p;\boldsymbol{\sigma},\boldsymbol{\lambda}) = \sum_n p_n P_n(x,p) \tag{49}$$

where the $P_n(x,p)$ are precisely the *oscillator* Wigner functions encountered at (43.7), and where the weights p_n are given by

$$p_n = (-)^n \tfrac{1}{b^2}\int_0^\infty e^{-\frac{1}{2}[1+\frac{1}{b^2}]z} L_n(z)\, dz = \tfrac{2}{b^2+1}\big[\tfrac{b^2-1}{b^2+1}\big]^n$$

It will be observed that

- $\sum_n p_n = 1$ for all values of b;
- if $b = 1$ then $p_0 = 1$ and all other p_n vanish;
- the p_n are non-negative for all values of n if and only if $b \geqslant 1$; violation of the minimality condition would therefore place us in violation of a fundamental principle of probability theory.

Remarkably, the p_n can be described

$$p_n = \tfrac{1}{Z}e^{-\beta(n+\frac{1}{2})} \quad \text{with} \quad Z = \sum_{n=0}^\infty e^{-\beta(n+\frac{1}{2})} = \frac{e^{-\frac{1}{2}\beta}}{1-e^{-\beta}}$$

provided we use

$$e^{-\beta} = \tfrac{b^2-1}{b^2+1} \qquad ;\text{ i.e.,} \qquad b^2 = \coth\tfrac{1}{2}\beta$$

to relate β to the fatness parameter b. Interestingly, the mathematical theory of fat Gaussians (48) has—on its face—nothing to do with the quantum physics of oscillators (and even less to do with the thermodynamics of equilibrated populations of such oscillators), but if we write $\beta \equiv \hbar\omega/kT$ then the two subjects turn out to be one and the same! In short: "fat Gaussians" are *hot* Gaussians.[36]

[36] Related results—differently motivated—are developed in §§2.4 & 4.4 of Hillery *et al.*[19]

Equation (49) provides a representation of a statement of the form

$$\boldsymbol{\rho}_{\text{fat}} = \sum_n p_n \boldsymbol{\rho}_n \tag{50}$$

where the $\boldsymbol{\rho}_n$ project onto states (oscillator eigenstates) which are known to be orthonormal. So (49)↔(50) refer in fact to the *spectral* representatiuon of $\boldsymbol{\rho}_{\text{fat}}$.

Does the fat Gaussian $P(x, p; \boldsymbol{\sigma}, \boldsymbol{\lambda})$ admit of (evidently *non*-spectral) representation more in line with (47)? Indeed it does, for if (borrowing notation from (0–98))

$$g(x;\sigma) \equiv \tfrac{1}{\sigma\sqrt{2\pi}} \exp\Big\{-\tfrac{1}{2}\big[\tfrac{x}{\sigma}\big]^2\Big\}$$

then

$$\begin{aligned}\int g(x-x_0;\sigma)g(x_0;u)\,dx_0 &= g\big(x;\sqrt{\sigma^2+u^2}\,\big)\\ &= g(x;\boldsymbol{\sigma}) \quad\text{with}\quad b = \sqrt{1+(u/\sigma)^2}\end{aligned}$$

So we have

$$P(x,p;\boldsymbol{\sigma},\boldsymbol{\lambda}) = \iint P(x-x_0, p-p_0;\sigma,\lambda)\,w(x_0,p_0)\,dx_0 dp_0 \tag{51}$$

if we set

$$w(x_0,p_0) = g(x_0;u)g(p_0,v) \quad\text{with}\quad \begin{cases} u = \sigma\sqrt{b^2-1}\\ v = \lambda\sqrt{b^2-1}\end{cases}$$

Notice that $w(x_0,p_0) \to \delta(x_0)\delta(p_0)$ as $b \downarrow 1$. In (51) a fat Gaussian $P(x,p;\boldsymbol{\sigma},\boldsymbol{\lambda})$ is portrayed as a "smeared minimal Gaussian." And that the smear function has been adapted to the Gaussian we intended to smear: to achieve $\boldsymbol{\sigma}/\boldsymbol{\lambda} = \sigma/\lambda$ we had to set $u/v = \sigma/\lambda$. Had we not done so, the smeared Gaussian would have become a fat Gaussian *of altered figure*.

Husimi's and other modifications of the Weyl-Wigner transforms. We have seen that

$$\psi_{\text{gaussian}}(x) \xrightarrow[\text{Wigner}]{} P_{\text{minimal gaussian}}(x,p)$$

and have observed that the non-negativity of $P_{\text{minimal gaussian}}(x,p)$ is atypical—so atypical as to be (I am tempted to conjecture) unique.[37] Let us adopt a simplified notation $\psi_{\text{gaussian}}(x) \equiv \psi_0(x) = (x|\psi_0)$, intended in part to recall to mind the fact that $|\psi_0)$ is the groundstate of some oscillator. We learned at (31) that if $|\varphi)$ is any state orthogonal to $|\psi_0)$ then

$$|(\psi_0|\varphi)|^2 = h\iint P_{\psi_0}(x,p)P_\varphi(x,p)\,dxdp = 0 \tag{52}$$

[37] Other non-negative Wigner functions exist in abundance, but—so runs the conjecture—are in every other case representative of *mixtures*. The terms "minimal dispersion," "Gaussian" and "pure state non-negativity" would, if the conjecture were confirmed (I possess no counterexample) become synomymous.

which, since $P_{\psi_0}(x,p)$ is nowhere negative, clearly forces $P_\varphi(x,p)$ to assume occasionally negative values.[38]

Because—as I have already twice remarked[12]—I consider quantum theory to be a "profoundly strange subject, entitled to its quirks" I have always been inclined to look upon the circumstance that Wigner distributions are actually *quasi-distributions, which assume occasionally negative values* as a small price to pay for the insights provided by the Weyl-Wigner-Moyal formalism. But for some people, in some contexts, it is a price too great. Some such people[39] take the view that the formalism should simply be abandoned, others[40] the view that it stands in need of "repair."

The standard mode of repair was first described by K. Husimi,[41] and later rediscovered by (among others) N. Cartwright.[42] The basic idea[43] is elementary. The "displaced minimal Gaussian"

$$G(x-x_0,p-p_0) \equiv \tfrac{1}{\sigma\sqrt{2\pi}}\tfrac{1}{\lambda\sqrt{2\pi}}\exp\Big\{-\tfrac{1}{2}\big[\tfrac{x-x_0}{\sigma}\big]^2-\tfrac{1}{2}\big[\tfrac{p-p_0}{\lambda}\big]^2\Big\} \tag{53}$$

—regarded as a Wigner distribution on $\{x,p\}$-space—refers, for each assignment of the parameters $\{x_0,p_0\}$, to a pure state; namely, to a "launched oscillator ground state," as was seen at (45.7). Call that state $|\psi_{00})$. The important point is that there *exists* such a state (as will be the case if and only if σ and λ satisfy the minimality condition $\sigma\lambda=\frac{1}{2}\hbar$). For that fact, by (31), insures that

$$|(\psi_{00}|\psi)|^2 = h\iint G(x'-x_0,p'-p_0)P_\psi(x',p')\,dx'dp' \geqslant 0 \;\;:\;\; \text{all } x_0,\, p_0$$

---

[38] This pretty argument is Wigner's own, and is consonant with the evidence of Figure 1.

[39] See, for example, P. R. Holland, *The Quantum theory of Motion: An Account of the de Broglie-Bohm Causal Interpretation of Quantum Mechanics* (1993), §8.4.3.

[40] R. F. Fox & T. C. Elston, "Chaos and a quantum-classical correspondence in the kicked pendulum," Phys. Rev. E **49**, 3683 (1994) and "Chaos and a quantum-classical correspondence in the kicked top," Phys. Rev. E **50**, 2553 (1994).

[41] "Some formal properties of the density matrix," Proc. Physico-Math. Soc. of Japan **22**, 264 (1940). It was, by the way, Ronald Fox who, on a recent visit to Reed College, directed my attention to Husimi's work. For indication of the locus of Kôdi Husimi's thought see p. 354 in Max Jammer's *The Philosopy of Quantum Mechanics* (1974). On pp. 422–425 Jammer has things to say about the general placement of the phase space formalism.

[42] "A non-negative Wigner-type distribution," Physica **83A**, 210 (1976). Nancy Cartwright is a philosopher of science at Stanford. The concluding essay "How the measurement problem is an artifact of the mathematics" in her *How the Laws of Physics Lie* (1983) may be of some continuing interest to some readers.

[43] I follow R. F. O'Connell & E. P. Wigner, "Some properties of a non-negative quantum-mechanical distribution function," Physics Letters **83A**, 121 (1981).

Dropping the subscripts $_0$ and drawing upon an obvious symmetry of G, we are led to what might be called "Husimi's adjustment:"

$$\begin{array}{l} P_\psi(x,p) \\ \Big\downarrow \text{Husimi} \\ \boldsymbol{P}_\psi(x,p) \equiv h\displaystyle\iint G(x-x',p-p')P_\psi(x',p')\,dx'dp' \end{array} \tag{54}$$

where, though my notation does not say so, the precise meaning of G awaits assignment of a value to σ (the consequent value of λ being then determined). The right side of (54) has precisely the convolutional structure encountered already at (47). My "poor man's bold" notation is intended to suggest that $\boldsymbol{P}_\psi$ is a *smeared* companion of P_ψ; other authors use a subscripted $_S$ to that same end. For the reasons already discussed,

$$\boldsymbol{P}_\psi(x,p) \geqslant 0 \quad \text{everywhere on phase space} \tag{55.1}$$

and from an elementary property of G if follows that

$$\iint \boldsymbol{P}_\psi(x,p)\,dxdp = \iint P_\psi(x',p')\,dx'dp' = 1 \tag{55.2}$$

so $\boldsymbol{P}_\psi(x,p)$ answers to all the requirements of a *proper* probability distribution. The conditions (55) jointly insure that

$$h\iint \boldsymbol{P}_\psi(x,p)\,dxdp < 1$$

which by (34) informs us that $\boldsymbol{P}_\psi(x,p)$ is representative of a *mixture*, and by an easy line of argument we see that Husimi's $\boldsymbol{P}_\psi(x,p)$ inherits from (26) the same upper bound as limited its Wignerian precursor:

$$0 \leqslant \boldsymbol{P}_\psi(x,p) \leqslant \tfrac{2}{h} \tag{56}$$

To illustrate the effect of Husimi's adjustment we look back again to the harmonic oscillator. The Wigner distributions descriptive of the ground state and first two excited states were—in dimensionless variables—described at (25) and plotted in Figure 1. Taking the "Husimi smear function" to be just the ground state Gaussian (as seems most natural in this context) we compute

$$\left.\begin{aligned} \boldsymbol{P}_0(x,p) &= \tfrac{1}{2}\,\tfrac{2}{h}e^{-\frac{1}{4}\mathcal{E}} \\ \boldsymbol{P}_1(x,p) &= \tfrac{1}{8}\,\tfrac{2}{h}e^{-\frac{1}{4}\mathcal{E}}\cdot\mathcal{E} \\ \boldsymbol{P}_2(x,p) &= \tfrac{1}{64}\,\tfrac{2}{h}e^{-\frac{1}{4}\mathcal{E}}\cdot\mathcal{E}^2 \end{aligned}\right\} \tag{57}$$

which (recall $\mathcal{E} \equiv \varkappa^2 + \wp^2$) are manifestly non-negative, and plotted in the following figure. With the assistance of *Mathematica* we readily confirm

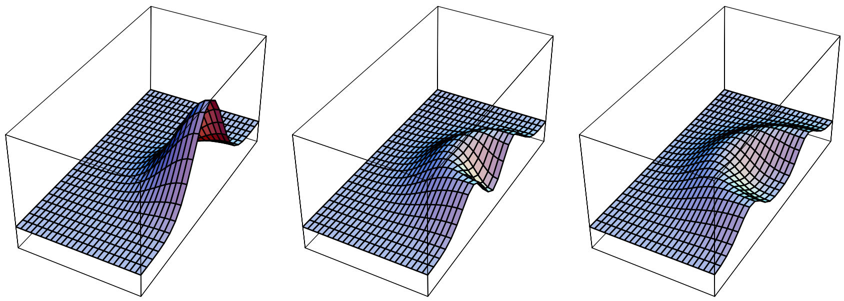

FIGURE 6: *Husimi transforms (57) of the Wigner functions (25) for the three lowest-lying energy eigenstates of a harmonic oscillator. The former "regions of negativity"—evident in Figure 1—have been extinguished.*

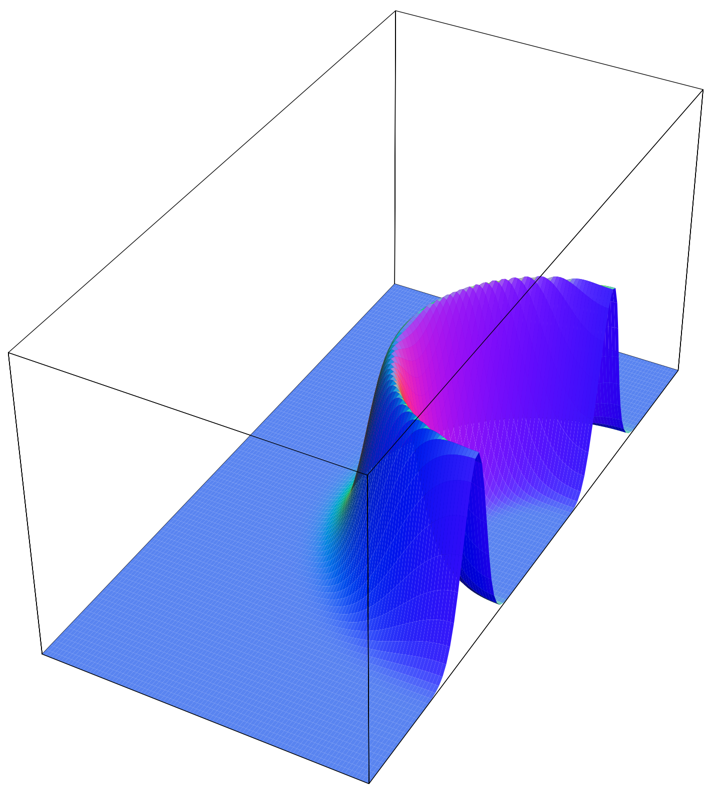

FIGURE 7: *Husimi transform of the Wigner function for the tenth state of a harmonic oscillator. The design is barely evident in the plot of* $\boldsymbol{P}_2(x,p)$*, but typical of* $\boldsymbol{P}_n(x,p)$ *for* n *large.*

(use $dxdp = \frac{h}{2}\frac{1}{2\pi}d\varkappa d\wp$) that

$$\begin{aligned}
\iint P_0(x,p)\,dxdp &= +\tfrac{1}{2\pi}\iint e^{-\frac{1}{2}\mathcal{E}}\,d\varkappa d\wp = 1\\
\iint P_1(x,p)\,dxdp &= -\tfrac{1}{2\pi}\iint e^{-\frac{1}{2}\mathcal{E}}(1-\mathcal{E})\,d\varkappa d\wp = 1\\
\iint P_2(x,p)\,dxdp &= +\tfrac{1}{2\pi}\iint e^{-\frac{1}{2}\mathcal{E}}(1-2\mathcal{E}+\tfrac{1}{2}\mathcal{E}^2\,)\,d\varkappa d\wp = 1
\end{aligned}$$

$$\begin{aligned}
h\iint P_0^2(x,p)\,dxdp &= \tfrac{1}{\pi}\iint e^{-\frac{1}{4}\mathcal{E}}\,d\varkappa d\wp = 1\\
h\iint P_1^2(x,p)\,dxdp &= \tfrac{1}{\pi}\iint e^{-\frac{1}{4}\mathcal{E}}(1-\mathcal{E})^2\,d\varkappa d\wp = 1\\
h\iint P_2^2(x,p)\,dxdp &= \tfrac{1}{\pi}\iint e^{-\frac{1}{4}\mathcal{E}}(1-2\mathcal{E}+\tfrac{1}{2}\mathcal{E}^2\,)^2\,d\varkappa d\wp = 1
\end{aligned}$$

$$\begin{aligned}
\iint \boldsymbol{P}_0(x,p)\,dxdp &= \tfrac{1}{2\pi}\iint \tfrac{1}{2}\,e^{-\frac{1}{4}\mathcal{E}}\,d\varkappa d\wp = 1\\
\iint \boldsymbol{P}_1(x,p)\,dxdp &= \tfrac{1}{2\pi}\iint \tfrac{1}{8}\,e^{-\frac{1}{4}\mathcal{E}}\cdot\mathcal{E}\,d\varkappa d\wp = 1\\
\iint \boldsymbol{P}_2(x,p)\,dxdp &= \tfrac{1}{2\pi}\iint \tfrac{1}{64}e^{-\frac{1}{4}\mathcal{E}}\cdot\mathcal{E}^2\,d\varkappa d\wp = 1
\end{aligned}$$

but

$$\left.\begin{aligned}
h\iint \boldsymbol{P}_0^2(x,p)\,dxdp &= \tfrac{1}{\pi}\iint\big[\,\tfrac{1}{2}\,e^{-\frac{1}{4}\mathcal{E}}\,\big]^2\,d\varkappa d\wp = \tfrac{1}{2} < 1\\
h\iint \boldsymbol{P}_1^2(x,p)\,dxdp &= \tfrac{1}{\pi}\iint\big[\,\tfrac{1}{8}\,e^{-\frac{1}{4}\mathcal{E}}\cdot\mathcal{E}\,\big]^2\,d\varkappa d\wp = \tfrac{1}{4} < 1\\
h\iint \boldsymbol{P}_2^2(x,p)\,dxdp &= \tfrac{1}{\pi}\iint\big[\tfrac{1}{64}e^{-\frac{1}{4}\mathcal{E}}\cdot\mathcal{E}^2\,\big]^2\,d\varkappa d\wp = \tfrac{3}{16} < 1
\end{aligned}\right\}\tag{58}$$

The latter inequalities inform us that the Husimi distributions $\boldsymbol{P}_0$, $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, ... refer to mixtures. We have already seen at (49) that the first of those can in spectral representation be displayed as a thermally weighted sum of oscillator eigenstates. I suspect that generating function techniques could be used to show that suitably altered weightings of those same Wigner functions supply representations of $\boldsymbol{P}_n$ $(n = 1, 2, \ldots)$, and the detailed argument would be both analytically interesting and physically informative ... but must save that discussion for another occasion.

With coaxing, *Mathematica* has labored heroically to obtain

$$\boldsymbol{P}_{10}(x,p) = \tfrac{2}{h}\tfrac{1}{7610145177600}e^{-\frac{1}{4}\mathcal{E}}\cdot\mathcal{E}^{10}$$

and, on the basis of this and preceding results, are led to speculate that it may be possible (and not too difficult) to show analytically that in the general case

$$\boldsymbol{P}_n(x,p) = \tfrac{2}{h}\tfrac{1}{2\cdot 4^n n!}e^{-\frac{1}{4}\mathcal{E}}\cdot\mathcal{E}^n \tag{59}$$

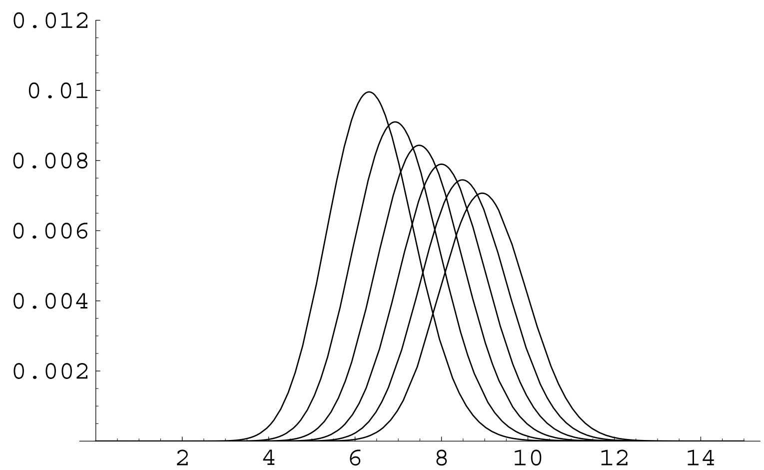

FIGURE 8: *The functions* $f_n(\varkappa) \equiv \frac{1}{2\pi}\frac{1}{2\cdot 4^n n!}e^{-\frac{1}{4}\varkappa^2}\cdot\varkappa^{2n}$, *which describe radial cross-sections of Husimi surfaces such as that shown in Figure 7, are plotted for* $n = 10, 12, 14, 16, 18, 20$. *The volume under such a curve-of-revolution can be described*

$$\int_0^\infty f_n(\varkappa)\cdot 2\pi\varkappa\,d\varkappa = 1 \quad : \quad \textit{all } n$$

since the integral is readily brought to the form of the Euler integral that defines the gamma function $\Gamma(n+1)$.

The Husimi distribution $\boldsymbol{P}_{10}(x,p)$ is plotted in Figure 7, which shows what we might describe as a "circular Gaussian ridge," centered on the $(\varkappa,\wp)$-plane. Radial sections of the n^{th} such ridge are described by the function $f_n(\varkappa)$ plotted above. From $f_n'(\varkappa) = (\text{numeric})\cdot e^{-\frac{1}{4}\varkappa^2}\cdot\varkappa^{2n-1}(-\frac{1}{2}\varkappa^2+2n) = 0$ we find that the

$$\text{radius } \varkappa_n \text{ of the } n^{\text{th}} \text{ Gaussian ridge} = \sqrt{4n}$$

which in dimensioned physical variables[13] becomes

$$x_n = \sqrt{2\hbar n/m\omega} = \text{classical amplitude of oscillator with energy } \hbar\omega n$$

In physical variables the annular ridge becomes elliptical, and peaks at the classical orbit of energy E_n.

One might anticipate that when one looks to the time-dependent theory one will see something like Gaussians *moving* along classical trajectories. But, as will emerge, that aspect of our subject provides some surprises.

Selection of a smear function (taken above to be the ground state Gaussian) is largely arbitrary. Selection of an alternative would change fine details, but leave unchanged the qualitative essentials of the results obtained above (or so I assert: the point merits closer examination).

In Chapter 1 I sketched—and in a recent seminar[44] have elaborated upon—a "theory of imperfect quantum measurement devices." The essential idea was that, whereas the devices contemplated by the standard quantum theory of measurement[45] prepare (and announce that they have prepared) pure states, imperfect devices prepare mixtures; I wrote

$$\boldsymbol{\rho}_{\text{in}} \longrightarrow \boxed{A\text{-meter announces "}a_0\text{"}} \longrightarrow \boldsymbol{\rho}_{\text{out}} = \boldsymbol{\rho}(a_0) \tag{60}$$

and proceeded from the presumption that $\boldsymbol{\rho}(a_0)$ could be described

$$\boldsymbol{\rho}(a_0) \equiv \int |a)(a| \cdot p(a_0; a)\, da$$

where in a typical case $p(a_0; a) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left[\frac{a-a_0}{\sigma}\right]^2\right\}$. I proposed, in other words, to "smudge the spectrum," to do my smearing *on the spectral line*. Husimi, however, has given us an alternative—and perhaps more natural—way to accomplish the same root objective.

Suppose we adopt Wignerian language to describe the action of an ideal device:

$$P_{\text{in}}(x,p) \longrightarrow \boxed{A\text{-meter announces "}a_0\text{"}} \longrightarrow P_{\text{out}}(x,p;a_0) \tag{60.1}$$

Here $P_{\text{out}}(x,p;a_0)$ refers to a pure state: $(x|a_0) \xrightarrow[\text{Wigner}]{} P_{\text{out}}(x,p;a_0)$.

To arrive at a modified theory of imperfect devices we have now to make only one tiny adjustment:

$$P_{\text{in}}(x,p) \longrightarrow \boxed{A\text{-meter announces "}a_0\text{"}} \longrightarrow \boldsymbol{P}_{\text{out}}(x,p;a_0) \tag{60.2}$$

where

$$P_{\text{out}}(x,p;a_0) \xrightarrow[\text{Husimi}]{} \boldsymbol{P}_{\text{out}}(x,p;a_0)$$

The smear operation is located now not on the spectral line, but *in phase space*. We have gained more secure contact with classical mechanics (which may or may not be a recommendation). And—if our experience with "fat Gaussians" can be generalized—may have positioned ourselves to speak in a natural way about a "temperature" characteristic of the noise introduced by the measurement process.

Wigner's innovation $|\psi) \longrightarrow P_\psi(x,p)$ makes it possible to contemplate assigning an "entropy" to pure quantum states

$$S[|\psi)] \equiv -\iint P_\psi(x,p) \log P_\psi(x,p)\, dxdp \tag{61.1}$$

[44] "Quantum measurement with imperfect devices" (10 February 2000).

[45] See, for example, J. Schwinger, *Quantum Kinematics & Dynamics* (1970), Chapter 1; or Kurt Gottfried, *Quantum Mechanics I: Fundamentals* (1979), Chapters IV & V.

but at some cost ... for at points where $P_\psi(x,p)$ becomes negative, $\log P_\psi(x,p)$ becomes *complex*. The proposed definition (61.1) leads to a notion of "complex entropy," and it is not at all clear that such a notion is viable.[46] But Husimi's adjustment, because it kills negativity (by diverting our attention from pure states to certain associated mixtures) permits us to contemplate a definition

$$S[|\psi)] \equiv -\iint \boldsymbol{P}_\psi(x,p)\log \boldsymbol{P}_\psi(x,p)\,dxdp \tag{61.1}$$

free from the complexity defect (if defect it be). Working from (59) to see what that revised definition has to say in the case of an oscillator, we have

$$S_n = -\iint \big[\tfrac{1}{2\pi}\tfrac{1}{2\cdot 4^n n!}e^{-\frac{1}{4}\mathcal{E}}\cdot \mathcal{E}^n\big]\log\big[\tfrac{1}{2\pi}\tfrac{1}{2\cdot 4^n n!}e^{-\frac{1}{4}\mathcal{E}}\cdot \mathcal{E}^n\big]\,d\varkappa d\wp$$

Introduce polar coordinates onto the dimensionless phase plane and obtain

$$\begin{aligned} &= \int_0^\infty \big[\tfrac{1}{2\pi}\tfrac{1}{2\cdot 4^n n!}e^{-\frac{1}{4}r^2}\cdot r^{2n}\big]\big\{\tfrac{1}{4}r^2 - 2n\log r + \log(2\pi\, 2\cdot 4^n n!)\big\}2\pi r\,dr \\ &= 1 + n + \log n! - n\text{PolyGamma}[0,n+1] + \log 4\pi \end{aligned} \tag{62.1}$$

where what *Mathematica* calls `PolyGamma[0,z]` is just the digamma function $\psi(z) \equiv \frac{d}{dz}\log\Gamma(z)$: see Chapter 44 of Spanier & Oldham.[21] Quick numerical calculation gives

$$\begin{aligned} S_0 &= 3.53102 \\ S_1 &= 4.10824 \\ S_2 &= 4.37860 \\ &\vdots \\ S_{10} &= 5.11721 \quad \sim \quad 5.16558 \\ &\vdots \\ S_{100} &= 6.25421 \quad \sim \quad 6.25919 \\ &\vdots \\ S_{1000} &= 7.40401 \quad \sim \quad 7.40451 \\ &\vdots \\ S_{10000} &= 8.55515 \quad \sim \quad 8.55520 \end{aligned}$$

while standard asymptotic expansions[47]

$$\begin{aligned} \log\Gamma(z) &\sim \log\sqrt{2\pi} - z + (z-\tfrac{1}{2})\log z + \tfrac{1}{12z} + \cdots \\ \psi(z) &\sim \log z - \tfrac{1}{2z} - \tfrac{1}{12z^2} + \cdots \end{aligned}$$

can be used to establish that as n becomes large we have

$$S_n \sim \big(\tfrac{1}{2} + \log 4\pi\sqrt{2\pi}\,\big) + \tfrac{1}{2}\log(n+1) + \tfrac{1}{6n} - \cdots \tag{62.2}$$

which was used to produce the second column of numbers. The accuracy of the

[46] I have explored the matter in §13 of "Gaussian wavepackets" (1998), with inconclusive results.

[47] See **43:6:7** and **44:6:5** in Spanier & Oldham.

approximation is impressive, but the point of deeper significance is that

$$\begin{aligned} S_n \sim \log\sqrt{n} &\sim \log\sqrt{\text{phase area enveloped by } n^{\text{th}} \text{ isoenergetic circle}} \\ &\sim \log\big\{\text{circumference of } n^{\text{th}} \text{ isoenergetic circle}\big\} \end{aligned} \tag{63}$$

What is the "deeper significance" of this result (which, by $S_n = -\langle \log \boldsymbol{P}_n \rangle$, is an almost-obvious consequence of the design of Figure 7)? In statistical mechanics one learns to associate "entropy" with "log of the density of states," and the latter concept with the "area of the isoenergetic surface."[48] In (63) we have obtained a similar result, even though the physical system has only a single degree of freedom (instead of the "very many" postulated by statistical mechanics), and we have made no evident use of the concept of "thermal equilibrium."

One should be aware that methods alternative to Husimi's method for constructing non-negative phase distributions have been devised, and found to offer advantages in certain contexts.[49] In view of the logic of the situation

HUSIMI decorated WIGNER implicit in WEYL

it is perhaps not surprising that those who first labored in this vinyard—mainly quantum opticians (who seem to have been ignorant of Husimi's work), together with a few people interested in quantum dynamical fundamentals—cultivated an interest in *alternatives to the Weyl correspondence* (7). Many of the results they achieved are interesting, but in my opinion they do not displace the historic main sequence of ideas.[50]

Dynamical motion of Wigner/Husimi distributions. We have several times had occasion to observe that the eigenkets of the Hamiltonian,[51] when launched into dynamical motion, simply sit there and buzz: $|n) \rightarrow e^{-i\omega_n t}|n)$. Harmonically and *unobservably*, since when we construct $(n|e^{+i\omega_n t}\mathbf{A}e^{-i\omega_n t}|n)$—here $\mathbf{A}$ is any observable with a steady definition—the buzz factors cancel. Cryptic quantum

---

[48] See MATHEMATICAL THERMODYNAMICS (1967), p. 72; or A. I. Khinchin, *Mathematical Foundations of Statistical Mechanics* (1949), pp. 33–35.

[49] See Y. Kano, "A new phase-space distribution function in the statistical theory of the electromagnetic field," J. Math. Phys. **6**, 1913 (1965). Wigner[18] refers also to some others.

[50] See R. J. Glauber, "Coherent & incoherent states of the radiation field," Phys. Rev. **131**, 2766 (1963); L. Cohen, "Generalized phase-space distribution functions," J. Math. Phys. **7**, 781 (1966); G. S. Agarwal & E. Wolf, "Calculus for functions of noncommutative operators and general phase-space methods in quantum mechanics. I. Mapping theorems and ordering of functions of noncommuting operators," Phys. Rev. D **2**, 2162 (1970) and "... II. Quantum methods in phase space," Phys. Rev. D **2**, 2187 (1970). Also S. S. Schweber, "On Feynman quantization," J. Math. Phys. **3**, 831 (1962). All of those papers provide elaborate bibliographies, but none contains a reference to Husimi.

[51] We assume the Hamiltonian to be time-independent, and work in the Schrödinger picture.

motion becomes manifest quantum motion only when $|\psi)_0$ has been assembled by *superposition of at least two* eigenkets:

$$|\psi)_0 = c_1|n_1)_0 + c_2|n_2)_0 \quad \text{with} \quad |c_1|^2 + |c_2|^2 = 1$$

Then $(\psi|\mathbf{A}|\psi)$ dithers; i.e., it exhibits a t-dependence of the form

$$(\psi|\mathbf{A}|\psi) = a + b\cos(\omega t + \delta) \quad \text{with} \quad \omega \equiv (E_2 - E_1)/\hbar$$

and will execute more complicated motion if more than two eigenkets are included in the superposition. In the interest of brevity, I illustrate this and subsequent points by looking again to the harmonic oscillator, reserving more general commentary for another occasion; I work in the familiar dimensionless variables,[13] and entrust all the labor to *Mathematica*.[51]

Look to the oscillator state $|\psi) = \frac{1}{\sqrt{2}}|2) + \frac{1}{\sqrt{2}}|3)$, which in $\varkappa$-representation at time t becomes

$$\psi(\varkappa, t) = \tfrac{1}{\sqrt{2}}\Psi[2,\varkappa]e^{-i\omega[2+\frac{1}{2}]t} + \tfrac{1}{\sqrt{2}}\Psi[3,\varkappa]e^{-i\omega[3+\frac{1}{2}]t} \tag{64}$$

giving

$$|\psi(\varkappa,\theta)|^2 = \tfrac{1}{2}\Psi[2,\varkappa]\Psi[2,\varkappa] + \Psi[2,\varkappa]\Psi[3,\varkappa]\cos\theta + \tfrac{1}{2}\Psi[3,\varkappa]\Psi[3,\varkappa] \tag{65}$$

where $\theta \equiv \omega t$ signifies "dimensionless time" (phase). The intent of (65) is illustrated below, but animation[52] makes a much more vivid statement; frames

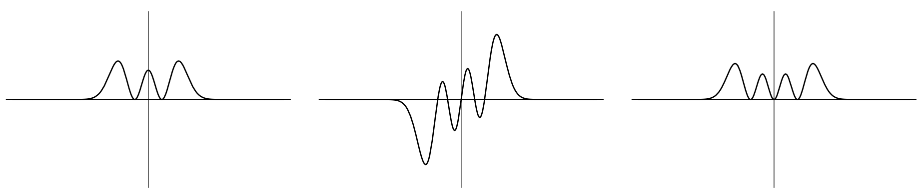

FIGURE 9: *Equation (65) asks us to add the outer functions to a* $\cos\theta$*-modulated copy of the central function.*

---

[51] My reader is encouraged to do computation in parallel with the text, so as to be in position to run the animations in which it will culminate. To that end, define `He[n_,x_]:=(1/`$\sqrt{2^n}$`)HermiteH[n,x/`$\sqrt{2}$`]` and

$$\texttt{Ψ[n\_,x\_]:=(1/}\sqrt{\texttt{n!}\sqrt{2\pi}}\texttt{)Exp[-}\tfrac{1}{4}\texttt{x}^2\texttt{]He[n,x]}$$

Verify that $\int \Psi[m,x]\Psi[n,x]\,dx = \delta_{mn}$. Further instruction will be provided as we proceed.

[52] Use `Table[Plot[`$\frac{1}{2}$`Ψ[2,x]`^{2}`+Ψ[2,x]Ψ[3,x]Cos[`$\frac{2\pi}{20}$`n]+`$\frac{1}{2}$`Ψ[3,x]`^{2}`, {x,-10,10}, PlotRange->{ 0,0.5}],{n,0,19}]` to construct a set of twenty figures. Select those, and from the Cell Menu select Animate Selected Graphics.

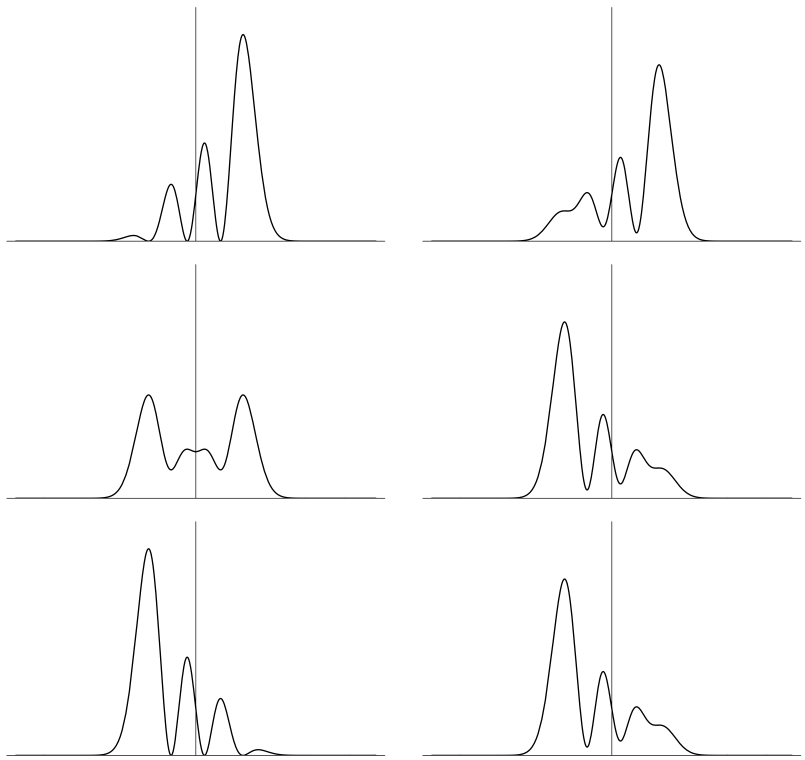

FIGURE 10: *Frames from an animation of (65). The figure is to be scanned like the page of a book. Successive frames correspond to $\theta = n\frac{\pi}{4}$ with $n = 0, 1, 2, 3, 4, 5$. The function sloshes from right to left, and starts back again.*

from such a filmstrip are presented above.

Look now to what Wigner has to say about such a situation. In our familiar dimensionless variables we have

$$P_\psi(\varkappa, \wp) = \tfrac{1}{2\pi}\int \psi^*(\varkappa + \xi)e^{i\,\wp\,\xi}\psi(\varkappa - \xi)\,d\xi$$

which, when we take ψ to be given by (64), becomes

$$P_\psi(\varkappa, \wp; \theta) = P_{22}(\varkappa, \wp) + \big\{P_{23}(\varkappa, \wp; \theta) + \text{conjugate}\big\} + P_{33}(\varkappa, \wp)$$

where, according to *Mathematica* (compare (25)),

$$\begin{aligned} P_{22}(\varkappa, \wp) &\equiv \tfrac{1}{2\pi}\int \tfrac{1}{\sqrt{2}}\Psi[2, \varkappa + \xi]e^{i\,\wp\,\xi}\,\tfrac{1}{\sqrt{2}}\Psi[2, \varkappa - \xi]\,d\xi \\ &= +\tfrac{1}{2}\tfrac{1}{2\pi}e^{-\frac{1}{2}\mathcal{E}}(1 - 2\mathcal{E} + \tfrac{1}{2}\mathcal{E}^2) \end{aligned}$$

$$\begin{aligned}
P_{33}(\varkappa,\wp) &\equiv \tfrac{1}{2\pi}\int \tfrac{1}{\sqrt{2}}\Psi[3,\varkappa+\xi]e^{i\wp\xi}\,\tfrac{1}{\sqrt{2}}\Psi[3,\varkappa-\xi]\,d\xi\\
&= -\tfrac{1}{2}\tfrac{1}{2\pi}e^{-\frac{1}{2}\mathcal{E}}(1-3\mathcal{E}+\tfrac{3}{2}\mathcal{E}^2-\tfrac{1}{6}\mathcal{E}^3)\\
P_{23}(\varkappa,\wp;\theta) &\equiv e^{-i\theta}\cdot\tfrac{1}{2\pi}\int \tfrac{1}{\sqrt{2}}\Psi[2,\varkappa+\xi]e^{i\wp\xi}\,\tfrac{1}{\sqrt{2}}\Psi[3,\varkappa-\xi]\,d\xi\\
&= e^{-i\theta}\cdot\tfrac{1}{2}\tfrac{1}{2\pi}\tfrac{6}{\sqrt{3}}(\varkappa-i\wp)(1-\mathcal{E}+\tfrac{1}{6}\mathcal{E}^2)
\end{aligned}$$

Assembling these results, we obtain

$$\begin{aligned}
P_\psi(\varkappa,\wp;\theta) = \tfrac{1}{2}\tfrac{1}{2\pi}e^{-\frac{1}{2}\mathcal{E}}\Big\{&(\mathcal{E}-\mathcal{E}^2+\tfrac{1}{6}\mathcal{E}^3)\\
&+\tfrac{12}{\sqrt{3}}(1-\mathcal{E}+\tfrac{1}{6}\mathcal{E}^2)(\varkappa\cos\theta-\wp\sin\theta)\Big\}
\end{aligned}\tag{66}$$

and verify that indeed $\iint P_\psi(\varkappa,\wp;\theta)\,d\varkappa d\wp = 1$. Note that the θ-dependent terms are *odd* functions of $\varkappa$ and $\wp$, so can make no net contribution to the integral.

Animation is the *only* way to grasp the story that (66) is trying to tell ... but a memory hog. In Figure 11 we see six frames (again: $\theta = n\frac{\pi}{4}$ with $n = 0,1,2,3,4,5$) taken from such a filmstrip. Note that the central asymmetry of the surface, and its steady rotation, have appeared spontaneously: we have done nothing so artificial as to "displace a copy of the groundstate Gaussian, release it, and watch it slosh back and forth."[53] Though we are deep within the quantum realm, the phase space formalism has exposed a motion strongly suggestive of classical oscillator motion ... but this, I think, is deceptive—an artifact of the circumstance that $|\psi)$ was constructed by superposition of *only two* energy eigenkets, and therefore contains only a single effective frequency.

Now take $P_0(\varkappa,\wp) = \frac{1}{2\pi}e^{-\frac{1}{2}\mathcal{E}}$ and look to the Husimi transform of (66). *Mathematica*, with a little gentle coaxing, supplies

$$\begin{aligned}
\boldsymbol{P}_\psi(\varkappa,\wp;\theta) &= \iint P_0(\varkappa-\varkappa',\wp-\wp')P_\psi(\varkappa',\wp')\,d\varkappa' d\wp'\\
&= \tfrac{1}{3072\pi}e^{-\frac{1}{2}\mathcal{E}}\cdot\mathcal{E}^2\big(12+\mathcal{E}+4\sqrt{3}(\varkappa\cos\theta-\wp\sin\varkappa)\big)
\end{aligned}\tag{67}$$

Again we compute $\iint \boldsymbol{P}_\psi(\varkappa,\wp;\theta)\,d\varkappa d\wp = 1$, which provides a weak check on the accuracy of the θ-independent terms in (67), but I have nothing sharp to say about the significance of the numerator: $3072 = 2^{10}\cdot 3$. Frames from the animation of (67) are displayed in Figure 12.

---

[53] See, for example, L. I. Schiff, *Quantum Mechanics* (3rd edition 1968) p. 74 or QUANTUM MECHANICS (1967), Chapter 2, pp. 89–93. This topic was first explored by Schrödinger himself, in 1926. An English translation of that paper, bearing the title "The continuous transition from micro- to macro-mechanics" can be found in his *Collected Papers on Wave Mechanics* (3rd English edition 1968).

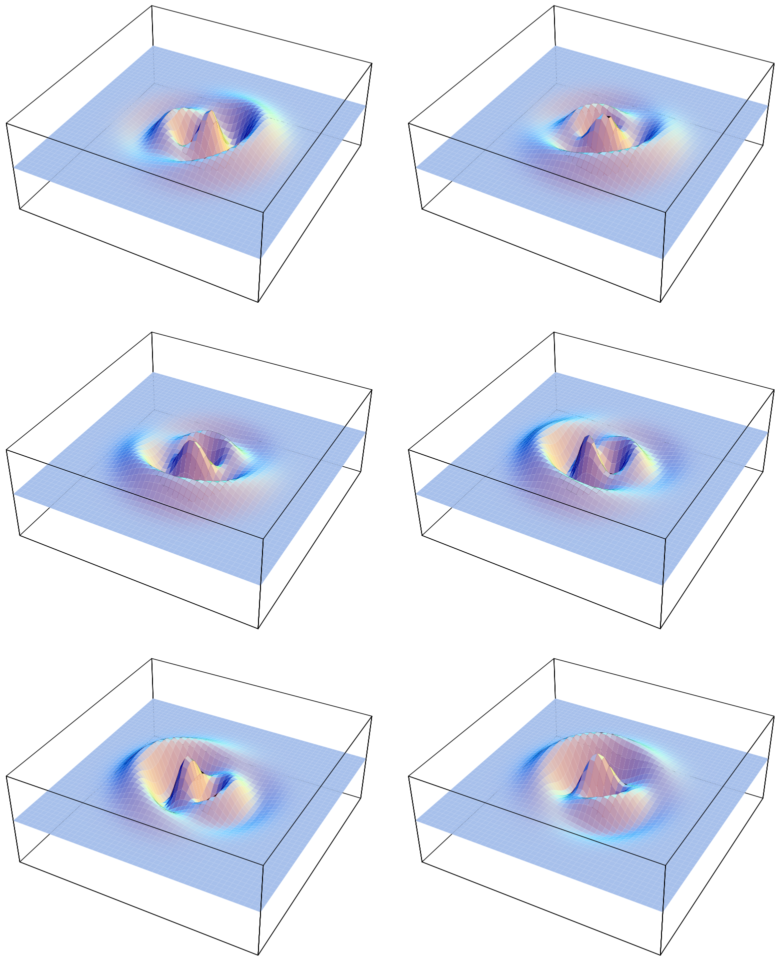

FIGURE 11: *Frames from an animation of (66), showing the Wigner transforms of the progressively evolved wave functions whose moduli are shown in Figure 10. The surface appears to spin clockwise about the origin.*

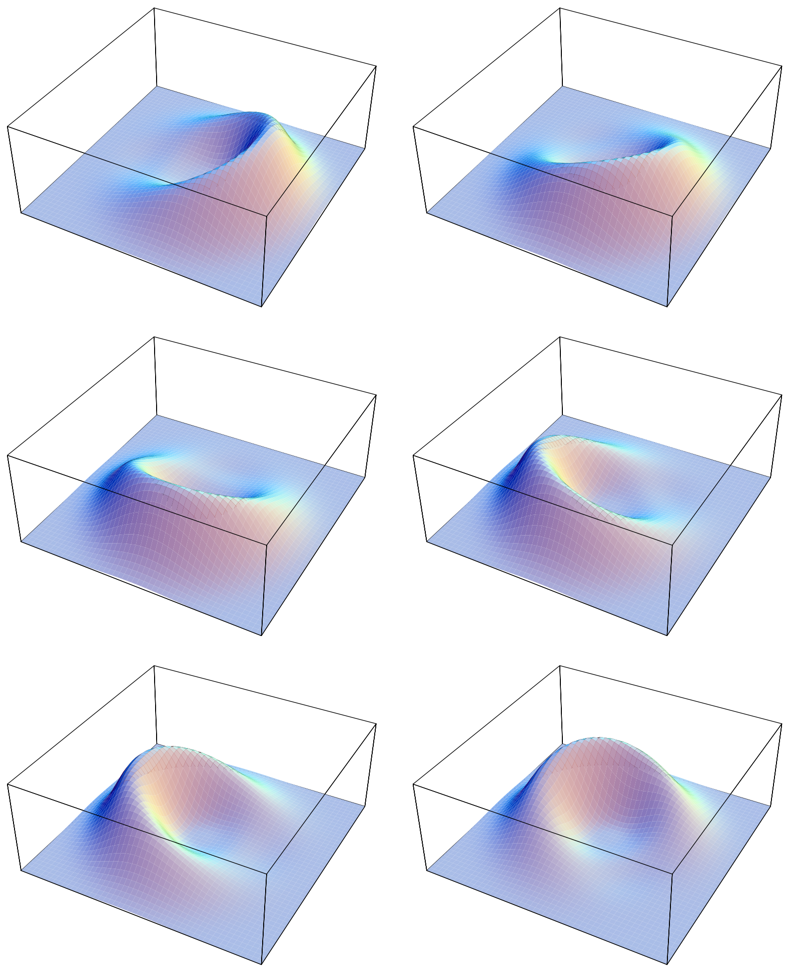

FIGURE 12: *Husimi transforms of the Wigner functions shown in the preceding figure. The progressive rotation is now more strikingly apparent. That the Husimi distribution is everywhere non-negative is made more vividly evident in the following figure.*

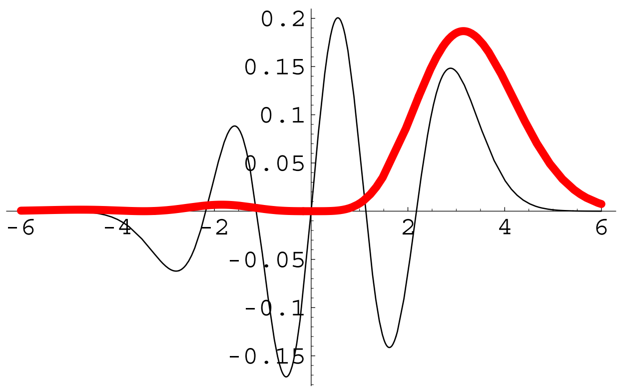

FIGURE 13: *Superimposed plots of the central cross-sections* ($\wp = 0$) *of the* $\theta = 0$ *Wigner/Husimi functions shown at upper left in Figures 11/12. The Husimi function—here multiplied by five to enhance clarity, and shown in red—never becomes negative, but does fall to zero at* $\varkappa = 0$ *and* $\varkappa = -2\sqrt{3} = -3.4641$.

It would be interesting and useful to examine other cases, other systems in order to gain a sharper sense of how generally applicable are the insights to which the oscillator has led us . . . but that is an undertaking which I must, for the moment, be content to reserve for an occasion when I have more time *and more computer memory* at my disposal. But I cannot in good conscience move on without remarking that in one important respect the physics to which Figure 12 alludes is a swindle. For the process we *have* been talking about might be diagrammed

$$\begin{array}{ccc} P(x,p;0) & \xrightarrow[\text{dynamical evolution by (38)}]{} & P(x,p;t) \\ & & \downarrow \\ & & \boldsymbol{P}(x,p;t) \end{array}$$

while the process we *should have* been talking about flows

$$\begin{array}{ccc} P(x,p;0) & & \\ \downarrow & & \\ \boldsymbol{P}(x,p;t) & \xrightarrow[\text{dynamical evolution by (38)}]{} & \boldsymbol{P}(x,p;t) \end{array}$$

The distinction is profound, and was first explored by O'Connell & Wigner.[19] It would, however, take me too far afield to pursue the matter here.

Quantum mechanics as a theory of interactive moments. Folklore alleges, and in some texts[54] it is explicitly—if, as will emerge, not quite correctly—asserted, that "quantum mechanical expectation values obey Newton's second law." The pretty point here at issue was first remarked by Paul Ehrenfest (1880–1933), in a paper scarcely more than two pages long.[55] Concerning the substance and impact of that little gem, Max Jammer, at p. 363 in his *The Conceptual Development of Quantum Mechanics* (1966), has this to say:

> *"That for the harmonic oscillator wave mechanics agrees with ordinary mechanics had already been shown by Schrödinger*[53]*...*
> *A more general and direct line of connection between quantum mechanics and Newtonian mechanics was established in 1927 by Ehrenfest, who showed 'by a short elementary calculation without approximations' that the expectation value of the time derivative of the momentum is equal to the expectation value of the negative gradient of the potential energy function. Ehrenfest's affirmation of Newton's second law in the sense of averages taken over the wave packet had a great appeal to many physicists and did much to further the acceptance of the theory. For it made it possible to describe the particle by a localized wave packet which, though eventually spreading out in space, follows the trajectory of the classical motion. ... Ehrenfest's theorem and its generalizations by Ruark*[56]*... do not conceptually reduce quantum dynamics to Newtonian physics. They merely establish an analogy—though a remarkable one in view of the fact that, owing to the absence of a superposition principle in classical mechanics, quantum mechanics and classical dynamics are built on fundamentally different foundations."*

The basic idea is elementary. Recall that the dynamical motion of $\langle \mathsf{A} \rangle \equiv (\psi|\mathsf{A}|\psi)$ can be described $\frac{d}{dt}\langle \mathsf{A}\rangle = \frac{1}{i\hbar}\langle \mathsf{AH} - \mathsf{HA}\rangle$ (which is, as it happens, picture-independent), so we can in particular write

$$\begin{aligned}\tfrac{d}{dt}\langle \mathsf{x}\rangle &= \tfrac{1}{i\hbar}\langle[\mathsf{x}\,,\mathsf{H}]\rangle\\ \tfrac{d}{dt}\langle \mathsf{p}\rangle &= \tfrac{1}{i\hbar}\langle[\mathsf{p}\,,\mathsf{H}]\rangle\end{aligned}$$

But if $\mathsf{H} = {}_{\mathsf{p}}\big[H_{px}(x,p)\big]_{\mathsf{x}} = \sum(\text{terms of the form } \mathsf{p}^m\mathsf{x}^n)$ then

$$\begin{aligned}[\mathsf{x}\,,\mathsf{H}] &= \sum(\text{terms of the form } +i\hbar\cdot m\,\mathsf{p}^{m-1}\mathsf{x}^n) = +i\hbar\,{}_{\mathsf{p}}\big[\tfrac{\partial H_{px}}{\partial p}\big]_{\mathsf{x}}\\ [\mathsf{p}\,,\mathsf{H}] &= \sum(\text{terms of the form } -i\hbar\cdot n\,\mathsf{p}^{m}\mathsf{x}^{n-1}) = -i\hbar\,{}_{\mathsf{p}}\big[\tfrac{\partial H_{px}}{\partial x}\big]_{\mathsf{x}}\end{aligned}$$

[54] See, for example, Griffiths, p. 17, Problem 1.12.

[55] "Bemerkung über die angenäherte Gültigkeit der klassichen Machanik innerhalb der Quanatenmechanik," Z. Physik **45**, 455–457 (1927).

[56] The allusion here is to A. E. Ruark, "... the force equation and the virial theorem in wave mechanics," Phys. Rev. **31**, 533 (1928).

Alternatively/equivalently (and with only the superficial appearance of bias in favor of the Weyl correspondence) we might write

$$\mathsf{H} \xleftarrow[\text{Weyl}]{} H(x,p)$$

and use (17.2) to obtain

$$\begin{aligned}[\mathsf{x}, \mathsf{H}] &\xleftarrow[\text{Weyl}]{} i\hbar\,[x, H] = +i\hbar\tfrac{\partial H}{\partial p}\\ [\mathsf{p}, \mathsf{H}] &\xleftarrow[\text{Weyl}]{} i\hbar\,[p, H] = -i\hbar\tfrac{\partial H}{\partial x}\end{aligned}$$

Or, in service of notational simplicity, we might assume the Hamiltonian to have the specific form $\mathsf{H} = \frac{1}{2m}\mathsf{p}^2 + U(\mathsf{x})$; it is then unarguable that

$$\begin{aligned}[\mathsf{x}, \mathsf{H}] &= +i\hbar\,\tfrac{1}{m}\,\mathsf{p}\\ [\mathsf{p}, \mathsf{H}] &= -i\hbar\,U'(\mathsf{x})\end{aligned}$$

giving

$$\left.\begin{aligned}\tfrac{d}{dt}\langle\mathsf{x}\rangle &= \tfrac{1}{m}\langle\,\mathsf{p}\,\rangle\\ \tfrac{d}{dt}\langle\mathsf{p}\rangle &= -\langle U'(\mathsf{x})\rangle\end{aligned}\right\} \tag{68}$$

It is to the latter class of systems that I shall—for the moment, as a matter of expository convenience—confine my specific remarks.

The first moments $\langle\mathsf{x}\rangle$ and $\langle\mathsf{p}\rangle$ will move "classically" only under those special circumstances which permit one to write $\langle U'(\mathsf{x})\rangle = U'(\langle\mathsf{x}\rangle)$. In classical statistical mechanics this would be achieved if $P(x,p)$ referred in fact to a "Dirac spike" in classical motion, but in quantum mechanics that degree of localization is disallowed (would be in conflict with the Heisenberg uncertainty principle). We are forced, therefore, to require that $U'(x)$ depend at most linearly on x:

$$U' = ax + b \quad : \quad \text{entails} \quad U(x) = \tfrac{1}{2}ax^2 + bx + c$$

This class of special cases includes the free particle, the particle in free fall and the harmonic oscillator; we are brought back, in short, to the famously tractable class of systems in which $H(x,p)$ *depends at most quadratically* on x and p. Let us look in this light to the

MOMENT THEORY OF THE HARMONIC OSCILLATOR Given the Hamiltonian $\mathsf{H} = \frac{1}{2m}\mathsf{p}^2 + \frac{1}{2}m\omega^2\mathsf{x}^2$ we look to the motion of $\langle\mathsf{p}\rangle$

$$\tfrac{d}{dt}\langle\mathsf{p}\rangle = \tfrac{1}{i\hbar}\langle[\mathsf{p}, \mathsf{H}]\rangle = -m\omega^2\langle\mathsf{x}\rangle \tag{69.1}$$

and find ourselves forced to look also to the motion of $\langle\mathsf{x}\rangle$:

$$\tfrac{d}{dt}\langle\mathsf{x}\rangle = \tfrac{1}{i\hbar}\langle[\mathsf{x}, \mathsf{H}]\rangle = \tfrac{1}{m}\langle\mathsf{p}\rangle \tag{69.2}$$

In (69) we have a closed system—a coupled pair of 1st-order differential equations, which conjointly entail $\frac{d^2}{dt^2}\langle \mathsf{x}\rangle_t + \omega^2\langle \mathsf{x}\rangle_t = 0$, giving (if we start our clock at the appropriate instant)

$$\langle \mathsf{x}\rangle_t = \langle \mathsf{x}\rangle_{\max}\cos\omega t \tag{69.3}$$

The remarkable implication is that for all pure states $|\psi)$—as, indeed, for all mixtures $\boldsymbol{\rho}$—the first moment $\langle \mathsf{x}\rangle_t$ *oscillates with angular frequency* ω (see again Figure 10), as also does $\langle \mathsf{p}\rangle_t$: returning with (69.3) to (69.2) we have

$$\begin{aligned}\langle \mathsf{p}\rangle_t &= -\langle \mathsf{p}\rangle_{\max}\cos\omega t \\ \langle \mathsf{p}\rangle_{\max} &= m\omega\,\langle \mathsf{x}\rangle_{\max}\end{aligned} \tag{69.4}$$

For energy eigenstates (also for mixtures of such states, but not for superpositions) the t-dependence must drop away; this, by the design of (69.3/4), entails that

$$\langle \mathsf{x}\rangle_{\text{any eigenstate}} = \langle \mathsf{p}\rangle_{\text{any eigenstate}} = 0 \tag{69.5}$$

which is a statement commonly attributed to parity properties of the oscillator eigenstates, but which has here been obtained without solving any differential equations, without exhibiting/examining the explicit descriptions of any eigenfunctions.

Look now to the motion of $\langle \mathsf{p}^2\rangle$ (which must not be confused with $\langle \mathsf{p}\rangle^2$). From

$$\tfrac{d}{dt}\langle \mathsf{p}^2\rangle = \tfrac{1}{i\hbar}\langle[\mathsf{p}^2, \mathsf{H}]\rangle = -\omega^2\langle \tfrac{1}{2}(\mathsf{x}\,\mathsf{p} + \mathsf{p}\,\mathsf{x})\rangle \tag{70.1}$$

we acquire an unanticipated interest in the motion of $\langle \tfrac{1}{2}(\mathsf{x}\,\mathsf{p} + \mathsf{p}\,\mathsf{x})\rangle$

$$\tfrac{d}{dt}\langle \tfrac{1}{2}(\mathsf{x}\,\mathsf{p} + \mathsf{p}\,\mathsf{x})\rangle = \tfrac{1}{i\hbar}\langle[\tfrac{1}{2}(\mathsf{x}\,\mathsf{p} + \mathsf{p}\,\mathsf{x}), \mathsf{H}]\rangle = \langle \tfrac{1}{m}\mathsf{p}^2 - m\omega^2\mathsf{x}^2\rangle \tag{70.2}$$

whence in the motion also of $\langle(\tfrac{1}{m}\mathsf{p}^2 - m\omega^2\mathsf{x}^2)\rangle$:

$$\tfrac{d}{dt}\langle(\tfrac{1}{m}\mathsf{p}^2 - m\omega^2\mathsf{x}^2)\rangle = \tfrac{1}{i\hbar}\langle[(\tfrac{1}{m}\mathsf{p}^2 - m\omega^2\mathsf{x}^2), \mathsf{H}]\rangle = -4\omega^2\langle \tfrac{1}{2}(\mathsf{x}\,\mathsf{p} + \mathsf{p}\,\mathsf{x})\rangle \tag{70.3}$$

$$\tfrac{d}{dt}\langle(\tfrac{1}{m}\mathsf{p}^2 + m\omega^2\mathsf{x}^2)\rangle = 0 \quad \text{by energy conservation} \tag{70.4}$$

But here the forced additions to our list of "operators of interest" stop, for were we to continue the procedure we would be led to the operators

$$\mathsf{C} \equiv \tfrac{1}{2}(\mathsf{x}\,\mathsf{p} + \mathsf{p}\,\mathsf{x}) \quad\text{and}\quad \mathsf{D} \equiv \tfrac{1}{m}\mathsf{p}^2 - m\omega^2\mathsf{x}^2 \tag{70.5}$$

<u>in alternating sequence</u> (the procedure has, in other words, terminated), and would have been led promptly to that same sequence had we started from $\frac{d}{dt}\langle \mathsf{x}^2\rangle$. We conclude from the design of (70.2/3) that in both cases the associated expectation values satisfy an equation of the form

$$\tfrac{d^2}{dt^2}\langle \mathsf{A}\rangle_t + 4\omega^2\langle \mathsf{A}\rangle_t = 0 \quad\Rightarrow\quad \text{oscillation with \underline{doubled frequency}}$$

Let us agree to write

$$\langle \mathsf{C} \rangle_t = \quad C\sin(2\omega t + \delta) \tag{70.6}$$

Then (70.2) supplies

$$\langle \mathsf{D} \rangle_t = 2\omega C\cos(2\omega t + \delta) \tag{70.7}$$

while by (70.1)

$$\langle \mathsf{p}^2 \rangle_t = P^2 + \quad \tfrac{1}{2}\omega C\cos(2\omega t + \delta) \tag{70.8}$$

where $P^2 \geqslant \frac{1}{2}\omega C$ is a constant of integration. Similarly

$$\langle \mathsf{x}^2 \rangle_t = X^2 - \tfrac{1}{m^2\omega^2}\tfrac{1}{2}\omega C\cos(2\omega t + \delta) \tag{70.9}$$

The integral of (70.4) serves to interrelate those second moments:

$$\tfrac{1}{2m}\langle \mathsf{p}^2 \rangle_t + \tfrac{1}{2}m\omega^2\langle \mathsf{x}^2 \rangle_t = \tfrac{1}{2m}P^2 + \tfrac{1}{2}m\omega^2X^2 = E \tag{70.10}$$

To summarize: simple coupled differential equations serve *universally* (i.e., without reference to the specific pure/mixed state of the quantum system) to describe the *motion of the moments*, and equalities/inequalities serve to relate/constrain the constants which appear in the solutions of those equations. In the example just studied, we found

- the motion of 1st moments to be ...
- decoupled from the (frequency doubled) motion of 2nd moments, which is
- decoupled from the (frequency trebled) motion of 3rd moments ...

and, moreover, that all moments move classically (the differential equations contain no exposed $\hbar$-factors). But the former circumstance is *special to the oscillator*, and the latter *special to systems with quadratic Hamiltonians*. To gain a glimpse of the more typical situation, consider the system

$$\mathsf{H} = \tfrac{1}{2m}\mathsf{p}^2 + \tfrac{1}{4}k\mathsf{x}^4 \tag{71.1}$$

The classical equations of motion read

$$\begin{aligned}\dot{p} &= -kx^3\\ \dot{x} &= \tfrac{1}{m}p\end{aligned}$$

while Ehrenfest's theorem (68) supplies

$$\left.\begin{aligned}\tfrac{d}{dt}\langle \mathsf{p}\rangle &= -k\langle \mathsf{x}^3\rangle\\ \tfrac{d}{dt}\langle \mathsf{x}\rangle &= \tfrac{1}{m}\langle \mathsf{p}\rangle\end{aligned}\right\} \tag{71.2}$$

from which we acquire an obligation to study the motion of $\langle \mathsf{x}^3\rangle$. Tedious computation gives

$$\begin{aligned}\tfrac{d}{dt}\langle \mathsf{x}^3\rangle &= \tfrac{3}{2m}\langle(\mathsf{x}^2\,\mathsf{p} + \mathsf{p}\,\mathsf{x}^2)\rangle\\ \tfrac{d}{dt}\langle(\mathsf{x}^2\,\mathsf{p} + \mathsf{p}\,\mathsf{x}^2)\rangle &= \tfrac{1}{2m}\langle(\mathsf{x}^2\,\mathsf{p} + 2\mathsf{x}\,\mathsf{p}\,\mathsf{x} + \mathsf{p}\,\mathsf{x}^2)\rangle - \tfrac{8}{3}k\langle \mathsf{x}^5\rangle\\ &\;\;\vdots\end{aligned}$$

Pretty clearly (since H introduces factors faster that $[\mathsf{x},\mathsf{p}] = i\hbar\mathbf{1}$ can kill them), these equations comprise only the leading members of an *infinite system of coupled first-order linear differential equations*. Other such systems interrelate such moments as are absent from the preceding system. The design of the complete set of systems is latent in the design of H; state data is written into the *initial values* of those moments. The intricate relationships which serve to distinguish possible initial values from impossible . . . are, for the most part, H-dependent, and resist general description.[57] But one class of universal constraints was identified by Schrödinger (1930), and is of invariably fundamental importance.[58] Let $(\Delta A)^2 \equiv \langle(\mathsf{A} - \langle\mathsf{A}\rangle)^2\rangle$. Then—in consequence ultimately of the Schwarz inequality—one for *all* observables A and B has

$$(\Delta A)^2(\Delta B)^2 \geq \Big\langle\frac{\mathsf{AB}-\mathsf{BA}}{2i}\Big\rangle^2 + \Big\{\Big\langle\frac{\mathsf{AB}+\mathsf{BA}}{2}\Big\rangle - \langle\mathsf{A}\rangle\langle\mathsf{B}\rangle\Big\}^2 \tag{72.1}$$

which in a particular case ($\mathsf{A}\mapsto\mathsf{x}$, $\mathsf{B}\mapsto\mathsf{p}$) entails

$$(\Delta x)^2(\Delta p)^2 \geq (\hbar/2)^2 + \Big\{\Big\langle\frac{\mathsf{xp}+\mathsf{px}}{2}\Big\rangle - \langle\mathsf{x}\rangle\langle\mathsf{p}\rangle\Big\}^2 \tag{72.2}$$

Schrödinger's inequality (72.1) provides a sharpened/generalized statement of the Heisenberg uncertainty principle.

At (72.2) we encounter once again the observable $\mathsf{C} \equiv \frac{1}{2}(\mathsf{x\,p} + \mathsf{p\,x})$ which at (70.5) was recommended to our attention by the physics of an oscillator, and which we first met when we had occasion to observe that

$$xp \xrightarrow[\text{Weyl}]{} \tfrac{1}{2}(\mathsf{x\,p} + \mathsf{p\,x})$$

Its most recent occurance can be understood as follows: Let $P(x,p)$ be a bi-variate distribution function in a pair of real random variables, x and p. One has $\langle x^m p^n\rangle = \iint x^m p^n P(x,p)\,dxdp$. If x and p were *independent* random variables, then $P(x,p)$ would factor—$P(x,p) = f(x)\cdot g(p)$—and we would have $\langle x^m p^n\rangle = \langle x^m\rangle\langle p^n\rangle$. The constructions $C_{m,n} \equiv \langle x^m p^n\rangle - \langle x^m\rangle\langle p^n\rangle$ serve (when they fail to vanish) to provide quantitative indication of the degree to which x and p *fail* to be statistically independent; i.e., of their "correlation." The leading such indicator is the

$$\text{“correlation coefficient”} \equiv \langle(x - \langle x\rangle)(p - \langle p\rangle)\rangle = \big\{\langle xp\rangle - \langle x\rangle\langle p\rangle\big\}$$

On the right side of (72.2) we have encountered the obvious *quantum analog* of that construction.

[57] Their description in particular cases is, however, worth the effort. For eigenstates are states that bring all moments to rest. The search for eigenstates is, therefore, equivalent to a search for *fixed points in the moment problem*.

[58] See "Remarks concerning the status & some ramifications of Ehrenfest's theorem" (1998), p. 7 for the proof and some historical references.

Continuing the preceding (pre-quantum mechanical) discussion ... group the mixed moments in ascending order

$$\begin{array}{c} 1 \\ \langle x\rangle \quad \langle p\rangle \\ \langle x^2\rangle \quad \langle xp\rangle \quad \langle p^2\rangle \\ \langle x^3\rangle \quad \langle x^2p\rangle \quad \langle xp^2\rangle \quad \langle p^3\rangle \\ \vdots \end{array}$$

and suppose all those moments to be *known*; one is then in position to construct

$$\begin{aligned} M(\alpha,\beta) &\equiv \sum_{k=0}^{\infty} \tfrac{1}{k!}\big(\tfrac{i}{\hbar}\big)^k \Big\{\sum_{n=0}^{k} \tbinom{k}{n}\big\langle x^n p^{k-n}\big\rangle \beta^n\alpha^{k-n}\Big\} \\ &= \sum_{k=0}^{\infty} \tfrac{1}{k!}\big(\tfrac{i}{\hbar}\big)^k \big\langle(\alpha p+\beta x)^k\big\rangle \\ &= \big\langle e^{\frac{i}{\hbar}(\alpha p+\beta x)}\big\rangle \\ &= \iint e^{\frac{i}{\hbar}(\alpha p+\beta x)} P(x,p)\,dxdp \end{aligned} \tag{73.1}$$

Immediately

$$P(x,p) = \tfrac{1}{h^2}\iint \underbrace{\big\langle e^{\frac{i}{\hbar}(\alpha p+\beta x)}\big\rangle}_{\text{moment data } \langle x^m p^n\rangle \text{ resides here}} e^{-\frac{i}{\hbar}(\alpha p+\beta x)}\,dqdy \tag{73.2}$$

so $M(\alpha,\beta)$—the *moment generating function* or "characteristic function," as it is called—is simply the Fourier transform of the distribution function.[59]

The idea now is to put that train of thought to quantum mechanical work. Immediately we are led—were, in effect, led already at (21)—to write

$$P_\psi(x,p) = \tfrac{1}{h^2}\iint M_\psi(\alpha,\beta)e^{-\frac{i}{\hbar}(\alpha p+\beta x)}\,d\alpha d\beta \tag{74}$$

$$M_\psi(\alpha,\beta) \equiv (\psi|e^{\frac{i}{\hbar}(\alpha\,\mathsf{p}+\beta\,\mathsf{x})}|\psi) = (\psi|\mathsf{E}(\alpha,\beta)|\psi) \tag{75}$$

which establishes an intimate connection between the quantum theory of moments and the Weyl–Wigner formalism. Evidently

$$\mathsf{E}(\alpha,\beta) = \sum_{k=0}^{\infty} \tfrac{1}{k!}\big(\tfrac{i}{\hbar}\big)^k \Big\{\sum_{n=0}^{k} \mathsf{M}_{k-n,n}\alpha^{k-n}\beta^n\Big\} \tag{76}$$

$$\begin{aligned} \mathsf{M}_{m,n} &\equiv \sum_{\text{all orderings}} m\ \mathsf{p}\text{-factors and } n\ \mathsf{x}\text{-factors} \\ &= \text{sum of } \tbinom{m+n}{n} \text{ terms altogether} \end{aligned} \tag{77}$$

---

[59] We note in passing that if x and p are statistically independent then $\big\langle e^{\frac{i}{\hbar}(\alpha p+\beta x)}\big\rangle = \big\langle e^{\frac{i}{\hbar}\alpha p}\big\rangle\big\langle e^{\frac{i}{\hbar}\beta x}\big\rangle$, and we recover $P(x,p) = f(x)g(p)$.

To the claim that quantum mechanics admits of formulation as a "theory of interactive moments" one natural response would be "Moments of what?" The momental quantum mechanics of oscillators (discussed previously) suggests one answer: "Moments of the operators which enter into the 'algebraic completion' of the operator of initial interest." But the result just achieved suggests another —complementary—answer: "The elements of the momental set $\big\{\langle \mathsf{M}_{m,n}\rangle\big\}$," where the primitive operators $\mathsf{M}_{m,n}$ of low order are displayed below:

$$
\begin{aligned}
\mathsf{M}_{0,0} &= \mathbf{1}\\[4pt]
\mathsf{M}_{1,0} &= \mathsf{p}\\
\mathsf{M}_{0,1} &= \mathsf{x}\\[4pt]
\mathsf{M}_{2,0} &= \mathsf{pp}\\
\mathsf{M}_{1,1} &= \mathsf{px}+\mathsf{xp}\\
\mathsf{M}_{0,2} &= \mathsf{xx}\\[4pt]
\mathsf{M}_{3,0} &= \mathsf{ppp}\\
\mathsf{M}_{2,1} &= \mathsf{ppx}+\mathsf{pxp}+\mathsf{xpp}\\
\mathsf{M}_{1,2} &= \mathsf{pxx}+\mathsf{xpx}+\mathsf{xxp}\\
\mathsf{M}_{0,3} &= \mathsf{xxx}\\
&\ \vdots
\end{aligned}
$$

The relation of those observables to their classical counterparts can be described

$$
\tfrac{1}{\text{number of terms}}\mathsf{M}_{m,n} \xleftarrow[\text{Weyl}]{\qquad\qquad} p^m x^n
$$

We began with an interest—Ehrenfest's interest—in the quantum dynamical motion of only a pair of moments ($\langle\mathsf{x}\rangle$ and $\langle\mathsf{p}\rangle$), but from the structure of (68) acquired an enforced collateral interest in *mixed moments of all orders*. Here I explore implications of some commonplace wisdom:

> *When one has interest in properties of an infinite set of objects, it is often simplest and most illuminating to look not to the objects individually but to their* generating function.

I look now, therefore, to the dynamical properties that the "Moyal function"

$$
M_\psi(\alpha,\beta;t) \equiv (\psi|\mathsf{E}(\alpha,\beta)|\psi) = \langle\mathsf{E}(\alpha,\beta)\rangle
$$

acquires from those of $|\psi)$. Immediately

$$
\tfrac{\partial}{\partial t}M_\psi(\alpha,\beta;t) = \tfrac{1}{i\hbar}\langle[\mathsf{E}(\alpha,\beta),\mathsf{H}]\rangle \tag{78}
$$

One might use

$$\mathbf{H} \xleftarrow[\text{Weyl}]{} H(x,p) = \iint h(\alpha',\beta')e^{\frac{i}{\hbar}(\alpha' p+\beta' x)}\,d\alpha' d\beta'$$

to obtain

$$\tfrac{\partial}{\partial t}M_\psi(\alpha,\beta;t) = \tfrac{1}{i\hbar}\iint h(\alpha',\beta')\big\langle[\mathbf{E}(\alpha,\beta),\mathbf{E}(\alpha',\beta')]\big\rangle\,d\alpha' d\beta'$$

and then this corollary of (9.2)

$$\begin{aligned}[\mathbf{E}(\alpha,\beta),\mathbf{E}(\alpha',\beta')] &= 2i\sin\vartheta\cdot\mathbf{E}(\alpha+\alpha',\beta+\beta')\\ \vartheta &\equiv \tfrac{1}{2\hbar}(\alpha\beta'-\beta\alpha')\end{aligned}$$

to be led, after a couple of lines, to a statement

$$\begin{aligned}\tfrac{\partial}{\partial t}M_\psi(\alpha,\beta;t) &= \iint \mathfrak{T}(\alpha,\beta;\alpha',\beta')\cdot M_\psi(\alpha',\beta';t)\,d\alpha' d\beta' \\ \mathfrak{T}(\alpha,\beta;\alpha',\beta') &\equiv \tfrac{2}{\hbar}h(\alpha'-\alpha,\beta'-\beta)\sin\big(\tfrac{\alpha\beta'-\beta\alpha'}{2\hbar}\big)\end{aligned} \tag{79}$$

reminiscent of a previously reported description[20] of $\frac{\partial}{\partial t}P_\psi(x,p;t)$—both of which are reminiscent of

$$\tfrac{\partial}{\partial t}(x|\psi) = \int (x|\mathbf{H}|x')dx'(x'|\psi)$$

Equation (79) is equivalent to a giant system of coupled first-order differential equations in the mixed moments of all orders; it asserts that the time derivatives of those moments are linear combinations of their instantaneous values, and that it is the responsibility of the Hamiltonian to answer the question "*What* linear combinations?" and thus to distinguish one dynamical system from another. But (79), while of theoretical interest, is awkward to use in specific cases: it is often advantageous to work directly from (78).

To illustrate how that is done, we look again to the harmonic oscillator, where (78) gives

$$\begin{aligned}\tfrac{\partial}{\partial t}M_\psi(\alpha,\beta;t) &= \tfrac{1}{i\hbar}\big\langle\tfrac{1}{2m}[\mathbf{E}(\alpha,\beta),\mathbf{p}^2] + \tfrac{1}{2}m\omega^2[\mathbf{E}(\alpha,\beta),\mathbf{x}^2]\big\rangle\\ &= \tfrac{1}{i\hbar}\big\langle\big\{\tfrac{1}{2m}\big(-2\tfrac{\hbar}{i}\beta\tfrac{\partial}{\partial\alpha}\big)+\tfrac{1}{2}m\omega^2\big(+2\tfrac{\hbar}{i}\alpha\tfrac{\partial}{\partial\beta}\big)\big\}\mathbf{E}(\alpha,\beta)\big\rangle\\ &= \big\{\tfrac{1}{m}\beta\tfrac{\partial}{\partial\alpha} - m\omega^2\alpha\tfrac{\partial}{\partial\beta}\big\}M_\psi(\alpha,\beta;t) \qquad (80.1)\\ &\downarrow\\ &= \tfrac{1}{m}\beta\tfrac{\partial}{\partial\alpha}M_\psi(\alpha,\beta;t)\quad\text{in the “free particle limit”} \qquad (80.2)\end{aligned}$$

Explicit expansion of the expression on the left gives

$$\begin{aligned}\tfrac{\partial}{\partial t}M_\psi(\alpha,\beta) = \tfrac{\partial}{\partial t}\Big\{&\langle\mathbf{1}\rangle + \tfrac{i}{\hbar}\big[\alpha\langle\mathbf{p}\rangle+\beta\langle\mathbf{x}\rangle\big]\\ &+\tfrac{1}{2}\big(\tfrac{i}{\hbar}\big)^2\big[\alpha^2\langle\mathbf{p}^2\rangle + \alpha\beta\langle\mathbf{p}\mathbf{x}+\mathbf{x}\mathbf{p}\rangle + \beta^2\langle\mathbf{x}^2\rangle\big] + \cdots\Big\}\end{aligned}$$

while expansion of the expression on the right gives

$$\begin{aligned}&\big\{\tfrac{1}{m}\beta\tfrac{\partial}{\partial\alpha} - m\omega^2\alpha\tfrac{\partial}{\partial\beta}\big\}M_\psi(\alpha,\beta)\\ &\quad= \tfrac{i}{\hbar}\big[\tfrac{1}{m}\beta\langle\mathsf{p}\rangle - m\omega^2\alpha\langle\mathsf{x}\rangle\big]\\ &\qquad+ \tfrac{1}{2}\big(\tfrac{i}{\hbar}\big)^2\big[\tfrac{1}{m}2\alpha\beta\langle\mathsf{p}^2\rangle + \big(\tfrac{1}{m}\beta^2 - m\omega^2\alpha^2\big)\langle\mathsf{px}+\mathsf{xp}\rangle - m\omega^2 2\alpha\beta\langle\mathsf{x}^2\rangle\big] + \cdots\end{aligned}$$

Term-by-term identification gives rise to a system of equations

$$\begin{aligned}\alpha^1 &\quad:\quad \tfrac{d}{dt}\langle\mathsf{p}\rangle = -m\omega^2\langle\mathsf{x}\rangle\\ \beta^1 &\quad:\quad \tfrac{d}{dt}\langle\mathsf{x}\rangle = \tfrac{1}{m}\langle\mathsf{p}\rangle\\ \alpha^2 &\quad:\quad \tfrac{d}{dt}\langle\mathsf{p}^2\rangle = -m\omega^2\langle\mathsf{xp}+\mathsf{px}\rangle\\ \alpha\beta &\quad:\quad \tfrac{d}{dt}\langle\mathsf{xp}+\mathsf{px}\rangle = \tfrac{2}{m}\langle\mathsf{p}^2\rangle - 2m\omega^2\langle\mathsf{x}^2\rangle\\ \beta^2 &\quad:\quad \tfrac{d}{dt}\langle\mathsf{x}^2\rangle = \tfrac{1}{m}\langle\mathsf{xp}+\mathsf{px}\rangle\\ &\quad\vdots\end{aligned}$$

which in the "free particle limit" become

$$\begin{aligned}\alpha^1 &\quad:\quad \tfrac{d}{dt}\langle\mathsf{p}\rangle = 0\\ \beta^1 &\quad:\quad \tfrac{d}{dt}\langle\mathsf{x}\rangle = \tfrac{1}{m}\langle\mathsf{p}\rangle\\ \alpha^2 &\quad:\quad \tfrac{d}{dt}\langle\mathsf{p}^2\rangle = 0\\ \alpha\beta &\quad:\quad \tfrac{d}{dt}\langle\mathsf{xp}+\mathsf{px}\rangle = \tfrac{2}{m}\langle\mathsf{p}^2\rangle\\ \beta^2 &\quad:\quad \tfrac{d}{dt}\langle\mathsf{x}^2\rangle = \tfrac{1}{m}\langle\mathsf{xp}+\mathsf{px}\rangle\\ &\quad\vdots\end{aligned}$$

These are precisely the results achieved earlier by other means. It seems, on the basis of such computation, fair to assert that *equations of type (80) provide a succinct expression of Ehrenfest's theorem in its most general form*.

In an essay previously cited[58] I discuss how (at least in favorable cases) one might undertake to *solve* partial differential equations of the class typified by (80). Notice that such equation have this in common with the Hamilton-Jacobi equation: both permit *populations of coupled ordinary differential equations* to be cast as *solitary partial differential equations* ("field equations," if you will). I allude here to no mere superficial formal similarity, but to what is in fact a deep physical interconnection...but must reserve for another occasion the argument in defense of that claim.

To summarize: I do not claim that one *should* look upon quantum theory as a "theory of interactive (or coupled) moments," only that it is *possible*, and that the exercise does serve (i) to expose the seldom-remarked fact that *the quantum motion of moments is, to a large extent, universal/state-independent*. It serves also (ii) to bring spontaneously into focus the importance of Moyal's most distinctive contribution to the phase space formulation of quantum mechanics, which by

$$P(x,p) \xleftrightarrow[\text{Fourier}]{} M(\alpha,\beta)$$

lies "on the other side of the Rue de Fourier" from Wigner's own seminal contribution: moment theory leads back in a fairly natural way to the whole Weyl-Wigner apparatus.

It would be of interest, on another occasion, to look to examine properties of the *Husimi transform* of Moyal's function $M(\alpha,\beta)$; i.e., to look to the closure of the scheme

$$\begin{array}{ccc} P(x,p) & \xleftrightarrow[\text{Fourier}]{} & M(\alpha,\beta) \\ \big\downarrow \text{Husimi} & & \\ \boldsymbol{P}(x,p) & \xleftrightarrow[\text{Fourier}]{} & \boldsymbol{M}(\alpha,\beta) \end{array}$$

Wigner-like distributions on "finite-dimensional phase spaces." Let

$$\mathbf{f} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_x \\ \vdots \\ f_n \end{pmatrix} \quad \text{and} \quad \mathbf{g} = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_p \\ \vdots \\ g_n \end{pmatrix}$$

be n-dimensional "probability vectors;" i.e., number sets whose non-negative elements are constrained to satisfy $\sum_x f_x = \sum_p g_p = 1$. Use that material to construct the dyadic

$$\mathbb{P}_0 \equiv \|P_0(x,p)\| \equiv \mathbf{f}\,\mathbf{g}^\mathsf{T} = \begin{pmatrix} f_1g_1 & f_1g_2 & \dots & f_1g_p & \dots & f_1g_n \\ f_2g_1 & f_2g_2 & \dots & f_2g_p & \dots & f_2g_n \\ \vdots & \vdots & & \vdots & & \vdots \\ f_xg_1 & f_xg_2 & \dots & f_xg_p & \dots & f_xg_n \\ \vdots & \vdots & & \vdots & & \vdots \\ f_ng_1 & f_ng_2 & \dots & f_ng_p & \dots & f_ng_n \end{pmatrix}$$

Look upon $P_0(x,p)$ as a description of how duplex events $\{x,p\}$ are distributed on a discrete $n\times n$ phase space. The product structure of $P_0(x,p) = f_x \cdot g_p$ signifies that x and p are independent random variables. Now let $\mathbb{C} = \|c_{xp}\|$ be an $n\times n$ matrix with the property that

$$\sum_x c_{xp} = 0 \ \text{(all } p) \qquad \text{and} \qquad \sum_p c_{xp} = 0 \ \text{(all } x)$$

Such a matrix might be constructed by (i) writing

$$\begin{pmatrix} a_1 & a_2 & a_3 & \dots & a_n \\ a_2 & & & & \\ a_3 & & & & \\ \vdots & & & & \\ a_n & & & & \end{pmatrix}$$

and requiring that $\sum_i a_i = 0$; then (ii) writing

$$\begin{pmatrix} a_1 & a_2 & a_3 & \dots & a_n \\ a_2 & b_2 & b_3 & \dots & b_n \\ a_3 & b_3 & & & \\ \vdots & \vdots & & & \\ a_n & b_n & & & \end{pmatrix}$$

and requiring of the b's that the elements in the second row sum to zero; (iii) continuing in that wise to the construction of

$$\begin{pmatrix} a_1 & a_2 & a_3 & \dots & a_n \\ a_2 & b_2 & b_3 & \dots & b_n \\ a_3 & b_3 & c_3 & \dots & c_n \\ \vdots & \vdots & \vdots & & \vdots \\ a_n & b_n & c_n & & z_n \end{pmatrix}$$

and finally (iv) subjecting the rows (ditto and independently, the columns) to arbitrary permutations. Finally construct

$$\mathbb{P} \equiv \|P(x,p)\| = \mathbb{P}_0 + \mathbb{C} = \|f_x g_p + c_{xp}\|$$

Evidently

$$\sum_x \sum_p P(x,p) = 1 \tag{81}$$

and the associated marginal distributions

$$\sum_p P(x,p) = f_x \qquad \text{and} \qquad \sum_x P(x,p) = g_p \tag{82}$$

are (compare (24)) good upstanding distribution functions. But not so $P(x,p)$ itself, which—depending upon the values ascribed to the elements of $\mathbb{C}$—can assume non-positive values!

All of which can be made quite explicit in the case $n = 2$, to which I henceforth limit my attention. Write

$$\mathbf{f} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} X \\ 1-X \end{pmatrix} \quad \text{and} \quad \mathbf{g} = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} = \begin{pmatrix} P \\ 1-P \end{pmatrix} \quad : \quad 0 \leqslant X, P \leqslant 1$$

Then

$$\mathbb{P} = \begin{pmatrix} XP + c & X(1-P) - c \\ (1-X)P - c & (1-X)(1-P) + c \end{pmatrix} \tag{83}$$

and the

$$\text{“correlation matrix”} \equiv \|P(x,p) - f_x g_p\| = \begin{pmatrix} c & -c \\ -c & c \end{pmatrix} = \mathbb{C}$$

If, in particular, we set $X = 0.2$ and $P = 0.7$ we have

$$\mathsf{f} = \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}, \quad \mathsf{g} = \begin{pmatrix} 0.7 \\ 0.3 \end{pmatrix}, \quad \mathbb{P} = \begin{pmatrix} 0.14 + c & 0.06 - c \\ 0.56 - c & 0.24 + c \end{pmatrix}$$

All elements of $\mathbb{P}$ lie all on the unit interval $[0, 1]$, and are therefore interpretable as ordinary “probabilities,” if and only if $-0.14 \leqslant c \leqslant +0.06$. But relaxation of that constraint does violence neither to the normalization condition (81) nor to a probabilistic interpretation of the associated marginal distributions (82). The latter conditions would be preserved and the elements of $\mathbb{P}$ would all lie on the expanded unit interval $[-1, +1]$ if we required of c only that $-0.44 \leqslant c \leqslant +0.76$. We might, in particular, set $c = 0.36$ to obtain a discrete bivariate distribution (better: “quasi-distribution”)

$$\mathbb{P} = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} = \begin{pmatrix} 0.5 & -0.3 \\ 0.2 & 0.6 \end{pmatrix}$$

which displays all (or at any rate many) of the properties most characteristic of Wigner distributions. From (see again (34))

$$(\text{sum of the elements of } \mathbb{P}^2) = 0.38 < (\text{sum of the elements of } \mathbb{P}) = 1$$

we might infer that $\mathbb{P}$ refers to some kind of a “mixture,” but leave in suspension the precise meaning of that conclusion.

Preceding remarks were motivated by aspects of the discussion to which I now turn:

Feynman on “negative probability” and Bell’s theorem. In 1981 Richard Feynman (1918–1988) was invited to give the keynote address at an MIT conference on the “Physics of Computation.” The published text of his remarks on that occasion[60] established Feynman as a visionary founding father in the expanding field of “quantum computation,” and is of interest to us for several reasons:

- In his §5 alludes—for the only time in print, so far as I am aware—to the Wigner function $P(x, p)$, and indicates that he considers the fact that $P(x, p)$ becomes sometimes negative to lie at the very core of quantum mechanics.

[60] “Simulating physics with computers,” International J. of Theo. Phys. **21**, 467 (1982). The paper (together with a rich collection of kindred papers) has been reprinted in *Feynman & Computation: Exploring the Limits of Computers* (1999), which was edited by A. J. G. Hey.

- To place himself in position to discuss that fact in simplest possible terms he pulls out of his hat the matrix

$$\mathbb{P}_{\text{Feynman}} = \begin{pmatrix} 0.6 & -0.1 \\ 0.3 & 0.2 \end{pmatrix}$$

  and observes that, though it contains a negative element, it does possess the properties (81/82) exhibited by all polite bivariate distributions.[61]

Feynman's motivation is multi-stranded, and so are his conclusions (which include an admission that he is "not sure there's no real problem" with quantum mechanics!). I propose to examine just a couple of those strands.

Feynman had developed in interest in the physical principles which limit the miniaturization of devices already by the mid-1950's. In an after-dinner speech presented at a 1959 meeting of the American Physical Society[62] he offered \$1000 prizes to "to the first guy who makes an operating electric motor... only $1/64^{\text{th}}$ inch cube" and "to the first guy who can take the information on a page of a book and put it on an area 1/25,000 smaller in linear scale in such a manner that it can be read by an electron microscope." The first prize was won within a year by one Bill McLellan, whose motor was $.006''$ in diameter and developed a reported 10^{-6} horsepower.[63] That semi-whimsical interest in microdevices ripened fairly naturally into an interest in quantum robotics and (especially) quantum computation.

A few pages prior to the presentation of $\mathbb{P}_{\text{Feynman}}$ Feynman had posed the question: "Can quantum systems be probabilistically simulated by a classical computer?" The question motivated him to sketch how quantum mechanics might be written so as to look as classical as possible, and led him to a broadly Wigneresque conception of quantum essentials. He concludes his §6 with these words:

---

[61] We, by the way, find the latter circumstance not at all surprising, since

$$\mathbb{P}_{\text{Feynman}} = (83) \quad \text{with} \quad X = 0.5,\ P = 0.9,\ c = 0.15$$

[62] The text is reprinted as "There's plenty of room at the bottom" in Hey's remarkable anthology (just cited), and served to establish Feynman as a founding father also of what has come to be called "nanotechnology."

[63] If those numbers are to be believed, then McLellan's motor had a power density of 4.6 hp/in^3, which is about 10^4 higher than is available in motors you can walk into a store and buy. When I, as a graduate student, first learned of Feynman's challenge I remarked to a friend that "the way to make such a motor is to make it out of meat," so had special interest when, shortly thereafter, biologists worked out the design of the "motor" that spins microbial flagella, the tails of spermatozoa, etc. By the way: text reduction at the scale stipulated by Feynman would permit the entire Encyclopedia Britanica to be printed on the head of a pin, and was first accomplished by Tom Newman, a graduate student at Stanford, in 1985.

> *"But* [*some of the elements of* $\mathbb{P}_{\text{Feynman}}$] *are not positive, and therein lies the great difficulty. The only difference between a probabilistic classical world and the equations of the quantum world is that somehow or other it appears as if the probabilities would have to go negative, and that we do not know, as far as I know, how to simulate. Okay, that's the fundamental problem. I don't know the answer to it, but I wanted to explain that if I try my best to make the equations look as near as possible to what would be imitable by a classical probabilistic computer, I get into trouble.*

At the beginning of his §7 he continues

> *I would like to show you why such minus signs* <u>cannot be avoided</u> [my emphasis], *or at least that you have some sort of difficulty.*

Feynman's argument seems to me to be, in some respects, a bit goofy. At its heart lies *Bell's theorem*, which he reconstructs without mention of Bell,[64] and in a way which (partly because the mathematics is developed verbally) I find much less clear than Bell's own demonstration.[65] Bell's accomplishment was to establish that certain specific statements that follow from orthodox quantum mechanics (statements which refer to the entangled states of 2-state systems, in the tradition of EPR—statements which have since been found to conform to the observational facts) *cannot be replicated by any deterministic local hidden variable theory in which ordinary probability theory is used to manage implications of the circumstance that the hidden variables themselves remain unknown/random.* Feynman (recall the title[60] of his talk) concludes on this basis that <u>no classical computer can be entirely successful as a simulator of quantum mechanics</u>.[66] But of more immediate interest to us is a remark which he drops in passing: "Such a formula cannot reproduce the quantum results if [certain details built into Bell's/Feynman's argument refer to] real probabilities ... but is easy if they are 'probabilities'—[allowed to become] negative [under] some conditions ..."

---

[64] Perhaps Feynman considered that his introductory "You probably have all heard of this example of the EPR paradox, but I will explain..." made explicit citation unnecessary. Or perhaps he is hinting that he had discovered "Bell's theorem"—by then eighteen years old—independently. Or perhaps there was some personal antagonism; I notice that J. S. Bell, in the twenty-two papers reproduced in his *Speakable and unspeakable in quantum mechanics* (1987), is unfailingly generous in recognizing those to whom he is indebted, but cites Feynman only once—glancingly.

[65] For a lucid reproduction of Bell's argument—stripped of a few fussy details, illuminated by some commentary—see Griffiths' §A.2. Bell's original paper ("On the Einstein-Podolsky-Rosen paradox," Physics **1**, 195–200 (1964)) is reprinted in *Speakable...*

[66] This conclusion inspires the speculation that "quantum computers" may possess capabilities beyond the reach of classical computers: "...it seems that the laws of physics present no barrier to reducing the size of computers until bits are the size of atoms, and quantum behavior holds dominant sway."

Getting down to work: Feynman speaks of "photons," and uses birefringent calcite crystals[68] to distinguish linearly polarized $\leftrightarrow$ states from $\updownarrow$; Griffiths looks to the respective spins of the electron/positron produced in a reaction of the type $e^- \longleftarrow \pi^0 \longrightarrow e^+$; Bell asks us to "consider a pair of spin one-half particles formed somehow in a singlet spin state and moving freely in opposite directions." Those distinctions are of no deep significance; I prefer to work in Bell's more generic language.

Feynman, Griffiths and Bell assume their readers to be familiar with a quantum mechanical result (the result which it is the assigned business of hidden variable theory to account for) which we must, as a preparatory set, digress to acquire. Let $\mathfrak{S}_1$ and $\mathfrak{S}_2$ be 2-state systems, from which we assemble the composite system $\mathfrak{S} = \mathfrak{S}_1 \otimes \mathfrak{S}_2$. We propose, in the EPR tradition, to examine $\mathfrak{S}$ first with

$$\mathbb{A} \equiv (a_1\boldsymbol{\sigma}_1 + a_2\boldsymbol{\sigma}_2 + a_3\boldsymbol{\sigma}_3) \otimes \mathbb{I} \quad : \quad a_1^2 + a_2^2 + a_3^2 = 1 \tag{84.1}$$

—which *in the absence of entanglement* would be to record a property of $\mathfrak{S}_1$—and then with[69]

$$\mathbb{B} \equiv \mathbb{I} \otimes (b_1\boldsymbol{\sigma}_1 + b_2\boldsymbol{\sigma}_2 + b_3\boldsymbol{\sigma}_3) \quad : \quad b_1^2 + b_2^2 + b_3^2 = 1 \tag{84.2}$$

Recalling from (1–7) the conventional definitions of the Pauli matrices, then using *Mathematica*'s

```
Outer[Times, first matrix, second matrix]//MatrixForm
```

to work out the Kronecker products, we find

$$\mathbb{A} = \begin{pmatrix} a_3 & 0 & a_1 - ia_2 & 0 \\ 0 & a_3 & 0 & a_1 - ia_2 \\ a_1 + ia_2 & 0 & -a_3 & 0 \\ 0 & a_1 + ia_2 & 0 & -a_3 \end{pmatrix} \tag{85.1}$$

$$\mathbb{B} = \begin{pmatrix} b_3 & b_1 - ib_2 & 0 & 0 \\ b_1 + ib_2 & -b_3 & 0 & 0 \\ 0 & 0 & b_3 & b_1 - ib_2 \\ 0 & 0 & b_1 + ib_2 & -b_3 \end{pmatrix} \tag{85.2}$$

Both matrices are manifestly Hermitian, and both (use `Eigenvalues[matrix]` and $\boldsymbol{a}\cdot\boldsymbol{a} = \boldsymbol{b}\cdot\boldsymbol{b} = 1$) have spectra which can be described $\{-1, -1, +1, +1\}$. Entrusting the heavy labor to *Mathematica* we are led at length to the *spectral representations* of $\mathbb{A}$ and $\mathbb{B}$:

$$\mathbb{A} = (+1)\,\mathbb{P}_+^{\mathsf{A}} + (-1)\,\mathbb{P}_-^{\mathsf{A}} \tag{86.1}$$

$$\mathbb{B} = (+1)\,\mathbb{P}_+^{\mathsf{B}} + (-1)\,\mathbb{P}_-^{\mathsf{B}} \tag{86.2}$$

---

[68] See E. Hecht & A. Zajac, *Optics* (1979) §§8.4.1–3.

[69] I write $\mathbb{A}$ and $\mathbb{B}$ where by former convention—introduced at (1–64)—I would have written $\mathbf{A}$ and $\mathbf{B}$ to emphasize the 4-dimensionality of the matrices in question.

where

$$\mathbb{P}_+^{\mathsf{A}} \equiv \begin{pmatrix} \frac{1}{2}(1+a_3) & 0 & +\frac{1}{2}(a_1-ia_2) & 0 \\ 0 & \frac{1}{2}(1+a_3) & 0 & +\frac{1}{2}(a_1-ia_2) \\ +\frac{1}{2}(a_1+ia_2) & 0 & \frac{1}{2}(1-a_3) & 0 \\ 0 & +\frac{1}{2}(a_1+ia_2) & 0 & \frac{1}{2}(1-a_3) \end{pmatrix} \tag{87.1}$$

$$\mathbb{P}_-^{\mathsf{A}} \equiv \begin{pmatrix} \frac{1}{2}(1-a_3) & 0 & -\frac{1}{2}(a_1-ia_2) & 0 \\ 0 & \frac{1}{2}(1-a_3) & 0 & -\frac{1}{2}(a_1-ia_2) \\ -\frac{1}{2}(a_1+ia_2) & 0 & \frac{1}{2}(1+a_3) & 0 \\ 0 & -\frac{1}{2}(a_1+ia_2) & 0 & \frac{1}{2}(1+a_3) \end{pmatrix} \tag{87.2}$$

$$\mathbb{P}_+^{\mathsf{B}} \equiv \begin{pmatrix} \frac{1}{2}(1+b_3) & +\frac{1}{2}(b_1-ib_2) & 0 & 0 \\ +\frac{1}{2}(b_1+ib_2) & \frac{1}{2}(1-b_3) & 0 & 0 \\ 0 & 0 & \frac{1}{2}(1+b_3) & +\frac{1}{2}(b_1-ib_2) \\ 0 & 0 & +\frac{1}{2}(b_1+ib_2) & \frac{1}{2}(1-b_3) \end{pmatrix} \tag{88.1}$$

$$\mathbb{P}_-^{\mathsf{B}} \equiv \begin{pmatrix} \frac{1}{2}(1-b_3) & -\frac{1}{2}(b_1-ib_2) & 0 & 0 \\ -\frac{1}{2}(b_1+ib_2) & \frac{1}{2}(1+b_3) & 0 & 0 \\ 0 & 0 & \frac{1}{2}(1-b_3) & -\frac{1}{2}(b_1-ib_2) \\ 0 & 0 & -\frac{1}{2}(b_1+ib_2) & \frac{1}{2}(1+b_3) \end{pmatrix} \tag{88.2}$$

I need not record the straightforwardly tedious details that led to the preceding statements, for one can readily *verify after the fact* that the matrices thus defined do possess all the required properties:

- $\mathbb{P}_+^{\mathsf{A}}$ and $\mathbb{P}_-^{\mathsf{A}}$ are Hermitian
- each is projective: $\mathbb{P}_+^{\mathsf{A}}\,\mathbb{P}_+^{\mathsf{A}} = \mathbb{P}_+^{\mathsf{A}}$, $\mathbb{P}_-^{\mathsf{A}}\,\mathbb{P}_-^{\mathsf{A}} = \mathbb{P}_-^{\mathsf{A}}$
- they are orthogonal and complementary: $\mathbb{P}_+^{\mathsf{A}}\,\mathbb{P}_-^{\mathsf{A}} = \mathbb{O}$ and $\mathbb{P}_+^{\mathsf{A}} + \mathbb{P}_-^{\mathsf{A}} = \mathbb{I}$
- they comply with (86.1): $\mathbb{P}_+^{\mathsf{A}} - \mathbb{P}_+^{\mathsf{A}} = \mathbb{A}$
- each has spectrum $\{1,1,0,0\}$, so projects onto a 2-space

and the same can be said of $\mathbb{P}_+^{\mathsf{B}}$ and $\mathbb{P}_-^{\mathsf{B}}$. All this *Mathematica* would be very quick to confirm. With that preparation behind us ... we assume $\mathfrak{S}$ to have been placed "somehow" in a singlet state—a state, that is, to say, with the property that it is killed by both $\mathbf{S}^2$ and $\mathbf{S}_3$; we found at (1–78.2)/(1–85) that the state in question can be described

$$|\text{singlet}) = \tfrac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ +1 \\ -1 \\ 0 \end{pmatrix} = \tfrac{1}{\sqrt{2}}(\uparrow\downarrow - \downarrow\uparrow)$$

Examination of such a state with our $\mathbf{A}$-meter yields

$$\begin{aligned} \mathbb{P}_+^{\mathsf{A}}\,|\text{singlet}) &\equiv |\mathsf{A}_+) \quad \text{when the } \mathbf{A}\text{-meter registers } +1 \\ \mathbb{P}_-^{\mathsf{A}}\,|\text{singlet}) &\equiv |\mathsf{A}_-) \quad \text{when the } \mathbf{A}\text{-meter registers } -1 \end{aligned}$$

By quick calculation

$$|\mathsf{A}_+) = \tfrac{1}{2\sqrt{2}}\begin{pmatrix} -(a_1 - ia_2) \\ +(1+a_3) \\ -(1-a_3) \\ +(a_1+ia_2) \end{pmatrix} \quad \text{and} \quad |\mathsf{A}_-) = \tfrac{1}{2\sqrt{2}}\begin{pmatrix} +(a_1 - ia_2) \\ +(1-a_3) \\ -(1+a_3) \\ -(a_1+ia_2) \end{pmatrix}$$

We observe that

$$|\mathsf{A}_+) + |\mathsf{A}_-) = |\text{singlet}) \quad \text{and} \quad (\mathsf{A}_+|\mathsf{A}_-) = 0$$

—which are gratifying but expected—and

$$(\mathsf{A}_+|\mathsf{A}_+) = (\mathsf{A}_-|\mathsf{A}_-) = \tfrac{1}{2}$$

from which we learn that the **A**-meter registers $+1$ else -1 with equal liklihood *for all assignments of* $\boldsymbol{a}$; the singlet state is—literally, if we imagine ourselves to be speaking about a spin system; formally if we assign a more abstract meaning to the phrase "2-state system"—"rotationally invariant."

The vectors $|\mathsf{A}_\pm)$ must be restored to unit length (multiplied by $\sqrt{2}$) before they become proper state descriptors. That done, we activate the **B**-meter, which—since it looks at one or another of two states, and in each case prepares one or another of two states—requires that we consider four distinct cases:

$$\begin{aligned}
\mathbb{P}^{\mathsf{B}}_+\sqrt{2}|\mathsf{A}_+) &\equiv |\mathsf{A}_+,\mathsf{B}_+) &:&\quad \||\mathsf{A}_+,\mathsf{B}_+)\|^2 = \tfrac{1}{2}(1-\boldsymbol{a}\cdot\boldsymbol{b})\\
\mathbb{P}^{\mathsf{B}}_-\sqrt{2}|\mathsf{A}_+) &\equiv |\mathsf{A}_+,\mathsf{B}_-) &:&\quad \||\mathsf{A}_+,\mathsf{B}_-)\|^2 = \tfrac{1}{2}(1+\boldsymbol{a}\cdot\boldsymbol{b})\\
\mathbb{P}^{\mathsf{B}}_+\sqrt{2}|\mathsf{A}_-) &\equiv |\mathsf{A}_-,\mathsf{B}_+) &:&\quad \||\mathsf{A}_-,\mathsf{B}_+)\|^2 = \tfrac{1}{2}(1+\boldsymbol{a}\cdot\boldsymbol{b})\\
\mathbb{P}^{\mathsf{B}}_-\sqrt{2}|\mathsf{A}_-) &\equiv |\mathsf{A}_-,\mathsf{B}_-) &:&\quad \||\mathsf{A}_-,\mathsf{B}_-)\|^2 = \tfrac{1}{2}(1-\boldsymbol{a}\cdot\boldsymbol{b})
\end{aligned}$$

The vectors $|\mathsf{A}_\pm,\mathsf{B}_\pm)$ are fairly complicated; *Mathematica* is happy enough to write them out, but it would serve no purpose to do so here (and, once again, they must be normalized before they become descriptors of the final state of the composite system). We arrive thus at this generalization of (1–86.2):

$$\left.\begin{aligned}
\text{if “+” then} &\begin{cases} \text{“+” with probability } \tfrac{1}{2}(1-\boldsymbol{a}\cdot\boldsymbol{b}) \\ \text{“−” with probability } \tfrac{1}{2}(1+\boldsymbol{a}\cdot\boldsymbol{b}) \end{cases}\\
\text{if “−” then} &\begin{cases} \text{“+” with probability } \tfrac{1}{2}(1+\boldsymbol{a}\cdot\boldsymbol{b}) \\ \text{“−” with probability } \tfrac{1}{2}(1-\boldsymbol{a}\cdot\boldsymbol{b}) \end{cases}
\end{aligned}\right\} \tag{89}$$

If we set $\boldsymbol{a} = \boldsymbol{b}$ then we recover precisely the situation (1–86.2) which Einstein *et al* found so perplexing. Bell, however, was looking beyond EPR's perplexity to another issue (the hidden variable question), and had the wit to allow $\boldsymbol{a}$ and $\boldsymbol{b}$ to be specified independently/arbitrarily. He noticed more particularly that the meter readings will be

$$\begin{aligned}
\text{coincident } (++ \text{ or } --) \text{ with probability } &\tfrac{1}{2}(1-\boldsymbol{a}\cdot\boldsymbol{b})\\
\text{anticoincident } (+- \text{ or } -+) \text{ with probability } &\tfrac{1}{2}(1+\boldsymbol{a}\cdot\boldsymbol{b})
\end{aligned}$$

and that the

$$\begin{aligned}\underline{\text{averaged product}}\text{ of the meter readings} &= \tfrac{1}{2}(1-\boldsymbol{a}\cdot\boldsymbol{b})-\tfrac{1}{2}(1+\boldsymbol{a}\cdot\boldsymbol{b})\\ &= -\,\boldsymbol{a}\cdot\boldsymbol{b} \qquad (90)\\ &\equiv -\cos\theta\end{aligned}$$

This is the quantum mechanical result which Bell/Feynman are content simply to report, and which Griffiths assigns as *****Problem 4.44**.[70]

Now to Bell's main argument: Suppose some "hidden variables" λ lurk within the physics of $\mathfrak{S}$, where they enter as arguments into a function

$$A(\boldsymbol{a},\lambda)=\begin{cases}+1 \text{ at some points in } \lambda\text{-space}\\ -1 \text{ on the complementary point set}\end{cases}$$

---

[70] When I asked David how he himself would go about solving Problem 4.44 he (with the disclaimer that his solution was "most <u>in</u>elegant") promptly supplied the following: "We may as well choose axes so that $\boldsymbol{a}$ lies along the z axis and $\boldsymbol{b}$ is in the xz plane. Then $S_a^{(1)}=S_z^{(1)}$ and $S_b^{(2)}=\cos\theta S_z^{(2)}+\sin\theta S_x^{(2)}$. We are to calculate $\langle 00|S_a^{(1)}S_b^{(2)}|00\rangle$. But

$$\begin{aligned}S_a^{(1)}S_b^{(2)}|00\rangle &= \tfrac{1}{\sqrt{2}}\big\{S_z^{(1)}(\cos\theta S_z^{(2)}+\sin\theta S_x^{(2)})(\uparrow\downarrow-\downarrow\uparrow)\big\}\\ &= \tfrac{1}{\sqrt{2}}\big\{(S_z\uparrow)(\cos\theta S_z\downarrow+\sin\theta S_x\downarrow)\\ &\qquad -(S_z\downarrow)(\cos\theta S_z\uparrow+\sin\theta S_x\uparrow)\big\}\\ &= \tfrac{1}{\sqrt{2}}\big\{\tfrac{\hbar}{2}\uparrow(\cos\theta(-\tfrac{\hbar}{2}\downarrow)+\sin\theta(\tfrac{\hbar}{2}\uparrow))\\ &\qquad -(-\tfrac{\hbar}{2}\downarrow)(\cos\theta(\tfrac{\hbar}{2}\uparrow)+\sin\theta(\tfrac{\hbar}{2}\downarrow))\big\} \qquad \text{by [4.145]}\\ &= \tfrac{\hbar^2}{4}\big\{\cos\theta\tfrac{1}{\sqrt{2}}(-\uparrow\downarrow+\downarrow\uparrow)+\sin\theta\tfrac{1}{\sqrt{2}}(\uparrow\uparrow+\downarrow\downarrow)\big\}\\ &= \tfrac{\hbar^2}{4}\big\{-\cos|00\rangle+\sin\theta(|11\rangle+|1\,{-1}\rangle)\big\}\end{aligned}$$

So

$$\begin{aligned}\langle S_a^{(1)}S_b^{(2)}\rangle=\langle 00|S_a^{(1)}S_b^{(2)}|00\rangle &= \tfrac{\hbar^2}{4}\langle 00|\big\{-\cos\theta|00\rangle+\sin\theta(|11\rangle+|1\,{-1}\rangle)\big\}\\ &= -\tfrac{\hbar^2}{4}\cos\theta\langle 00|00\rangle+0 \qquad \text{by orthogonality}\\ &= -\tfrac{\hbar^2}{4}\cos\theta \qquad \text{QED"}\end{aligned}$$

David's argument—which is in a well-established tradition—presumes skill with the $\uparrow\downarrow$-notation (which I, for some reason, have never been able to acquire), but is so brief as to make my own argument (which I would have been reluctant to undertake without the assistance of *Mathematica*) look longwindedly pedantic. My argument makes no special assumption (such as David makes at the outset), and makes clear its reliance upon quantum mechanical first principles, but presumes familiarity with the Kronecker product approach to composite system theory. We are agreed that Bell himself probably argued that the right side of (90) *has necessarily on transformation-theoretic grounds* to be of the form $p+q\boldsymbol{a}\cdot\boldsymbol{b}$, and used a couple of special cases (EPR with $\boldsymbol{b}=\pm\boldsymbol{a}$?) to enforce $p=0$ and $q=-1$.

that deterministically directs $\mathfrak{S}_1$ how to interact with the **A**-meter, and a second function

$$B(\boldsymbol{b},\lambda) = \begin{cases} +1 \text{ at some points in } \lambda\text{-space} \\ -1 \text{ on the complementary point set} \end{cases}$$

that similarly controls the interaction of $\mathfrak{S}_2$ how to interact with the **B**-meter. EPR tell us that

$$B(\boldsymbol{a},\lambda) = -A(\boldsymbol{a},\lambda) \quad : \quad \text{all } \lambda$$

The value of λ is assumed to be "known" precisely to $\mathfrak{S}$, but only statistically to us: $\rho(\lambda)$.[71] Within the framework thus established, we have

$$\begin{aligned} P(\boldsymbol{a},\boldsymbol{b}) &\equiv \underline{\text{averaged product}} \text{ of the meter readings} \\ &= \int \rho(\lambda) A(\boldsymbol{a},\lambda) B(\boldsymbol{b},\lambda)\, d\lambda \qquad (\star) \\ &= -\int \rho(\lambda) A(\boldsymbol{a},\lambda) A(\boldsymbol{b},\lambda)\, d\lambda \end{aligned}$$

Therefore, for any $\boldsymbol{c}$,

$$P(\boldsymbol{a},\boldsymbol{b}) - P(\boldsymbol{a},\boldsymbol{c}) = -\int \rho(\lambda) \Big\{ A(\boldsymbol{a},\lambda)A(\boldsymbol{b},\lambda) - A(\boldsymbol{a},\lambda)A(\boldsymbol{c},\lambda) \Big\}\, d\lambda$$

which by $[A(\boldsymbol{b},\lambda)]^2 = (\pm 1)^2 = 1$ can be expressed

$$= -\int \rho(\lambda) \Big\{ 1 - A(\boldsymbol{b},\lambda)A(\boldsymbol{c},\lambda) \Big\} A(\boldsymbol{a},\lambda)A(\boldsymbol{b},\lambda)\, d\lambda$$

giving

$$\begin{aligned} \big|P(\boldsymbol{a},\boldsymbol{b}) - P(\boldsymbol{a},\boldsymbol{c})\big| &= \left| \int \underbrace{\rho(\lambda)\Big\{1 - A(\boldsymbol{b},\lambda)A(\boldsymbol{c},\lambda)\Big\}}_{\text{non-negative}} A(\boldsymbol{a},\lambda)A(\boldsymbol{b},\lambda)\, d\lambda \right| \qquad (\star\star) \\ &\leqslant \int \rho(\lambda)\Big\{1 - A(\boldsymbol{b},\lambda)A(\boldsymbol{c},\lambda)\Big\} \underbrace{\big|A(\boldsymbol{a},\lambda)A(\boldsymbol{b},\lambda)\big|}_{\text{unity}}\, d\lambda \end{aligned}$$

whence

$$\big|P(\boldsymbol{a},\boldsymbol{b}) - P(\boldsymbol{a},\boldsymbol{c})\big| \leqslant 1 + P(\boldsymbol{b},\boldsymbol{c}) \qquad (91)$$

which is the celebrated *Bell inequality*.

Were (91) to hold quantum mechanically we would, by (90), have

$$\big|\boldsymbol{a}\cdot(\boldsymbol{b}-\boldsymbol{c})\big| + \boldsymbol{b}\cdot\boldsymbol{c} \leqslant 1 \qquad (92)$$

Bell noticed that it is possible to select $\{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}\}$ so that (92) is *violated*; it is possible, in other words, to devise quantum mechanical situations for which no

[71] How we would come to know even $\rho(\lambda)$-much about the hidden variables is an interesting question, but one which, for the purposes at hand, need not concern us.

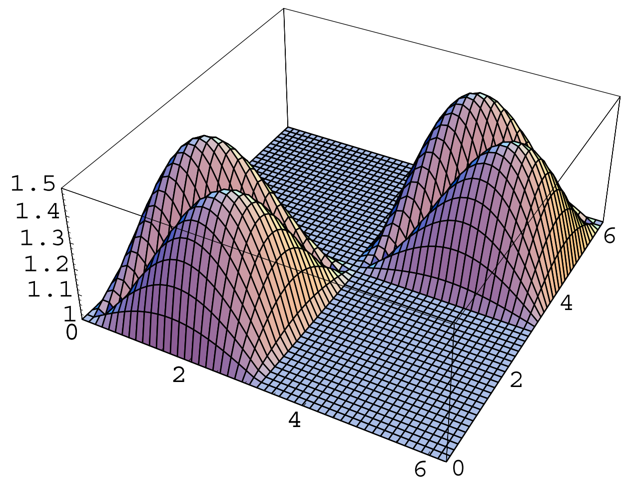

FIGURE 14: *Indication of the points in (β,γ)-space where (93) is violated. β runs $\searrow$, γ runs $\nearrow$. The specific command was*

```
Plot3D[Sqrt[(Cos[β]-Cos[γ])^2]+Cos[γ−β], {β, 0, 2π},
{γ, 0, 2π}, PlotRange->{1.0,1.5}, PlotPoints->50]
```

Note the $\sqrt{\text{square}}$ technique used to mimic the effect of absolute value bars; also the use of `PlotRange` *to eclipse the points where (93) is* not *violated.*

hidden variable theory (within the broad class of such theories contemplated by Bell) can account. A single example serves to establish the point: Bell/Griffiths assume $\{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}\}$ to be co-planar, $\boldsymbol{a}\perp\boldsymbol{b}$, and $\boldsymbol{c}$ to be the bisector of that right angle; then (92) reads $\left|0-\frac{1}{\sqrt{2}}\right|+\frac{1}{\sqrt{2}}=\sqrt{2}\leqslant 1$, which is absurd. To gain a somewhat more comprehensive view of the situation, I retain the assumption that $\{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}\}$ are co-planar and write $\boldsymbol{a}\cdot\boldsymbol{b}=\cos\beta$, $\boldsymbol{a}\cdot\boldsymbol{c}=\cos\gamma$. Then (92) reads

$$|\cos\beta-\cos\gamma\,|+\cos(\gamma-\beta)\leqslant 1 \tag{93}$$

Points where this instance of Bell's inequality is violated are shown in the figure.

Feynman observed that the force of Bell's argument would be lost if at $(\star\star)$ one allowed $\rho(\lambda)$ to assume negative values. To make the point notationally more vivid he assigns distinct sets α and β of hidden parameters to $\mathfrak{S}_1$ and $\mathfrak{S}_2$. In place of $(\star)$ he writes

$$P(\boldsymbol{a},\boldsymbol{b})=\iint\rho(\alpha,\beta)A(\boldsymbol{a},\alpha)B(\boldsymbol{b},\beta)\,d\alpha d\beta$$

and seems speculatively prepared to assign Wigner-like properties to $\rho(\alpha,\beta)$. His point seems to be that *when one looks closely to the quantum/classical* *connection one can expect to encounter "negative probability" ... if not on one side of the equation, then on the other.*[72]

Feynman admits to being quite at a loss when it comes to the question "What does 'negative probability' actually mean?" but appears to regard its occurance as a symptom of other, deeper problems. Quoting from his concluding remarks

> *"It seems to be almost ridiculous that you can squeeze [the difficulty of quantum mechanics] to a numerical question that one thing is bigger than another. But there you are— ... It is interesting to try to discuss the possibilities. I mentioned something about the possibility of time—of things being affected not just by the past, but also by the future, and therefore that our probabilities are in some sense 'illusory.' We only have the information from the past, and we try to predict the next step, but in reality it depends upon the near future which we can't get at, or something like that. A very interesting question is the origin of the probabilities in quantum mechanics. . . "*

The notion that we might have to adopt an altered view of time itself is being advanced here by the physicist who taught the world to look upon positrons as "electrons running backward in time,"[73] and who collaborated with John Wheeler to develop the "absorber theory"[74] which assigns equal weight to the advanced and retarded potentials of classical electrodynamics.[75]

Or perhaps the problem has to do with our still-imperfect understanding

---

[72] A somewhat related point was made by Bell himself, a few years later: see "EPR correlations and EPW distributions" (1986), which is reprinted as the penultimate essay in *Speakable ...* The paper begins with these words

> *It is known that with Bohm's example of EPR correlations involving particles with spin, there is an irreducible non-locality [which] cannot be removed by the introduction of hypothetical variables unknown to ordinary quantum mechanics. How is it with the original EPR example involving two particles of zero spin? Here we will see that the Wigner phase space distribution illuminates the problem.*

and continues to relate the occurance of negative probability to violation of a certain inequality. The "EPW" in the title is, of course, a cute reference, to E. P. Wigner, to whom the essay is dedicated.

[73] "The theory of positrons," Phys. Rev. **76**, 749 (1949).

[74] See F. Rohrlich, *Classical Charged Particles* (1965) §7.2 for brief discussion and references.

[75] As I student I was struck by the fact that the temporal unidirectionality of the diffusion equation can, in reference to a simple random walk model, be traced to the operation of ordinary probability theory. I asked: "Can 'negative probability' be used to construct a theory of backward diffusion?" The question has borne no fruit, but I admit to being still susceptible to its vague charm.

of what we should mean when we refer to the "state" of a quantum system. Bell has many times stressed, and so have many others, that in this area our axiomatic principles speak with the appearance of a mathematical precision much sharper than the facts of the matter are able to support. Perhaps we should adopt the principle that

If it talks "negative probability" it ain't a state

and look to Husimi for guidance toward the implementation of that principle.

Concluding remarks. I have reviewed the essentials of the Weyl–Wigner–Moyal "phase space formulation of non-relativistic quantum mechanics," and tried to indicate why the existence of such a formalism is worthy of notice. And I have explored a few of the theory's nooks and crannies. But the discussion could be much extended, for I have said not a word about (for example)

- Wigner functions on phase spaces of $n > 2$ dimensions
- why the subject is of special interest to chemical physicists
- applications to quantum optics
- applications to the study of quantum chaos.

For all of that and more I must refer my reader to the vast literature.

A few pages ago I finally received mail from the University of Minnesotta library which permitted me for—the first time—actually to *examine* Husimi's long neglected but recently much cited "Some formal properties of the density matrix."[41] The paper turns out to be a 50-page critical review of essentially all that had been learned about the density matrix in the dozen years since its invention. The author has an evidently deep familiarity with the European (especially the German) literature of the 1930's; he cites many/most of major figures of the period (Dirac, von Neumann, Courant & Hilbert, Szegö, Peierls, Delbrück, Uhlenbeck, Fock . . . but the work of not a single Japanese physicist), but—curiously—seems <u>ignorant of the work of Wigner</u> (work[9] which Wigner in 1932 did not claim had anything to do with the density matrix), work which Husimi is motivated in his §5 to sort of (but only sort of) re-invent. Buried in that discussion is the work for which he is now mainly remembered. I get the impression that Kôdi Husimi was an exceptionally capable young theoretical physicist, yet he was, so far as I am aware, never heard from again. I hope one day to learn the circumstances of his life . . . and death.