GAUSSIAN WAVEPACKETS

Nicholas Wheeler, Reed College Physics Department
July 1998

Introduction. Recent work—and a remark dropped casually by my statistical colleague, Albyn Jones—has led me to the speculative perception that "the quantum world, under certain circumstances, appears classical" for ultimately the same reason that experiments repeated many times over can be expected to yield results of high accuracy, and that the statistical properties of thermalized systems are so sharp as to become susceptible to analysis by the methods of thermodynamics.

The classical theory of errors is dominated by the *normal* distribution for reasons rooted in the *central limit theorem*. The train of thought to which I have alluded leads me to contemplate the existence of a quantum analog of the central limit theorem, phrased in terms not of probability distributions but of probability *amplitudes* or—equivalently but (as I will argue) more naturally—*Wigner* distributions.

My purpose here is simply to collect together, for the convenience of future reference, material pertaining to the "Gaussian quantum mechanics" which will be central to any effort to put meat on the bare bones of my present intuition. The rudiments of this subject are, of course, treated in every introductory quantum text,[1] but closer examination turns up a number of subtleties and complications, and exposes a variety of methodological options, to which I will draw attention. I borrow freely from some informal notes[2] which were written in support of the thesis research of a former student[3] and inspired originally by quite a different train of thought. Many of the results reported in the latter sections of the essay are, so far as I am aware, new.

I work mainly in one-dimension, and will borrow mainly from the quantum physics of *free* particles to lend dynamical substance to my remarks.

---

[1] See, for example, David Griffiths, *Introduction to Quantum Mechanics* (1995), Problems 2.22 (p. 50) and 2.40 (p. 69), and §3.4.2 (p. 111).

[2] "Classical motion of quantum Gaussians," (1991).

[3] Rodney Yoder, *The Phase Space Formulation of Quantum Mechanics and the Problem of Negative Probabilities* (Reed College thesis, 1992).

1. Initial construction of the Gaussian packet. To express the circumstance that "x-measurement (performed at time $t = 0$ with an instrument of imperfect resolution) has shown the particle to reside in the vicinity of the point $x = a$" we write

$$P(x,0) \equiv |\psi(x,0)|^2 = \begin{cases} \text{some properly positioned and} \\ \text{shaped distribution function} \end{cases}$$

and notice that such a statement supplies only limited information about the structure of $\psi(x,0)$ itself:

$$\psi(x,0) = \sqrt{P(x,0)} \cdot e^{i\alpha(x,0)} \quad : \quad \text{phase factor remains at present arbitrary}$$

The phase factor has entered with simple innocence upon the stage, but is destined to play a leading role as the drama unfolds.

Whether we proceed from some tentative sense of the operating characteristics of instruments of finite resolution or seek only to model such statements in a concrete but analytically tractable way, it becomes fairly natural to look to the special case

$$P(x,0) = \tfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}[\frac{x-a}{\sigma}]^2} \tag{1}$$

The Gaussian on the right defines the "normal distribution" with

$$\begin{aligned} \text{mean} \; &: \; \langle x \rangle = a \\ \text{variance} \equiv (\text{uncertainty})^2 \; &: \; \langle (x-a)^2 \rangle = \sigma^2 \end{aligned}$$

and the associated wave function reads

$$\psi(x,0) = \left[\tfrac{1}{\sigma\sqrt{2\pi}}\right]^{\frac{1}{2}} e^{-\frac{1}{4}[\frac{x-a}{\sigma}]^2} \cdot e^{i\alpha(x,0)} \tag{2}$$

I write $\psi(x,0) \equiv (x|\psi)_0$ to draw attention to the fact that we have worked thus far in the x-representation. Passage to the momentum representation is accomplished

$$\begin{aligned} \varphi(p,0) \equiv (p|\psi)_0 &= \int (p|x)\, dx\, (x|\psi)_0 \\ &= \tfrac{1}{\sqrt{h}} \int e^{-\frac{i}{\hbar} px} \psi(x,0)\, dx \end{aligned} \tag{3}$$

but cannot be carried out in detail until the phase factor has been specified.[4] If, in the Gaussian case (2), we set $\alpha = 0$ then (3) gives

$$\varphi(p,0) = \left[\tfrac{1}{\lambda\sqrt{2\pi}}\right]^{\frac{1}{2}} e^{-\frac{1}{4}[\frac{p}{\lambda}]^2} \cdot e^{-\frac{i}{\hbar} ap} \quad \text{with} \quad \lambda \equiv \hbar/2\sigma \tag{4}$$

---

[4] The statement

$$\text{normalization of } P(x) \equiv |\psi(x)|^2 \;\Rightarrow\; \text{normalization of } Q(p) \equiv |\varphi(p)|^2$$

is, however, phase-insensitive, and is the upshot of Parseval's theorem: see P. Morse & H. Feshbach, *Methods of Theoretical Physics* (1953), p. 456.

whence

$$Q(p,0) = \tfrac{1}{\lambda\sqrt{2\pi}}e^{-\frac{1}{2}[\frac{p}{\lambda}]^2} \tag{5}$$

which is again normal, but centered at the origin of p-space:

$$\begin{aligned}\langle p\rangle &= 0\\ \langle (p-0)^2\rangle &= \lambda^2\end{aligned}$$

In

$$\sigma\lambda = \Delta x\cdot\Delta p = \tfrac{1}{2}\hbar \tag{6}$$

we have encountered an instance of optimal compliance with the Heisenberg uncertainty principle: $\Delta x\cdot\Delta p \geq \frac{1}{2}\hbar$.

To achieve arbitrary placement of the origin of the normal distribution in momentum space—i.e., to achieve

$$(5) \quad\longrightarrow\quad Q(p,0) = \tfrac{1}{\lambda\sqrt{2\pi}}e^{-\frac{1}{2}[\frac{p-b}{\lambda}]^2}$$

—it might appear most natural in place of (4) simply to write

$$\varphi(p,0) = \big[\tfrac{1}{\lambda\sqrt{2\pi}}\big]^{\frac{1}{2}}e^{-\frac{1}{4}[\frac{p-b}{\lambda}]^2}\cdot e^{-\frac{i}{\hbar}a(p-b)}$$

But then

$$\psi(x,0) = \tfrac{1}{\sqrt{h}}\int e^{+\frac{i}{\hbar}px}\varphi(p,0)\,dp \tag{7}$$

is, according to *Mathematica*, a mess; to achieve a simpler result I tentatively omit the phase factor, writing

$$\begin{aligned}\psi(x,0) &= \tfrac{1}{\sqrt{h}}\int e^{+\frac{i}{\hbar}px}\Big\{\big[\tfrac{1}{\lambda\sqrt{2\pi}}\big]^{\frac{1}{2}}e^{-\frac{1}{4}[\frac{p-b}{\lambda}]^2}\Big\}\,dp\\ &= \big[\tfrac{1}{\sigma\sqrt{2\pi}}\big]^{\frac{1}{2}}e^{-\frac{1}{4}[\frac{x}{\sigma}]^2}\cdot e^{+\frac{i}{\hbar}bx} \quad\text{with}\quad \sigma\equiv\hbar/2\lambda\end{aligned}$$

The resulting distribution function

$$P(x,0) = \tfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}[\frac{x}{\sigma}]^2}$$

is (compare (5)) is again normal but centered at the origin in configuration space; it is precisely (1) with $a = 0$. These simple results indicate that if we desire to displace $P(x,0)$ and $Q(p,0)$ simultaneously from their respective origins we are going to have to learn how to "manage the mess." They bear directly upon the problem of the "launched Gaussian wavepacket," to which I turn in §5.

2. Standing motion of a Gaussian wavepacket. Dynamical evolution

$$|\psi)_0 \xrightarrow[\text{free particle}]{} |\psi)_t$$

is particularly easy to describe in the momentum representation (essentially because p and $\mathsf{H} = \frac{1}{2m}\mathsf{p}^2$ trivially commute); we have

$$\begin{aligned}\varphi(p,t) &= \int \underbrace{(p|e^{-\frac{i}{\hbar}(\mathsf{p}^2/2m)\,t}|q)}_{=\,e^{-\frac{i}{\hbar}(q^2/2m)\,t}\delta(q-p)}\varphi(q,0)\,dq\\ &= \varphi(p,0)\cdot e^{-\frac{i}{\hbar}(p^2/2m)\,t}\end{aligned}\tag{8}$$

The t-dependence of $\varphi(p,t)$ is so simple as to imply

$$Q(p,t) = Q(p,0)\tag{9}$$

Because the particle has been assume to move *freely*, almost nothing is going on in momentum space...but enough (see again (8)) to cause interesting things to happen in configuration space:

Working from

$$\psi(x,t) = \tfrac{1}{\sqrt{h}}\int e^{+\frac{i}{\hbar}px}\varphi(p,t)\,dp\tag{10}$$

with

$$\begin{aligned}\varphi(p,t) &= \Big\{\big[\tfrac{1}{\lambda\sqrt{2\pi}}\big]^{\frac{1}{2}}e^{-\frac{1}{4}[\frac{p}{\lambda}]^2}\cdot e^{-\frac{i}{\hbar}ap}\Big\}\cdot e^{-\frac{i}{\hbar}(p^2/2m)\,t}\\ &= \big[\tfrac{1}{\lambda\sqrt{2\pi}}\big]^{\frac{1}{2}}\exp\Big\{-\tfrac{[1+i(t/\tau)]p^2}{4\lambda^2}\Big\}\cdot e^{-\frac{i}{\hbar}ap}\end{aligned}$$

NOTE: We have introduced here the "natural time" $\tau \equiv \hbar m/2\lambda^2 = 2m\sigma^2/\hbar$

we find

$$\psi(x,t) = \Big[\tfrac{1}{\sigma[1+i(t/\tau)]\sqrt{2\pi}}\Big]^{\frac{1}{2}}\exp\Big\{-\tfrac{1}{4}\tfrac{(x-a)^2}{\sigma^2[1+i(t/\tau)]}\Big\}\tag{11}$$

It becomes natural at this point to define

$$\sigma(t) \equiv \sigma\sqrt{1+(t/\tau)^2}\tag{12}$$

Then $\sigma[1+i(t/\tau)] = \sigma(t)e^{i\arctan(t/\tau)}$ and $\frac{1}{\sigma^2[1+i(t/\tau)]} = [1-i(t/\tau)]/\sigma^2(t)$, so we have

$$P(x,t) = |\psi(x,t)|^2 = \tfrac{1}{\sigma(t)\sqrt{2\pi}}\exp\Big\{-\tfrac{1}{2}\big[\tfrac{x-a}{\sigma(t)}\big]^2\Big\}\tag{13}$$

which is normal, with fixed mean ($\langle x\rangle = a$: all t) and growing variance.

That the $\psi(x,t)$ of (11) is in fact a solution of the Schrödinger equation is confirmed by calculation; *Mathematica* informs us that

$$\begin{aligned}&\Big\{-\tfrac{\hbar^2}{2m}\big(\tfrac{\partial}{\partial x}\big)^2 - i\hbar\tfrac{\partial}{\partial t}\Big\}\psi(x,t)\\ &\qquad = \underbrace{(\hbar\tau - 2m\sigma^2)}_{0\text{ by definition of }\tau}\cdot(\text{complicated factor})\cdot e^{-\frac{1}{4}[\frac{x-a}{\sigma(t)}]^2}\end{aligned}$$

3. Relationship to the free particle propagator. It follows also from the definition $\tau \equiv 2m\sigma^2/\hbar$ that

$$\lim_{\sigma\downarrow 0}\sigma^2[1+i(t/\tau)] = i\hbar t/2m$$

so, working from (11), we have

$$\begin{aligned}\lim_{\sigma\downarrow 0}\psi(x,t) &\cong \sqrt{\sigma}\cdot\Big[\tfrac{2m}{i\hbar t\sqrt{2\pi}}\Big]^{\frac{1}{2}}\exp\Big\{\tfrac{i}{\hbar}\tfrac{m(x-a)^2}{2t}\Big\}\\ &= \sqrt{2\sigma\sqrt{2\pi}}\cdot\underbrace{\sqrt{\tfrac{m}{iht}}\,e^{\frac{i}{\hbar}\frac{m}{2}\frac{(x-a)^2}{t}}}_{K(x,t;a,0)\ \ :\ \text{ the } \textit{free particle propagator!}}\end{aligned} \tag{14}$$

We are inspired by this development to notice that

$$\begin{aligned}S(x,t;a,0) &\equiv \tfrac{\hbar}{i}\Big\{-\tfrac{1}{4}\tfrac{(x-a)^2}{\sigma^2[1+i(t/\tau)]}\Big\}\\ &\downarrow\\ &= \tfrac{m}{2}\tfrac{(x-a)^2}{t} \quad \text{in the limit } \sigma\downarrow 0\end{aligned} \tag{15}$$

is (according to *Mathematica*) a solution of the Hamilton-Jacobi equation

$$\tfrac{1}{2m}\big(\tfrac{\partial S}{\partial x}\big)^2 + \tfrac{\partial S}{\partial t} = 0 \tag{16}$$

and gives

$$\begin{aligned}\tfrac{\partial^2 S}{\partial x\partial a} &= \tfrac{\hbar}{2i\sigma^2[1+i(t/\tau)]}\\ &\downarrow\\ &= -\tfrac{m}{t} \quad \text{in the limit } \sigma\downarrow 0\end{aligned}$$

so (11) can be written

$$\psi(x,t) = \sqrt{2\sigma\sqrt{2\pi}}\cdot\sqrt{\tfrac{i}{\hbar}\tfrac{\partial^2 S}{\partial x\partial a}}\,e^{\frac{i}{\hbar}S(x,t;a,0)} \tag{17}$$

which is exact, but serves in the limit to cast useful light on the origin of (14). The factor $\sqrt{\frac{i}{\hbar}\frac{\partial^2 S}{\partial x\partial a}}\,e^{\frac{i}{\hbar}S}$ originates in early work of J. H. Van Vleck,[5] and has for more than half a century (i.e., since the invention of the Feynman formalism[6]) stood guard at the portal through which quantum and classical mechanics communicate. But it should, in this connection, be noticed that

[5] "The correspondence principle in the statistical interpretation of quantum mechanics," P. N. A. S. **14**, 178 (1928).

[6] See Chapter 7 of *Pauli Lectures on Physics: Volume 6. Selected Topics in Field Quantization*. This is the English translation (1973) of material which Pauli presented at a research seminar held at the ETH in Zürich during the academic year 1950-51.

the function $S(x,t;a,0)$ introduced at (15) cannot arise from orthodox classical mechanics, for it is *complex*—this we might emphasize by writing

$$S(x,t;a,0)=\hbar\tfrac{1}{4}\big[\tfrac{x-a}{\sigma(t)}\big]^2\cdot\big\{\tfrac{t+i\tau}{\tau}\big\}$$

—and its real/imaginary parts become entangled by the non-linearity of the Hamilton-Jacobi equation (16).

The factor $\sqrt{2\sigma\sqrt{2\pi}}$ which intrudes at (14) derives from this circumstance: the propagator $K(x,t;a,0)$ is the solution of the Schrödinger equation which has by design the property that

$$\lim_{t\downarrow 0}K(x,t;a,0)=\delta(x-a)$$

while (11) refers to the solution which can (in a certain limit, and somewhat informally) be said to evolve from $\sqrt{\delta(x-a)}$; we have

$$\begin{aligned}&\int \text{Gaussian}\,dx=1\quad\text{with}\quad \text{Gaussian}\equiv\tfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}[\frac{x}{\sigma}]^2}\\ &\int\sqrt{\text{Gaussian}}\,dx=\sqrt{2\sigma\sqrt{2\pi}}\end{aligned}$$

which captures the analytical essence of the situation. Pursuing this remark in finer detail: it follows from (11) that

$$\begin{aligned}\psi(x,0)=\Big[\tfrac{1}{\sigma\sqrt{2\pi}}\Big]^{\frac{1}{2}}\exp\Big\{-\tfrac{1}{4}\tfrac{(x-a)^2}{\sigma^2}\Big\}&=\sqrt{\text{Gaussian}}\\ &=\sqrt{\delta(x-a)}\quad\text{in the limit }\sigma\downarrow 0\end{aligned}\tag{18}$$

and by computation

$$\begin{aligned}\psi(x,t)&=\int K(x,t;y,0)\psi(y,0)\,dy\\ &=\int\sqrt{\tfrac{m}{iht}}\,e^{\frac{i}{\hbar}\frac{m}{2}\frac{(x-y)^2}{t}}\sqrt{\tfrac{1}{\sigma\sqrt{2\pi}}}e^{-\frac{1}{4}[\frac{y-a}{\sigma}]^2}\,dy\\ &=\sqrt{\tfrac{\tau}{4\pi\sigma^2 it}\tfrac{1}{\sigma\sqrt{2\pi}}}\int e^{-\frac{1}{4\sigma^2}\big\{[1-i(\tau/t)]y^2-2[a-i(\tau/t)x]y+[a^2-i(\tau/t)x^2]\big\}}\,dy\\ &=\sqrt{\tfrac{\tau}{4\pi\sigma^2 it}\tfrac{1}{\sigma\sqrt{2\pi}}}\sqrt{\pi\tfrac{4\sigma^2}{[1-i(\tau/t)]}}\exp\Big\{\tfrac{1}{4\sigma^2}\tfrac{[a-i(\tau/t)x]^2-[1-i(\tau/t)][a^2-i(\tau/t)x^2]}{[1-i(\tau/t)]}\Big\}\end{aligned}\tag{19}$$

we do in fact (after simplifications) recover precisely (11).

We observe finally that results obtained above by the process $\sigma\downarrow 0$ could equally well have been achieved by $t\uparrow\infty$; i.e., that (11) entails

$$\frac{1}{\sqrt{2\sigma\sqrt{2\pi}}}\cdot\psi(x,t)\longrightarrow K(x,t;a,0)\quad\text{for }t\gg\tau\tag{20}$$

For an electron we have

$$\begin{aligned} \tau &= \tfrac{2\hbar}{mc^2} \quad \text{if } \sigma \text{ set equal to the Compton length: } \sigma = \tfrac{\hbar}{mc} \\ &\sim 2.6 \times 10^{-21} \text{ seconds} \end{aligned}$$

so the process (20) is prompt; it becomes (by the factor $(\frac{1}{137})^2 = 5.3 \times 10^{-5}$) even more prompt if we set

$$\sigma = \text{classical electron radius} = \tfrac{e^2}{\hbar c} \cdot \text{Compton length}$$

But for a grain of sand (cubic millimeter of SiO_2, which has a mass of 2.65×10^{-3} gm.) we have

$$\begin{aligned} \tau &= 5.0 \times 10^{16} \text{ seconds if } \sigma = 10^{-4} \text{ centimeters (one micron)} \\ &\sim 1.6 \times 10^{9} \text{ years} \end{aligned}$$

which is certainly *not* prompt; even for such a precisely located smallish lump of macroscopic stuff (4.4×10^{-5} mole of quarz, with its more than 10^{20} internal degrees of freedom) the process (20) takes roughly the age of the universe to approach to completion.

4. Remarks concerning the "natural time" parameter. The (non-relativistic) quantum dynamics of a free particle supplies two dimensioned constants (m and $\hbar$), from which it is *not* possible to construct a "natural time." But particular solutions of the Schrödinger equation have distinctive *shapes*, and therefore supply characteristic assortments of numbers with the physical dimension of length; it was thus that we came to associate a

$$\text{"natural time" } \tau \equiv \tfrac{2m}{\hbar} \cdot \text{(minimal variance)} = \tfrac{2m\sigma^2}{\hbar}$$

with the Gaussian solutions (11). We were motivated by (20) to look to the numerical value assumed by τ in some characteristic cases, but (20) touches upon a fairly esoteric point; the simple essence of the matter was implicit already in (12), which written

$$\sigma^2(t) = \sigma^2[1 + (t/\tau)^2]$$

informs us that

$$\tau = \begin{cases} \text{time required for the variance } \sigma^2(t) \text{ to} \\ \text{grow to twice its initial/minimal value} \end{cases}$$

We have

$$\left.\begin{aligned} \sigma(t) &\cong ut \\ u &\equiv \tfrac{\hbar}{2m\sigma} \end{aligned}\right\} \quad \text{if} \quad t \gg \tau \tag{21}$$

For an electron $\hbar/2m = \text{cm}^2\,\text{sec}^{-1}$

$$\begin{aligned} u &= 5.8 \times 10^{8} \text{ cm sec}^{-1} \quad \text{if } \sigma = \text{Bohr radius} \\ &= \tfrac{1}{2}(c/137) = \tfrac{\text{orbital speed}}{2} \end{aligned}$$

To set $u = c$ is to obtain

$$\sigma = \tfrac{1}{2}(\text{Compton length})$$

which establishes the sense in which the quantum mechanical expansion of a primeval Gaussian mimics the expansion of the universe. Note also that by the Heisenberg uncertainty principle $\Delta p \sim \hbar/2\sigma = mu$; this elementary remark establishes a sense in which (21) is not at all surprising.

But from another point of view (21) *is* surprising: The "fundamental solution"[7] of the heat/diffusion equation $\nabla^2\psi = D\frac{\partial}{\partial t}\psi$ is Gaussian

$$\begin{aligned} k(x,t) &\equiv \tfrac{1}{\sqrt{4\pi t/D}} e^{-\frac{x^2}{4t/D}} \\ &= \tfrac{1}{\sigma(t)\sqrt{2\pi}} e^{-\frac{1}{2}[\frac{x}{\sigma(t)}]^2} \quad \text{with} \quad \sigma(t) = \sqrt{2t/D} \end{aligned}$$

and exposes this fact:

$$\sigma \sim t^{\frac{1}{2}} \quad : \quad \text{characteristic signature of a "diffusion processes"}$$

If, formally, we set $D = -i\frac{2m}{\hbar}$ then the heat equation becomes the free-particle Schrödinger equation, and $k(x,t)$ becomes precisely the propagator $K(x,t)$, as defined at (14). Looking back again to (11–13) we see that *it is the complexity of the construction* $\sigma\sqrt{1 + i(t/\tau)}$ that accounts for

$$\sigma \sim t^{\frac{1}{2}} \xrightarrow[\text{surprising}]{} \sigma \sim t$$

5. The "launched" Gaussian wavepacket. The Gaussian packet (11) sits in one place and grows slowly fat. By "launching" such a packet we place ourselves in position to model the quantum mechanical motion of a projectile, which departs (neighborhood of) the origin at time $t = 0$ with velocity v and with momenta which lie in the neighborhood Δp of $p = mv$. I discuss three distinct but equivalent routes to the construction of such a "launched Gaussian wavepacket."

It is an implication (set $a = 0$) of the argument that culminated in (11) that

$$\begin{aligned} \psi(x,t) &= \underbrace{\Big[\tfrac{1}{\sigma[1+i(t/\tau)]\sqrt{2\pi}}\Big]^{\frac{1}{2}} \exp\Big\{-\tfrac{1}{4}\tfrac{x^2}{\sigma^2[1+i(t/\tau)]}\Big\}}_{\text{Gaussian standing at } x=0} \\ &= \tfrac{1}{\sqrt{h}} \int e^{+\frac{i}{\hbar}px} \varphi(p,t)\,dp \\ &\qquad\qquad \varphi(p,t) = \Big\{\big[\tfrac{1}{\lambda\sqrt{2\pi}}\big]^{\frac{1}{2}} e^{-\frac{1}{4}[\frac{p}{\lambda}]^2}\Big\} \cdot e^{-\frac{i}{\hbar}(p^2/2m)\,t} \end{aligned}$$

[7] Also called the "source solution;" see p. 10 and Chapter 3 of D. V. Widder, *The Heat Equation* (1975).

The idea now is to *shift the momentum distribution* ($p \to p - p_0$ with $p_0 \equiv mv$) and then work backwards, writing

$$\begin{aligned}
\psi(x,t) &= \tfrac{1}{\sqrt{h}}\int e^{+\frac{i}{\hbar}px}\Big\{\big[\tfrac{1}{\lambda\sqrt{2\pi}}\big]^{\frac{1}{2}}\exp\big\{-\tfrac{1}{4}\big[\tfrac{p-p_0}{\lambda}\big]^2\big\}\Big\}\cdot e^{-\frac{i}{\hbar}(p^2/2m)\,t}\,dp\\
&= \big[\tfrac{1}{h\lambda\sqrt{2\pi}}\big]^{\frac{1}{2}}\int e^{-(Ap^2+2Bp+C)}\,dp
\end{aligned}$$

$$\begin{aligned}
A &\equiv \tfrac{1}{4\lambda^2}[1+i(t/\tau)]\\
B &\equiv -\tfrac{1}{4\lambda^2}[p_0+i\tfrac{m}{\tau}x]\\
C &\equiv \tfrac{1}{4\lambda^2}p_0^2
\end{aligned}$$

$$\begin{aligned}
&= \big[\tfrac{1}{h\lambda\sqrt{2\pi}}\big]^{\frac{1}{2}}\big[\tfrac{\pi}{A}\big]^{\frac{1}{2}}e^{\frac{B^2-AC}{A}}\\
&= \big[\tfrac{\pi}{h\lambda A\sqrt{2\pi}}\big]^{\frac{1}{2}}\exp\Big\{\tfrac{[p_0+i(m/\tau)x]^2-p_0^2[1+i(t/\tau)]}{4\lambda^2[1+i(t/\tau)]}\Big\}\\
&= \big[\tfrac{1}{\sigma[1+i(t/\tau)]\sqrt{2\pi}}\big]^{\frac{1}{2}}\exp\Big\{-\tfrac{x^2}{4\sigma^2[1+i(t/\tau)]}+\underbrace{\tfrac{i}{\hbar}\tfrac{p_0x-(p_0^2/2m)t}{1+i(t/\tau)}}_{\text{new term}}\Big\}
\end{aligned}\tag{22}$$

From

$$= \big[\text{etc.}\big]^{\frac{1}{2}}\exp\Big\{\tfrac{1}{4\sigma^2(t)}\big[-(x-vt)^2+i\tfrac{t}{\tau}(x^2-v^2\tau^2)\big]\Big\}\cdot\exp\Big\{\tfrac{i}{\hbar}\tfrac{p_0x}{1+(t/\tau)^2}\Big\}\tag{23}$$

it becomes clear that the associated probability density

$$P(x,t) = \tfrac{1}{\sigma(t)\sqrt{2\pi}}\exp\big\{-\tfrac{1}{2}\big[\tfrac{x-vt}{\sigma(t)}\big]^2\big\}\tag{24}$$

does as anticipated describe a Gaussian which drifts to the right with speed v, growing fat in the familiar way as it does so.

At time $t=0$ (22) gives

$$\psi(x,0) = \big[\tfrac{1}{\sigma\sqrt{2\pi}}\big]^{\frac{1}{2}}e^{-\frac{1}{4}[\frac{x}{\sigma}]^2}\cdot e^{\frac{i}{\hbar}p_0x}\tag{25}$$

which is a specialized instance of (2): set $a=0$ and $\alpha(x,0)=\frac{1}{\hbar}p_0x$. If we were able to argue on some grounds that *introduction of the phase factor* $e^{\frac{i}{\hbar}p_0x}$ *serves to "launch" the Gaussian packet* then we should be able to recover (22) by dynamical propagation of (25). Returning to (19) (where the same idea was developed in the case $p_0=0$), we have

$$\begin{aligned}
&\int K(x,t;y,0)\psi(y,0)\,dy\\
&\quad= \int\sqrt{\tfrac{m}{iht}}\,e^{\frac{i}{\hbar}\frac{m}{2}\frac{(x-y)^2}{t}}\sqrt{\tfrac{1}{\sigma\sqrt{2\pi}}}e^{-\frac{1}{4}[\frac{y}{\sigma}]^2}\cdot e^{iky}\,dy\quad\text{with } k\equiv p_0/\hbar\\
&\quad= \sqrt{\tfrac{\tau}{4\pi\sigma^2it}\tfrac{1}{\sigma\sqrt{2\pi}}}\int e^{-\frac{1}{4\sigma^2}\{[1-i(\tau/t)]y^2+2i[-2\sigma^2k+(\tau/t)x]y+[-i(\tau/t)x^2]\}}\,dy\\
&\quad= \sqrt{\tfrac{\tau}{4\pi\sigma^2it}\tfrac{1}{\sigma\sqrt{2\pi}}}\sqrt{\pi\tfrac{4\sigma^2}{[1-i(\tau/t)]}}\exp\Big\{-\tfrac{1}{4\sigma^2}\tfrac{[-2\sigma^2k+(\tau/t)x]^2+[1-i(\tau/t)][-i(\tau/t)x^2]}{[1-i(\tau/t)]}\Big\}\\
&\quad= \Big[\tfrac{1}{\sigma[1+i(t/\tau)]\sqrt{2\pi}}\Big]^{\frac{1}{2}}\exp\Big\{-\tfrac{x^2}{4\sigma^2[1+i(t/\tau)]}+i\tfrac{kx-(\sigma^2/\tau)k^2t}{1+i(t/\tau)}\Big\}
\end{aligned}$$

which (by $\sigma^2/\tau = \hbar/2m$ and $k = p_0/\hbar$) does in fact precisely reproduce (22). But how might we have *foreseen* that the adjustment

$$\Big[\tfrac{1}{\sigma\sqrt{2\pi}}\Big]^{\frac{1}{2}} e^{-\frac{1}{4}[\frac{x}{\sigma}]^2} \quad\longrightarrow\quad \Big[\tfrac{1}{\sigma\sqrt{2\pi}}\Big]^{\frac{1}{2}} e^{-\frac{1}{4}[\frac{x}{\sigma}]^2}\cdot e^{\frac{i}{\hbar}p_0 x} \tag{26}$$

serves to "launch" the standing Gaussian packet on the left? The question motivates me to pose a related question (which might have been posed already in connection with (2)): If $\psi(x,t)$ is a solution of the free particle Schrödinger equation $\psi_{xx} + i\kappa\psi_t = 0$ ($\kappa \equiv 2m/\hbar$), what condition is imposed upon $\alpha(x,t)$ by the requirement that $\psi\cdot e^{i\alpha}$ be also a solution? From

$$(\partial_x^2 + i\kappa\partial_t)\psi e^{i\alpha} = \big\{\underbrace{(\psi_{xx} + i\kappa\psi_t)}_{0} + (i\alpha_{xx} - \alpha_x^2 - \kappa\alpha_t)\psi + 2i\alpha_x\psi_x\big\}e^{i\alpha}$$

we see that necessarily

$$\alpha_{xx} + i\alpha_x^2 + i\kappa\alpha_t + 2\alpha_x\frac{\partial\log\psi}{\partial x} = 0$$

In the present application (see again (22))

$$\frac{\partial\log\psi}{\partial x} = -\frac{x}{2\sigma^2[1+i(t/\tau)]} = -\frac{mx}{\hbar\tau[1+i(t/\tau)]}$$

and *Mathematica* assures us that $\alpha(x,t) \equiv \frac{p_0x - (p_0^2/2m)t}{\hbar[1+i(t/\tau)]}$ does in fact satisfy

$$\alpha_{xx} + i\alpha_x^2 + i\tfrac{2m}{\hbar}\alpha_t - \tfrac{2mx}{\hbar\tau[1+i(t/\tau)]}\alpha_x = 0 \tag{27}$$

This result is confirmatory of (22), but provides disappointingly little insight into the mechanism that lies at the base of (26).

Many years ago I had occasion[8] to describe the structure of the group of transformations which send

$$\text{solutions} \quad\longrightarrow\quad \text{solutions of the free particle Schrödinger equation}$$

That work—in which the Gaussian "fundamental solution"

$$g(x,t) = \frac{1}{\sqrt{4\pi a t}}e^{-\frac{1}{4at}x^2} \quad\text{with}\quad a \equiv \frac{\hbar}{2m}$$

plays a central role—is by nature an elaborate generalization of the little argument just concluded, and *does* serve to provide insight into (26). But to venture down that seldom-traveled road would be ask too much of my reader; I turn now, therefore, to a simpler line of argument—a selected detail from that more comprehensive work—which captures the essence of the point at issue.

---

[8] APPELL, GALILEAN & CONFORMAL TRANSFORMATIONS IN CLASSICAL/QUANTUM FREE PARTICLE DYNAMICS (1976). I drew inspiration there from several sources, but especially from §6 of Widder's introductory chapter.

An inertial observer glides by, contemplating the physics (our §2) of a standing Gaussian. The following discussion proceeds from the remark that *to us, his Gaussian will appear to have been "launched."* Let us suppose that our inertial friend—so that we can continue to use $\{x,t\}$—uses $\{X,T\}$ to coordinate the points of spacetime, with

$$\begin{aligned} x &= X + vT \\ t &= T \end{aligned} \quad ; \text{ inversely} \quad \begin{aligned} X &= x - vt \\ T &= t \end{aligned}$$

He writes (see again (11))

$$\Psi(X,T) = \Big[\tfrac{1}{\sigma[1+i(T/\tau)]\sqrt{2\pi}}\Big]^{\frac{1}{2}} \exp\Big\{-\tfrac{1}{4}\tfrac{X^2}{\sigma^2[1+i(T/\tau)]}\Big\}$$

to describe his Gaussian-at-the-origin, and notes that $\Psi(X,T)$ is a solution of the Schrödinger equation, which he writes

$$\Big\{\big(\tfrac{\partial}{\partial X}\big)^2 + i\kappa\tfrac{\partial}{\partial T}\Big\}\Psi = 0$$

Those pronouncements, in our variables, read

$$\Psi(x-vt,t) = \Big[\tfrac{1}{\sigma[1+i(t/\tau)]\sqrt{2\pi}}\Big]^{\frac{1}{2}} \exp\Big\{-\tfrac{1}{4}\tfrac{(x-vt)^2}{\sigma^2[1+i(t/\tau)]}\Big\}$$

and (since $\frac{\partial}{\partial X} = \frac{\partial}{\partial x}$ and $\frac{\partial}{\partial T} = v\frac{\partial}{\partial x} + \frac{\partial}{\partial t}$)

$$\Big\{\big(\tfrac{\partial}{\partial x}\big)^2 + i\kappa\big[v\tfrac{\partial}{\partial x} + \tfrac{\partial}{\partial t}\big]\Big\}\Psi(x-vt,t) = 0$$

The latter equation does *not* possess free particle Schrödinger form, except in the trivial case $v = 0$. We observe, however, that we can multiply the latter equation by any non-zero factor without distroying its validity, and that we have at our disposal the operator identity ("shift rule")

$$\begin{aligned} e^{\varphi}\Big\{\big(\tfrac{\partial}{\partial x}\big)^2 + i\kappa\big[v\tfrac{\partial}{\partial x} + \tfrac{\partial}{\partial t}\big]\Big\} &= \Big\{\big(\tfrac{\partial}{\partial x} - \varphi_x\big)^2 + i\kappa\big[v\big(\tfrac{\partial}{\partial x} - \varphi_x\big) + \big(\tfrac{\partial}{\partial t} - \varphi_t\big)\big]\Big\}e^{\varphi} \\ &= \Big\{\big(\tfrac{\partial}{\partial x}\big)^2 + i\kappa\tfrac{\partial}{\partial t}\Big\}e^{\varphi} \\ &\qquad + \Big\{\big(-2\varphi_x + i\kappa v\big)\tfrac{\partial}{\partial t} + \big(\varphi_x^2 - \varphi_{xx} - i\kappa v\varphi_x - i\kappa v\varphi_t\big)\Big\}e^{\varphi} \end{aligned}$$

The implication is that $\psi(x,t) \equiv e^{\varphi}\cdot\Psi(x-vt,t)$ *will* satisfy the free particle Schrödinger equation if

$$\big(-2\varphi_x + i\kappa v\big) = \varphi_x^2 - \varphi_{xx} - i\kappa v\varphi_x - i\kappa v\varphi_t = 0$$

These conditions are readily seen to entail

$$\begin{aligned} \varphi(x,t) &= \tfrac{1}{2}i\kappa vx - \tfrac{1}{4}i\kappa v^2 t + \text{constant} \\ &= \tfrac{i}{\hbar}\big\{mvx - \tfrac{1}{2}mv^2 t\big\} + \varphi_0 \end{aligned}$$

so we obtain

$$\begin{aligned}\psi(x,t) &= \Big[\tfrac{1}{\sigma[1+i(t/\tau)]\sqrt{2\pi}}\Big]^{\frac{1}{2}}\exp\Big\{-\tfrac{1}{4}\tfrac{(x-vt)^2}{\sigma^2[1+i(t/\tau)]}\Big\}\cdot e^{\frac{i}{\hbar}\{mvx-\frac{1}{2}mv^2t\}}\\ &= \Big[\tfrac{1}{\sigma[1+i(t/\tau)]\sqrt{2\pi}}\Big]^{\frac{1}{2}}\exp\Big\{\tfrac{1}{\sigma^2[1+i(t/\tau)]}\Big[-\tfrac{x^2}{4}+\tfrac{i}{\hbar}\big(mvx-\tfrac{1}{2}mv^2t\big)\Big]\Big\}\end{aligned}$$

which precisely reproduces (22). This line of argument traces (26) to the circumstance that *Galilian covariance of the Schrödinger equation requires that the wavefunction acquires a factor* when transformed. We might on dimensional grounds argue that the only factors available (at $t=0$) have the form

$$e^{(\text{numeric})\frac{i}{\hbar}mvx}$$

and that detailed analysis has served only to establish that

$$\text{numeric} = 1$$

6. The associated Wigner distributions. In 1932 E. P. Wigner had occasion[9] to pull from his hat (or perhaps from that of Leo Szilard) the definition

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

and by 1949 J. E. Moyal[10] had traced the fact that the so-called "Wigner distribution function" is so richly endowed with wonderful properties to the circumstance that $P_\psi(x,p)$ stands in Weyl correspondence with the density matrix

$$hP_\psi(x,p) \;\xleftrightarrow[\text{Weyl}]{}\; |\psi)(\psi|$$

My present objective is to construct a description of $P_{\text{gaussian}}(x,p)$.

Looking first to the simple static Gaussian[11]

$$\psi(x) = \big[\tfrac{1}{\sigma\sqrt{2\pi}}\big]^{\frac{1}{2}}e^{-\frac{1}{4}[\frac{x-a}{\sigma}]^2}$$

Mathematica supplies

$$\begin{aligned}P_\psi(x,p) &= \tfrac{2}{h}e^{-\frac{1}{2}\{[\frac{x-a}{\sigma}]^2+[\frac{p}{\lambda}]^2\}}\\ &= \tfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}[\frac{x-a}{\sigma}]^2}\cdot\tfrac{1}{\lambda\sqrt{2\pi}}e^{-\frac{1}{2}[\frac{p}{\lambda}]^2} \qquad (28)\\ &= |\psi(x)|^2\cdot|\varphi(p)|^2\end{aligned}$$

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

[10] "Quantum mechanics as a statistical theory," Proc. Camb. Phil. Soc. **45**, 92 (1949). For historical comments, additional references and a summary of the mathematical/physical details see §6 in "Status and some ramifications of Ehrenfest's theorem" (1998).

[11] This is (2), with $\alpha = 0$.

where (as henceforth) I have allowed myself to make tacit use of (6): $\lambda = \hbar/2\sigma$. These *general* properties of the Wigner distribution

$$\iint P_\psi(x,p)\,dxdp = 1 \tag{29.1}$$

$$\int P_\psi(x,p)\,dp = |\psi(x)|^2 \quad\text{and}\quad \int P_\psi(x,p)\,dx = |\varphi(p)|^2 \tag{29.2}$$

$$P_\psi(x,p) \text{ is } \textit{bounded}\text{: } |P_\psi(x,p)| \le 2/h \tag{29.3}$$

pertain transparently to the $P_\psi(x,p)$ encountered at (28), but this "most distinctive quirk" of the Wigner distribution

$$P_\psi(x,p) \text{ is not precluded from assuming } \textit{negative values} \tag{29.4}$$

does not: (28) describes a function which is everywhere positive, and is therefore not a "quasi-distribution" but a true distribution function; it is, in plain words, a *bivariate normal distribution, defined on phase space*.

Working now from the launched Gaussian wavepacket (22)—written

$$\begin{aligned}\psi(x,t) =&\Psi\cdot\exp\Big\{\tfrac{1}{1+i\theta}\big[-\tfrac{1}{4\sigma^2}x^2+\tfrac{i}{\hbar}p_0x\big]\Big\}\\ &\Psi\equiv\big[\tfrac{1}{\sigma[1+i\theta]\sqrt{2\pi}}\big]^{\frac{1}{2}}\exp\Big\{-\tfrac{i}{\hbar}\tfrac{(p_0^2/2m)t}{1+i\theta}\Big\}\end{aligned}$$

with $\theta\equiv t/\tau$—we have

$$\begin{aligned}P_\psi(x,p) &= \tfrac{2}{h}|\Psi|^2\int\exp\Big\{\tfrac{1}{1-i\theta}\big[-\tfrac{1}{4\sigma^2}(x+\xi)^2-\tfrac{i}{\hbar}p_0(x+\xi)\big]\Big\}e^{2\frac{i}{\hbar}p\xi}\\ &\qquad\qquad\cdot\exp\Big\{\tfrac{1}{1+i\theta}\big[-\tfrac{1}{4\sigma^2}(x-\xi)^2+\tfrac{i}{\hbar}p_0(x-\xi)\big]\Big\}\,d\xi\\ &= \tfrac{2}{h}|\Psi|^2\int e^{-[A\xi^2+2B\xi+C]/2\hbar\sigma^2(t)}\,d\xi \;:\; \begin{cases}A\equiv\hbar\\ B\equiv i[\hbar x\theta+2p_0\sigma^2-2p\sigma^2(t)]\\ C\equiv\hbar x^2-4p_0x\sigma^2\theta\end{cases}\\ &= \underbrace{\tfrac{2}{h}|\Psi|^2\cdot\big[\tfrac{\pi}{\hbar}2\hbar\sigma^2(t)\big]^{\frac{1}{2}}}_{=\,\frac{2}{h}}\exp\Big\{\tfrac{-[\hbar x\theta+2mv\sigma^2-2p\sigma^2(t)]^2-\hbar[\hbar x^2-4mvx\sigma^2\theta]}{2\hbar^2\sigma^2(t)}\Big\}\\ &\qquad\qquad\exp\Big\{-\tfrac{2\hbar mv^2\sigma^2t\theta}{2\hbar^2\sigma^2(t)}\Big\}\\ &= \tfrac{2}{h}\exp\Big\{-x^2\tfrac{1}{2\sigma^2}+xp\tfrac{\theta}{\sigma\lambda}-p^2\tfrac{(1+\theta^2)}{2\lambda^2}+p\tfrac{mv}{\lambda^2}-\tfrac{m^2v^2}{2\lambda^2}\Big\}\\ &= \tfrac{2}{h}\exp\Big\{-\tfrac{1}{2}\big[\tfrac{x}{\sigma}-\theta\cdot\tfrac{p}{\lambda}\big]^2-\tfrac{1}{2}\big[\tfrac{p-mv}{\lambda}\big]^2\Big\} &(30.1)\\ &= \tfrac{2}{h}\exp\Big\{-\tfrac{1}{2}\big[\tfrac{x-vt}{\sigma}-\theta\cdot\tfrac{p-mv}{\lambda}\big]^2-\tfrac{1}{2}\big[\tfrac{p-mv}{\lambda}\big]^2\Big\} &(30.2)\end{aligned}$$

It is gratifying to observe that (30) gives back (28) when we simultaneously turn off the drift (set $v=0$), execute a spatial displacement $x\to x-a$ and go to the origin of time ($\theta=0$). And that (as one readily verifies) the implied marginal distribution $\int P_\psi(x,p)\,dp$ reproduces precisely (24). We observe—as we observed already at (28)—that the prefactor

$$\tfrac{2}{h} \text{ can be written } \tfrac{1}{\sigma\sqrt{2\pi}}\cdot\tfrac{1}{\lambda\sqrt{2\pi}}, \text{ and is time-independent}$$

and that—surprisingly?—it is not $\sigma(t)$ but $\sigma=\sigma(0)$ which enters into the design both of the prefactor and of the exponent. We note finally that (30) describes a Wigner distribution which is *non-negative*, everywhere and always.

7. Minimal dispersion, and the emergence of correlation. From the statement

$$\Delta x\,\Delta p \geq \tfrac{1}{2}\hbar \quad : \quad \text{all states } |\psi)$$

of the Heisenberg uncertainty principle it becomes natural to ask: For what states $|\psi)$ is equality achieved? It is in response to this natural question that one is led to the so-called "states of minimal dispersion" which are, as it emerges, Gaussian; it is largely (though by no means exclusively) from this fact that physicists acquire their special interest in "Gaussian wavepackets."

In 1930 Schrödinger observed[12] that if A and B are self-adjoint operators associated with *any* pair of observables, then

$$\begin{aligned}(\Delta A)^2(\Delta B)^2 &\geq \Big\langle\frac{\mathsf{AB}-\mathsf{BA}}{2i}\Big\rangle^2 + \Big\{\Big\langle\frac{\mathsf{AB}+\mathsf{BA}}{2}\Big\rangle - \langle\mathsf{A}\rangle\langle\mathsf{B}\rangle\Big\}^2 \\ &\geq \Big\langle\frac{\mathsf{AB}-\mathsf{BA}}{2i}\Big\rangle^2\end{aligned} \tag{31}$$

with equality if and only if it is simultaneously the case that

- the "quantum correlation coefficient" $\Big\langle\frac{\mathsf{AB}+\mathsf{BA}}{2}\Big\rangle - \langle\mathsf{A}\rangle\langle\mathsf{B}\rangle$ vanishes, and
- the vectors $\big\{\mathsf{A}-\langle\mathsf{A}\rangle\big\}|\psi)$ and $\big\{\mathsf{B}-\langle\mathsf{B}\rangle\big\}|\psi)$ are parallel

These conditions, in the special case $\mathsf{A}\to\mathsf{x}$ and $\mathsf{B}\to\mathsf{p}$, are readily shown[13] to entail—compare (25)—

$$\psi(x) \equiv (x|\psi) = \big[\tfrac{1}{\sigma\sqrt{2\pi}}\big]^{\frac{1}{2}} e^{-\frac{1}{4}[\frac{x-\langle x\rangle}{\sigma}]^2} e^{\frac{i}{\hbar}\langle p\rangle x} \quad \text{with} \quad \sigma = \Delta x$$

Turning now (from "quantum pre-dynamics") to quantum *dynamics*, one sets

$$\psi(x,0) = \text{minimal dispersion wavepacket}$$

[12] The detailed argument is reviewed in §2 of "Status and some ramifications of Ehrenfest's theorem" (1998). See also the discussion in §7.1 of Max Jammer's *The Conceptual Development of Quantum Mechanics* (1966).

[13] See David Bohm, *Quantum Mechanics* (1951) §10.10; David Griffiths, *Introduction to Quantum Mechanics* (1995) §3.4.2 or virtually any other good quantum text.

and looks to the dispersive properties of

$$\psi(x,t) = \int K(x,t;y,0)\psi(y,0)\,dy \tag{32}$$

It is a familiar fact[14] that the ground state of an oscillator is Gaussian

$$\psi_{\text{oscillator}}(x,0) = \big[\tfrac{m\omega}{\hbar\pi}\big]^{\frac{1}{4}} e^{-\frac{m\omega}{2\hbar}x^2} \quad : \quad \text{evidently } \sigma = \Delta x = \sqrt{\hbar/2m\omega}$$

and, because we are talking here about an eigenstate, we know even without appeal to the propagator that

$$\psi_{\text{oscillator}}(x,t) = \big[\tfrac{m\omega}{\hbar\pi}\big]^{\frac{1}{4}} e^{-\frac{m\omega}{2\hbar}x^2}\cdot e^{-\frac{i}{\hbar}[\frac{1}{2}\hbar\omega]t}$$

We have here exhibited a particular state of a particular system with the property that *minimal dispersiveness is persistent*. If, more generally (and more interestingly), we take

$$\psi(x,0) = \big[\tfrac{m\omega}{\hbar\pi}\big]^{\frac{1}{4}} e^{-\frac{m\omega}{2\hbar}(x-A)^2} \quad : \quad \text{displaced copy of the groundstate}$$

and work from (32), we at length[15] obtain

$$\psi(x,t) = \big[\tfrac{m\omega}{\hbar\pi}\big]^{\frac{1}{4}} e^{-\frac{m\omega}{2\hbar}(x-A\cos\omega t)^2}\cdot e^{-if(x,t)}$$

where $f(x,t) \equiv \frac{\omega t}{2} + \frac{m\omega}{2\hbar}\big[2xA\sin\omega t - \frac{1}{2}A^2\sin 2\omega t\big]$ is periodic; then

$$\begin{aligned}|\psi(x,t)|^2 &= \big[\tfrac{m\omega}{\hbar\pi}\big]^{\frac{1}{2}} e^{-\frac{m\omega}{\hbar}(x-A\cos\omega t)^2}\\ &= \text{Gaussian, sloshing rigidly back and forth, with amplitude } A\end{aligned}$$

The results just summarized are of interest not least because they are so *atypical*: generally, minimal dispersiveness is *not* persistent. Indeed, we saw already in §4 that

$$\Delta x\,\Delta p = \tfrac{1}{2}\hbar \xrightarrow[\text{free particle}]{} \tfrac{1}{2}\hbar\sqrt{1+(t/\tau)^2} \tag{33}$$
$$\tau \equiv \tfrac{\Delta x}{u} \text{ with } u = \tfrac{\Delta p}{m}$$

I propose now to discuss, from a rather novel point of view, how (33) comes about.[16]

---

[14] See Bohm, §13.8.

[15] See QUANTUM MECHANICS (1967), Chapter 2, pp. 89–95 for the details. One needs to know that the oscillator propagator can be described

$$K_{\text{oscillator}}(x,t,y,0) = \sqrt{\frac{m\omega}{2\pi i\hbar\sin\omega t}}\exp\Big\{-\frac{m\omega}{2i\hbar\sin\omega t}\big[(x^2+y^2)\cos\omega t - 2xy\big]\Big\}$$

[16] Closely related remarks can be found in Bohm's §10.8, but he was obliged to proceed without the assistance of a tool which (as I hope to demonstrate) lends itself particularly well to such discussion—the notion of a Wigner distribution.

Equation (30) has the form

$$\begin{aligned} P_\psi(x,p;t) &= \tfrac{2}{h}e^{\text{polynomial of degree 2 in variables } x \text{ and } p} \\ &= \tfrac{2}{h}e^{\text{quadratic in ``displaced variables'' } X\equiv x-vt \text{ and } P\equiv p-mv} \end{aligned}$$

and can be notated

$$P_\psi(x,p;t) = \tfrac{1}{2\pi}\sqrt{\det\mathbb{A}}\; e^{-\frac{1}{2}\boldsymbol{\xi}^\mathsf{T}\mathbb{A}\,\boldsymbol{\xi}} \tag{34}$$

with

$$\boldsymbol{\xi}\equiv\begin{pmatrix} X \\ P\end{pmatrix} \quad\text{and}\quad \mathbb{A}\equiv\begin{pmatrix} \frac{1}{\sigma^2} & -\theta\frac{1}{\sigma\lambda} \\ -\theta\frac{1}{\sigma\lambda} & \frac{1+\theta^2}{\lambda^2}\end{pmatrix} \tag{35}$$
$$\Downarrow$$
$$\tfrac{1}{2\pi}\sqrt{\det\mathbb{A}} = \tfrac{1}{2\pi}\tfrac{1}{\sigma\lambda} = \tfrac{2}{h}$$

I wrote (34) to establish contact with the standard *theory of bivariate normal distributions*, and digress now to review the bare essentials of that subject.[17] We have these Gaussian integral formulæ:

$$\tfrac{1}{2\pi}\sqrt{ac-b^2}\iint\begin{pmatrix}1\\x\\y\\x^2\\xy\\y^2\end{pmatrix}e^{-\frac{1}{2}(ax^2+2bxy+cy^2)}dxdy = \begin{pmatrix}1\\0\\0\\+c/(ac-b^2)\\-b/(ac-b^2)\\+a/(ac-b^2)\end{pmatrix}$$

subject only to the conditions $a>0$, $c>0$ and $ac-b^2>0$. Evidently

$$\begin{aligned} F(x,y) &\equiv \tfrac{1}{2\pi}\sqrt{ac-b^2}e^{-\frac{1}{2}(ax^2+2bxy+cy^2)} \\ &= \tfrac{1}{2\pi}\sqrt{\det\mathbb{A}}\,\exp\left\{-\tfrac{1}{2}\begin{pmatrix}x\\y\end{pmatrix}^\mathsf{T}\underbrace{\begin{pmatrix}a & b\\ b & c\end{pmatrix}}_{\mathbb{A}}\begin{pmatrix}x\\y\end{pmatrix}\right\} \end{aligned}$$

can be understood to describe a bivariate distribution, with vanishing means

$$\iint\begin{pmatrix}x\\y\end{pmatrix}F(x,y)\,dxdy = \begin{pmatrix}0\\0\end{pmatrix}$$

and (trivially centered) second moments—variances and "covariances"—given by

$$\begin{pmatrix}\sigma_{xx}^2 & \sigma_{xy}^2\\ \sigma_{yx}^2 & \sigma_{yy}^2\end{pmatrix}\equiv\iint\begin{pmatrix}xx & xy\\ yx & yy\end{pmatrix}F(x,y)\,dxdy = \tfrac{1}{\det\mathbb{A}}\begin{pmatrix}+c & -b\\ -b & +a\end{pmatrix} = \mathbb{A}^{-1}$$

---

[17] See Harald Cramér, *Mathematical Methods of Statistics* (1946), §21.12.

In terms of the latter, we have

$$\begin{aligned}
\mathbb{A} &= \tfrac{1}{\sigma_{xx}^2\sigma_{yy}^2-\sigma_{xy}^2\sigma_{yx}^2}\begin{pmatrix} \sigma_{yy}^2 & -\sigma_{xy}^2 \\ \\ -\sigma_{yx}^2 & \sigma_{xx}^2 \end{pmatrix} \\
&= \tfrac{1}{\sigma_x^2\sigma_y^2(1-\rho^2)}\begin{pmatrix} \sigma_y\sigma_y & -\rho\sigma_x\sigma_y \\ -\rho\sigma_x\sigma_y & \sigma_x\sigma_x \end{pmatrix} = \begin{pmatrix} \sigma_x\sigma_x & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y\sigma_y \end{pmatrix}^{-1} \\
&= \tfrac{1}{1-\rho^2}\begin{pmatrix} \frac{1}{\sigma_x^2} & -\rho\frac{1}{\sigma_x\sigma_y} \\ \\ -\rho\frac{1}{\sigma_x\sigma_y} & \frac{1}{\sigma_y^2} \end{pmatrix}
\end{aligned} \tag{36}$$

where I have abandoned the phoney distinction between σ_{xy} and σ_{yx}, adopted the notational simplifications $\sigma_{xx} \to \sigma_x$ and $\sigma_{yy} \to \sigma_y$, and introduced[18] the (dimensionless) "correlation coefficient"

$$\rho \equiv \frac{\sigma_{xy}^2}{\sigma_x\sigma_y} \tag{37}$$

The distribution $F(x,y)$ is constant on curves

$$\begin{pmatrix} x \\ y \end{pmatrix}^{\mathsf{T}}\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \equiv \boldsymbol{x}^{\mathsf{T}}\mathbb{A}\,\boldsymbol{x} = \text{constant}$$

which (by $\det\mathbb{A} = \frac{1}{\sigma_x^2\sigma_y^2(1-\rho^2)} > 0$, which entails $-1 \le \rho \le +1$) are in fact concentric ellipses, centered at the origin, and of a shape/orientation which is set by the spectral properties of $\mathbb{A}$. One has

$$\begin{aligned}
\text{eigenvalues} &= \frac{(\sigma_x^2+\sigma_y^2) \pm \sqrt{(\sigma_x^2-\sigma_y^2)^2+4\rho^2\sigma_x^2\sigma_y^2}}{2\sigma_x^2\sigma_y^2(1-\rho^2)} \\
&= \tfrac{1}{2}(\text{sum}) \pm \tfrac{1}{2}(\text{difference}) \\
&\qquad (\text{sum}) = \tfrac{\sigma_x^2+\sigma_y^2}{\sigma_x^2\sigma_y^2(1-\rho^2)} = \tfrac{1}{1-\rho^2}\left[\tfrac{1}{\sigma_x^2}+\tfrac{1}{\sigma_y^2}\right] = \text{tr}\mathbb{A}
\end{aligned} \tag{38}$$

To discover the orientation of the ellipse we draw upon the fact[19] that every 2×2 symmetric matrix can be written

$$\begin{aligned}
&\begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix}^{\mathsf{T}}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix} \\
&\quad = \begin{pmatrix} \frac{1}{2}(\text{sum}) + \frac{1}{2}(\text{difference})\cos 2\vartheta & -\frac{1}{2}(\text{difference})\sin 2\vartheta \\ -\frac{1}{2}(\text{difference})\sin 2\vartheta & \frac{1}{2}(\text{sum}) - \frac{1}{2}(\text{difference})\cos 2\vartheta \end{pmatrix}
\end{aligned}$$

[18] See Cramér, p. 265.

[19] See §1 of "Non-standard applications of Mohr's construction" (1998).

Working from (36) and (38), we therefore have

$$\left.\begin{aligned}\sin 2\vartheta &= \frac{2\rho\sigma_x\sigma_y}{\sqrt{(\sigma_x^2-\sigma_y^2)^2+4\rho^2\sigma_x^2\sigma_y^2}}\\ \cos 2\vartheta &= \frac{\sigma_x^2-\sigma_y^2}{\sqrt{(\sigma_x^2-\sigma_y^2)^2+4\rho^2\sigma_x^2\sigma_y^2}}\end{aligned}\right\}\tag{39}$$

Looking to what (38) and (39) have to say in some special cases: if $\rho = 0$ then it is obvious from (36) but also an implication of (38) that the eigenvalues of $\mathbb{A}$ are $1/\sigma_x^2$ and $1/\sigma_y^2$, while (39) gives $\vartheta = 0$ (no rotation is required to diagonalize the already-diagonal matrix: the principal axes coincide with the coordinate axes): no cross-term appears in the exponent, so $F(x, y)$ factors

$$F(x, y) = f(x)\cdot g(y)$$

($f(x)$ and $g(y)$ are, as it happens, both Gaussian) and the random variables x and y have become "independent." At $\rho = \pm 1$ the right side of (38) becomes singular, but useful information can be obtain by setting $\rho = 1 - \epsilon$ (else $\rho = -1 + \epsilon$) and studying the *approach* to singularity (limit $\epsilon \downarrow 0$); we have

$$\text{eigenvalues} = \begin{cases}\frac{1}{\epsilon}\frac{\sigma_x^2+\sigma_y^2}{2\sigma_x^2\sigma_y^2} + \cdots \longrightarrow \quad \infty\\ \quad\frac{1}{\sigma_x^2+\sigma_y^2} + \cdots \longrightarrow \frac{1}{\sigma_x^2+\sigma_y^2}\end{cases}\tag{40}$$

and so obtain concentric elongated ellipses of vanishing area—in short: a line, of

$$\begin{aligned}\text{slope} = \tan\vartheta &= \pm\sqrt{\tfrac{1-\cos 2\vartheta}{1+\cos 2\vartheta}}\\ &= \pm\tfrac{\sigma_y}{\sigma_x}\end{aligned}\tag{41}$$

(I will not linger to resolve the sign ambituity); the random variables x and y have become perfectly (anti)correlated, which is to say "dependent."[20]

Returning with these classic resources to our physical problem: comparison of (35) with (36) gives

$$\begin{aligned}\sigma_X^2 &= \qquad\ \tfrac{1}{1-\rho^2}\sigma^2\\ \sigma_P^2 &= \tfrac{1}{1+\theta^2}\tfrac{1}{1-\rho^2}\lambda^2\\ \sigma_X\sigma_P &= \qquad\tfrac{\rho}{\theta}\tfrac{1}{1-\rho^2}\sigma\lambda\end{aligned}$$

[20] We are in position now to use 1st and 2nd moment data, whether obtained from experimental observation or from some theoretical $\mathcal{F}(x, y)$, to construct a "bivariate normal distribution of best fit." I have no present reason to pursue the details (a variant of the program will be taken up in a companion essay), but see Cramér's Chapter 21 or the resources in *Mathematica*'s "Multinormal Distribution" package to gain a sense of what lies down that much-traveled road.

The first two of those equations entail $\sigma_X\sigma_P = \frac{1}{\sqrt{1+\theta^2}}\frac{1}{1-\rho^2}\sigma\lambda$, which upon comparison with the third gives

$$\rho = \tfrac{\theta}{\sqrt{1+\theta^2}} \qquad \text{i.e.,} \qquad \tfrac{1}{1-\rho^2} = 1+\theta^2 \tag{42}$$

whence

$$\left.\begin{aligned} \sigma_X &= \sigma\sqrt{1+\theta^2} \\ \sigma_P &= \lambda \end{aligned}\right\} \tag{43}$$

which precisely reproduce (12) and its momental companion. But while (12) refers to a property of the Gaussian wavepacket $\psi(x,t)$ described by (11), the equations (43) refer to a more complex (but physically equivalent) object: the $P_\psi(x,p;t)$ of (30).

We are, by (42), placed in position to state (in reference to the quantum dynamics of a free Gaussian wavepacket) that the correlation coefficient

$$\rho(t) = \tfrac{t/\tau}{\sqrt{1+(t/\tau)^2}} = \begin{cases} (t/\tau) + \cdots & \longrightarrow 0 \quad \text{at } \ t = 0 \\ 1 - \frac{1}{2}(\tau/t)^2 + \cdots & \longrightarrow 1 \quad \text{for } t \gg \tau\text{; i.e., as } t \uparrow \infty \end{cases}$$

That x-measurements and p-measurements yield results which are initially *un*correlated is plausible enough, but that there exists any sense in which

asymptotically in time, p-measurements become
redundant *with x-measurements!*

is counterintuitive... though susceptible to interpretation. Asymptotically in time, we have $P = \text{slope}\cdot X$; i.e.,

$$p - mv = \text{slope}\cdot(x - \langle x\rangle)$$

while it follows from (41) by (43) that

$$\text{slope} = \tfrac{\lambda}{\sigma\sqrt{1+\theta^2}} \sim \tfrac{\lambda\tau}{\sigma t} = \tfrac{m}{t} \longrightarrow 0$$

The distribution—which at finite times was constant on concentric ellipses with moving centers—has in the limit $t \uparrow \infty$ become axially symmetric about the "line of zero slope" $p = mv$; we have become unable to say anything useful concerning the likely outcome of an x-measurement, but retain our initial ability to speak usefully about p-measurements. Looking to (30.1) for direct support of this conclusion, we are reminded that, *to the extent that $P_\psi(x,p;t)$ has acquired axial symmetry, it has lost its normalizability*; the thing to notice is that the marginal distribution

$$\int \Big\{\text{right side of (30.1)}\Big\}\,dx = \tfrac{1}{\lambda\sqrt{2\pi}}\exp\Big\{-\tfrac{1}{2}\big[\tfrac{p-mv}{\lambda}\big]^2\Big\}$$

is θ-independent: it supports our conclusion at all finite times, and therefore does so also in the limit.

Looking again to (30), we notice that

$$P_\psi(x,p;t) \text{ is maximal at the moving point } \begin{cases} x(t) = vt \\ p(t) = mv \end{cases} : \text{ constant}$$
$$\text{maximal value is at all times given by } \tfrac{2}{h}$$

and that the maximal value is in fact greatest-possible for a (bivariate) Wigner distribution. The valuation drops to $\frac{2}{h}e^{-\frac{1}{2}}$ (60% of maximum) on the moving curve

$$\left[\tfrac{x-vt}{\sigma} - \theta\cdot\tfrac{p-mv}{\lambda}\right]^2 + \left[\tfrac{p-mv}{\lambda}\right]^2 = 1 \tag{44}$$

which at $t = 0$ reads

$$\left[\tfrac{x}{\sigma}\right]^2 + \left[\tfrac{p-mv}{\lambda}\right]^2 = 1 \tag{45}$$

The latter equation describes a ellipse (center displaced a distance mv up the p-axis), with principal axes parallel to the coordinate axes, and

$$\text{area} = \pi\sigma\lambda = \tfrac{h}{4} \tag{46}$$

Generally, the area of the ellipse $\boldsymbol{x}^{\mathsf{T}}\mathbb{A}\,\boldsymbol{x} = 1$ $(\det\mathbb{A} > 0)$ can be described

$$\text{area} = \pi\sqrt{\text{product of eigenvalues of } \mathbb{A}^{-1}} = \pi\sqrt{\det\mathbb{A}^{-1}} = \pi/\sqrt{\det\mathbb{A}}$$

In connection with the general theory of bivariate Gaussian distributions we have already had occasion to remark that $\det\mathbb{A}^{-1} = \sigma_x^2\sigma_y^2(1-\rho^2)$, and from this it now follows (see again (35)) that the area of the ellipse (44) is given (not just initially but) at *all* times by (44): the deformation (45) $\xrightarrow[t]{}$ (44) is, in fact, *area-preserving*.

The "incompressibility of dynamical phase flow"[21] is in classical mechanics a celebrated implication of the fact that dynamical phase flow is canonical, and has important applications especially to (classical) statistical mechanics. It is a notion not common to quantum mechanics, but has here been encountered in connection with a particular application of the "phase space formulation" of quantum mechanics. I now show that (45) $\xrightarrow[t]{}$ (44) can, in fact, be extracted from the *classical* dynamics of a free particle:

Generally, if

$$\left.\begin{matrix} x_0 \\ p_0 \end{matrix}\right\} \xrightarrow[t]{} \begin{cases} x = f(x_0,p_0;t) \\ p = g(x_0,p_0;t) \end{cases}$$

describes an invertible (not necessarily dynamical) t-parameterized map, and if

$$\varphi(x_0,p_0) = 0$$

[21] Liouville's theorem: see §9-8 in H. Goldstein, *Classical Mechanics* (1980).

serves implicitly to describe a curve $\mathcal{C}_0$ inscribed on phase space, then we write

$$\varphi(x,p)=0 \xrightarrow[t]{} \varphi(x,p;t)\equiv\varphi(f^{-1}(x,p;t),g^{-1}(x,p;t)) \tag{47}$$

to described the flow-induced deformation $\mathcal{C}_0\to\mathcal{C}_t$ of the curve. The dynamical flow generated by the free particle Hamiltonian can be described

$$\left.\begin{matrix} x_0\\ p_0\end{matrix}\right\} \xrightarrow[t]{} \left\{\begin{matrix} x = x_0+\frac{1}{m}p_0t\\ p=p_0\end{matrix}\right. \quad : \quad \text{“free particle sheer”}$$

and as an instance of (47) induces

$$\big[\tfrac{x}{\sigma}\big]^2+\big[\tfrac{p-mv}{\lambda}\big]^2=1 \xrightarrow[t]{} \big[\tfrac{x-(p/m)t}{\sigma}\big]^2+\big[\tfrac{p-mv}{\lambda}\big]^2=1 \tag{48}$$

But

$$\begin{aligned}\big[\tfrac{x-vt}{\sigma}-\theta\cdot\tfrac{p-mv}{\lambda}\big] &= \tfrac{1}{\sigma}\big[x-vt-\tfrac{\sigma}{\tau\lambda}(p-mv)t\big]\\ &\qquad\qquad \tfrac{\sigma}{\tau\lambda}=\tfrac{1}{m}\\ &=\big[\tfrac{x-(p/m)t}{\sigma}\big]\quad\text{after simplifications}\end{aligned}$$

so at (48) we have in fact recovered (45) $\xrightarrow[t]{}$ (44): see the following figure.

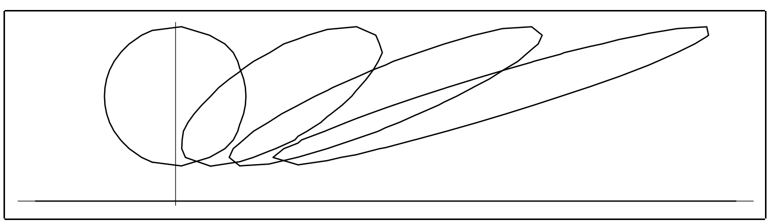

FIGURE 1: *Mechanism responsible for the dynamical development of correlation. The figure derives from (48), in which I have set* σ, λ, m *and* v *all equal to unity, and* $t=\{0,1,2,3\}$. *A similar graphic appears on p. 204 of Bohm's text, but is claimed by him to refer only to the* classical *physics 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.

8. Quantum analog of Liouville's theorem. It is a frequently reenforced lesson of computational experience (unsupportable, so far as I am aware, by any direct appeal to physical intuition) that the dialog between classical and quantum mechanics becomes uniquely felicitous when the Hamiltonian depends at most quadratically upon its arguments, and it is for this reason that most experienced physicists would, I anticipate, hold that the quantum/classical confluence developed in the preceding section and summarized in the figure is "not very surprising." My objective here will be to show how the phase space formalism can be used to make such a train of argument clear and precise.

In ordinary quantum mechanics the motion (in the Schödinger picture) of the density matrix $\boldsymbol{\rho} \equiv |\psi)(\psi|$ can be described

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

The preceding equation can be shown without much difficulty[22] to stand in Weyl correspondence with the following fairly awesome equation:

$$\tfrac{\partial}{\partial t}P_\psi(x,p;t) = \tfrac{2}{\hbar}\sin\Big\{\tfrac{\hbar}{2}\Big(\frac{\partial}{\partial x_H}\frac{\partial}{\partial p_P} - \frac{\partial}{\partial x_P}\frac{\partial}{\partial p_H}\Big)\Big\}H(x,p)P_\psi(x,p;t) \tag{49}$$

in connection with which it is to be understood that $H(x,p) \underset{\text{Weyl}}{\longleftrightarrow} \mathsf{H}$; that

$$\begin{aligned}\sin\big\{\tfrac{\hbar}{2}\big(\text{etc.}\big)\big\} &\equiv \big\{\tfrac{\hbar}{2}\big(\text{etc.}\big)\big\} - \tfrac{1}{3!}\big\{\tfrac{\hbar}{2}\big(\text{etc.}\big)\big\}^3 + \tfrac{1}{3!}\big\{\tfrac{\hbar}{2}\big(\text{etc.}\big)\big\}^5 - \cdots\\ &= \text{differential operator of infinite order}\end{aligned}$$

and that $\frac{\partial}{\partial x_H}$ sees the x-dependence of H but is blind to that of P, etc. The point to which I would draw attention is that (49) simplifies greatly

$$\begin{aligned}&\downarrow\\ &= \Big(\frac{\partial}{\partial x_H}\frac{\partial}{\partial p_P} - \frac{\partial}{\partial x_P}\frac{\partial}{\partial p_H}\Big)HP_\psi \quad \text{ if } H \text{ is } \textit{quadratic} \text{ in its arguments}\\ &= \frac{\partial H}{\partial x}\frac{\partial P}{\partial p} - \frac{\partial P}{\partial x}\frac{\partial H}{\partial p}\\ &= [H, P_\psi] \quad : \quad \text{Poisson bracket}\end{aligned}$$

Equation (49) is, within the phase space formalism, the analog of—and conveys precisely the same information as—the Schrödinger equation. It can be written

$$\tfrac{\partial}{\partial t}P_\psi(x,p;t) = [H, P_\psi] + \text{power series in } \hbar$$

(which suggests the special utility of the formalism to semi-classical lines of argument) and in quadratic cases assumes precisely the structure of the Liouville equation

$$\tfrac{\partial}{\partial t}D(x,p;t) = [H, D\,] \tag{50}$$

[22] See Chapter 3, p. 110 of QUANTUM MECHANICS (1967).

—the upshot of Liouville's theorem, according to which the local density of any population of points sprinkled on phase space is, with respect to any flowing point, constant:

$$\begin{aligned}\tfrac{d}{dt}D\big(x(t),p(t);t\big) &= \tfrac{\partial D}{\partial t} + \dot{p}\tfrac{\partial D}{\partial p} + \dot{x}\tfrac{\partial D}{\partial x}\\ &= \tfrac{\partial D}{\partial t} - \tfrac{\partial H}{\partial x}\tfrac{\partial D}{\partial p} + \tfrac{\partial H}{\partial p}\tfrac{\partial D}{\partial x}\\ &= \tfrac{\partial D}{\partial t} - [H,D\,]\\ &= 0 \end{aligned}\tag{51}$$

Liouville's theorem is, as previously remarked, an expression ultimately of the "incompressibility of classsical phase flow." The quantum dynamical equation (49) can in this light be construed to be a *modified analog* of Liouville's equation, an assertion that "phase flow in the quantum world is—except for quadratic Hamiltonians—*not* incompressible, but squishy."

Returning now from generalities to the Gaussian's of immediate interest, we observe that the Wigner distribution encountered at (30)—which we are in position now to write

$$P(x,p;t) = \tfrac{1}{\sigma\sqrt{2\pi}}\tfrac{1}{\lambda\sqrt{2\pi}}\exp\Big\{-\tfrac{1}{2}\big[\tfrac{x-(p/m)t}{\sigma}\big]^2 - \tfrac{1}{2}\big[\tfrac{p-mv}{\lambda}\big]^2\Big\}\tag{52}$$

—is, almost trivially, a solution of the Liouville equation (50). And that, since

$$H_{\text{free}} = \tfrac{1}{2m}p^2$$

is quadratic, means that it is (to say the same thing another way) a solution also of the "Schrödinger equation" (49).

9. A paradox, and its resolution. In accounting thus for the "quantum/classical confluence" we have achieved almost too much of a good thing. For (52) contains no reference to $\hbar$, and describes a correctly normalized Gaussian solution of (49/50) *for all values of σ and λ—even those that stand in violation of the Heisenberg/minimality conditions $\sigma\lambda \geq \frac{1}{2}\hbar$.* But the uncertainty principle admits in quantum mechanics of *no* exceptions!

My first approach to the resolution of this paradoxical development hinges on the seldom-remarked fact that *only conditionally can a solution $P(x,p;t)$ of (49) be interpreted to be the Wigner distribution associated with some wave function $\psi(x,t)$.*

When one is—as we were, at beginning of §6—introduced to the "forward Wigner construction"

$$\psi(x) \xrightarrow[\text{Wigner}]{} P_\psi(x,p) \equiv \tfrac{2}{h}\int \psi^*(x+\xi)\, e^{2\frac{i}{\hbar}p\,\xi}\,\psi(x-\xi)\,d\xi\tag{53}$$

and told that it lies at the foundation of an elaborately developed "phase space formulation of quantum mechanics" which is entirely equivalent to the standard formulation (but in some respects more attractive), it becomes natural to ask: "How does one pass backwards across the bridge that interconnects the two formalisms? How does one describe $\psi(x) \xleftarrow[\text{Wigner}]{} P_\psi(x,p)$?" Curiously, the literature known to me is silent on this point. The desultory attention which I had given to the problem over a span of forty years had served only to convince me that, since the solution entailed a kind of de-convolution, "extraction of a functional square root," it was probably hard. I was therefore amazed when my colleague Mark Beck (who has an *experimental* interest in the phase space formalism) referred casually, in private conversation, to this elegant solution:[23]

By Fourier transformation of (53) we have

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

Select a point a at which $\int P_\psi(a,p)\,dp = \psi^*(a)\,\psi(a) \neq 0$.[24] Set $\hat{\xi} = a - x$ to obtain

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

which by notational adjustment $2x - a \mapsto x$ gives

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

Evidently $|\psi(0)|^2 = \int P_\psi(0,p)\,dp$, so we have Beck's formula—the "backward Wigner construction"

$$\psi(x) = \frac{e^{i\alpha}}{\sqrt{\int P_\psi(0,p)\,dp}}\cdot\int P_\psi(\tfrac{x}{2},p)\,e^{\frac{i}{\hbar}px}\,dp \xleftarrow[\text{Wigner}]{} P_\psi(x,p) \tag{54}$$

The wavefunction ψ is delivered to us already normalized, fixed to within specification of an unphysical phase factor.

Insertion of the $P(x,p;0)$ of (52)[25] into Beck's formula (I discard the phase

---

[23] The following account of "Beck's trick" is taken from §6 of "Status and some ramifications of Ehrenfest's theorem" (1998).

[24] 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 $a = 0$.

[25] It is simply to avoid irrelevant notational complexity that I have set $t = 0$.

factor) gives

$$\psi(x,0) = \frac{\int \frac{1}{\sigma\sqrt{2\pi}}\frac{1}{\lambda\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left[\frac{x}{2\sigma}\right]^2-\frac{1}{2}\left[\frac{p-mv}{\lambda}\right]^2\right\}e^{\frac{i}{\hbar}px}\,dp}{\sqrt{\int \frac{1}{\sigma\sqrt{2\pi}}\frac{1}{\lambda\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left[\frac{p-mv}{\lambda}\right]^2\right\}dp}}$$

$$= \frac{\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{4}\left[\frac{1}{2\sigma^2}+\frac{2\lambda^2}{\hbar^2}\right]x^2\right\}e^{\frac{i}{\hbar}mvx}}{\sqrt{\frac{1}{\sigma\sqrt{2\pi}}}}$$

This function is supposed to be "already normalized," but *is* normalized—and in fact reproduces precisely (25)—if and only if $\left[\frac{1}{2\sigma^2}+\frac{2\lambda^2}{\hbar^2}\right]=\frac{1}{\sigma^2}$, and this is equivalent to the *stipulation that the minimality condition* $\sigma\lambda=\hbar/2$ *be satisfied*.

We are brought thus to this resolution of our paradox:

Each of the $\{\sigma,\lambda\}$-parameterized functions (52) is a solution of the Liouville equation (50). And each, therefore, is a solution of the "Schrödinger equation" (49). But it is possible to set up an association of the form

$$P(x,p;t)\xleftrightarrow[\text{Wigner}]{\hspace{2cm}}\psi(x,t)\text{ if and only if }\sigma\lambda=\tfrac{1}{2}\hbar \tag{55}$$

Quantum mechanics enforces the minimality condition, which was seen already at (46) to be interpretable as a statement that certain ellipses (those of Figure 1) have area $=\frac{1}{4}\hbar$. This development is superficially reminiscent of the principle used by Planck to "quantize" the classical mechanics of an oscillator, and of the subsequent "Bohr-Sommerfeld quantization condition," and might reward closer study.

I discuss now an alternative approach to resolution of our paradox. We recall that (to within a dimensionally enforced factor of h, and as was first appreciated by Moyal[10])

$$\text{Wigner distribution }P(x,p)\xleftrightarrow[\text{Weyl}]{\hspace{2cm}}\text{density matrix }\boldsymbol{\rho} \tag{56}$$

and that in the general (or "mixed") case the density matrix

$$\begin{aligned}\boldsymbol{\rho} &= \sum\text{statistically weighted projection operators}\\ &= \sum p_k|\psi_k)(\psi_k|\\ &\downarrow\\ &= \sum\text{single term of unit weight}\\ &= |\psi)(\psi| \qquad \text{in the "pure case"}\end{aligned}$$

An algebraic characterization of the mixed/pure distinction is provided by the statement

$$\text{tr}\,\boldsymbol{\rho}^2\begin{cases}=1 & \text{in the "pure" case}\\ <1 & \text{in the "mixed" case}\end{cases} \tag{57}$$

(Notice in this connection that $0 < \text{tr}$ is automatic.) Weyl transform theory supplies the information that if

$$\mathbf{A} \xleftrightarrow[\text{Weyl}]{} A(x,p) \quad \text{and} \quad \mathbf{B} \xleftrightarrow[\text{Weyl}]{} B(x,p)$$

then[26]

$$\text{tr}\,\mathbf{AB} = \tfrac{1}{h}\iint A(x,p)B(x,p)\,dxdp$$

It follows that if—generalizing remarks presented at the beginning of §6—we write

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

then (57) can, within the phase space formalism, be expressed

$$h\iint P(x,p)P(x,p)\,dxdp \begin{cases} = 1 & \text{in the ``pure'' case} \\ < 1 & \text{in the ``mixed'' case} \end{cases} \tag{58}$$

If, in particular, we take $P(x,p)$ to be given (at time t) by the right side of (52) then the $\iint$ is trivial, and we (at all times) obtain

$$h\frac{1}{4\pi\sigma\lambda} \begin{cases} = 1 & \text{in the ``pure'' case} \\ < 1 & \text{in the ``mixed'' case} \end{cases}$$

The gratifying implication is that

$$\sigma\lambda \begin{cases} < \frac{1}{2}\hbar & \text{is quantum mechanically precluded} \\ = \frac{1}{2}\hbar & \text{in the ``pure'' case} \\ > \frac{1}{2}\hbar & \text{in the ``mixed'' case} \end{cases} \tag{59}$$

Only in the pure case is it sensible to search for an "associated wavefunction $\psi(x,t)$" (Beck has taught us how to conduct that search), and only in that case are we authorized to adopt the more emphatic notation $P_\psi(x,p;t)$. In the mixed case we expect to be able to write something like

$$\begin{aligned} \text{right side of (52)} &= \sum\!\!\!\!\!\!\int p_k P_{\psi_k}(x,p;t) \\ &\ \ \updownarrow \text{Weyl} \\ &= \text{spectral resolution of } \boldsymbol{\rho} \end{aligned}$$

The problem thus posed will, as it relates to "fat Gaussians," be solved in §12. We touch here upon a particular instance of an important general problem to which I hope to return in the near future. Our "paradox" has, in any event, been laid neatly—and informatively—to rest.

[26] See QUANTUM MECHANICS (1967), Chapter 2, p. 109.

10. Non-Gaussian wavepackets. The "wavepacket" concept is very general, and admits in principle of infinitely many realizations. So also is the notion of a "bivariate (quasi)-distribution." The "phase space formulation" of quantum mechanics serves to fuse those two primitive notions.

Talk about "wavepackets" tends to degenerate quickly into illustrative talk about "Gaussian wavepackets," not because the packets encountered in laboratory situations are known to be Gaussian, but simply because the function e^{-x^2} is (in all relevant respects) analytically so tractable. Similarly, talk about "bivariate distributions" tends to focus upon properties of the "bivariate normal distribution," but in the latter connection "analytical tractability" is a lucky accident; the deeper reason is that, in many applications, normal distributions acquire (from the central limit theorem) special importance. The phase space formalism permits that "special importance" to be shipped back into quantum mechanical discourse (or does it? That is the question which, as indicated in my introductory remarks, serves to motivate my present work.)

My objectives here will be to emphasize—working all the while within the bounds of "analytical tractability"—that many of the methods/results which we have used/obtained in reference to Gaussians are in fact not specific to Gaussians, and to show that some of the information brought thus to light is of physical/mathematical interest in its own right.

Let $P(x,p)$ be *any* real-valued function of the indicated arguments, subject only to the normalization condition

$$\iint P(x,p)\,dxdp = 1 \tag{60}$$

The side-condition $P(x,p) \geq 0$ (all x and p) would, if imposed, cause $P(x,p)$ to become a "bivariate distribution function," but will *not* be imposed. We might, for example, take $P(x,p)$ to be given by any of the following expressions:

$$\left.\begin{aligned}
P_0(x,p) &\equiv +\tfrac{1}{2\pi\cdot\sigma\lambda}e^{-\frac{1}{2}z}\\
P_1(x,p) &\equiv -\tfrac{1}{2\pi\cdot\sigma\lambda}e^{-\frac{1}{2}z}\big\{1-z\big\}\\
P_2(x,p) &\equiv +\tfrac{1}{2\pi\cdot\sigma\lambda}e^{-\frac{1}{2}z}\big\{1-2z+\tfrac{1}{2}z^2\big\}\\
&\ \ \vdots\\
P_n(x,p) &\equiv (-)^n\tfrac{1}{2\pi\cdot\sigma\lambda}e^{-\frac{1}{2}z}\big\{L_n(z)\big\}
\end{aligned}\right\} \tag{61}$$

where
$$z \equiv \big[\tfrac{x}{\sigma}\big]^2 + \big[\tfrac{p}{\lambda}\big]^2$$

and $L_n(z) \equiv \frac{1}{n!}e^z\big(\frac{d}{dz}\big)^n e^{-z}z^n$ is the Laguerre polynomial of order n. We are assured by *Mathematica* that each of those functions does in fact (all positive σ and λ) satisfy (60), and are informed additionally that in each case

$$h\iint P(x,p)P(x,p)\,dxdp = 1 \quad \text{if and only if} \quad \sigma\lambda = \tfrac{1}{2}\hbar$$

—none of which remains surprising when I confess that I borrowed (61) from an antique discussion[27] of the phase space formulation of the quantum theory of an oscillator. When the minimality condition *is* satisfied, $P_0(x,p)$ reproduces precisely (28), which when (consistently with minimality) we set

$$\sigma = \sqrt{\tfrac{\hbar\omega}{2m}}\,\tfrac{1}{\omega} \quad\text{and}\quad \lambda = \sqrt{\tfrac{\hbar\omega}{2m}}\,m$$

reads and entails

$$\begin{aligned}
P_0(x,p) &= \tfrac{2}{h}e^{-\frac{1}{2}z} \quad\text{with}\quad z = \tfrac{2m}{\hbar\omega}\big\{(\omega x)^2 + (p/m)^2\big\}\\
&\Big\downarrow \text{backward Wigner, according to Beck}\\
\psi_0(x) &= \sqrt{\tfrac{1}{\sigma\sqrt{2\pi}}}e^{-\frac{1}{4}[\frac{x}{\sigma}]^2}\\
&= \big[\tfrac{m\omega}{\hbar\pi}\big]^{\frac{1}{4}}e^{-\frac{m\omega}{2\hbar}x^2} \quad : \quad \text{oscillator groundstate (see again §7)}
\end{aligned}$$

Similarly

$$\begin{aligned}
P_n(x,p) &= (-)^n\tfrac{2}{h}e^{-\frac{1}{2}z}L_n(z)\\
&\Big\updownarrow \text{Wigner}\\
\psi_n(x) &= \sqrt{\tfrac{1}{n!\sigma\sqrt{2\pi}}}e^{-\frac{1}{4}[\frac{x}{\sigma}]^2}H_n(\tfrac{x}{\sigma}) \quad : \quad n^{\text{th}}\text{ oscillator eigenstate}
\end{aligned}$$

so the functions (61) are actually not at all esoteric, and would be "well-known" if only the phase space formalism were! It is interesting to notice that oscillator eigenfunctions become localized "by action of the spring." If we used the spring to assemble a copy of the ground state, but at time $t = 0$ *snipped the spring*, we would be returned to precisely the familiar physics of §1. But...

If we used spring to assemble an *excited* oscillator state and then got similarly busy with our snips, we would at times $t > 0$ find ourselves watching the evolution of a *non-Gaussian wavepacket*. And (to return to my main point) we would find, moreover, that our former methods still served: we would, on the basis of the *classical* flow pattern, construct[28]

$$P_n(x,p;t) \equiv P_n(x-(p/m)t,\, p-mv)$$

---

[27] QUANTUM MECHANICS (1967), Chapter 3, pp. 116–120. My original sources were U. Uhlhorn, Arkiv für Fysik **11**, 87 (1956) §5; M. S. Bartlett & J. E. Moyal, Proc. Camb. Phil. Soc. **45**, 545 (1949) and the appendix to G. A. Baker, Jr., Phys. Rev. **109**, 2198 (1958).

[28] Notice that we could have adopted this modified definition

$$z \equiv \big[\tfrac{x-a}{\sigma}\big]^2 + \big[\tfrac{p-b}{\lambda}\big]^2$$

without changing the force of preceding discussion. I have set $a = 0$ and $b - mv$ in order to "lauch" our "dissociated oscillator packets."

We would observe that (since $H = \frac{1}{2m}p^2$ for $t > 0$) the argument that gave (51) still pertains, and that (since the Hamiltonian is quadratic) it is therefore automatic that $P_n(x, p; t)$ satisfies the quantum dynamical equation (49).

The statements (59) pertain not just to $P_0(x, p)$ but to *each* of the quasi-distributions $P_n(x, p)$, each of which is—subject only to the proviso $\sigma\lambda \geq \frac{1}{2}\hbar$ —quantum mechanically unexceptionable. It becomes interesting in this light to notice that

$$P_n(x, p) \text{ is } \textit{classically precluded} \text{ except in the case } n = 0$$

The elementary reason for this state of affairs is made evident by Figure 2:

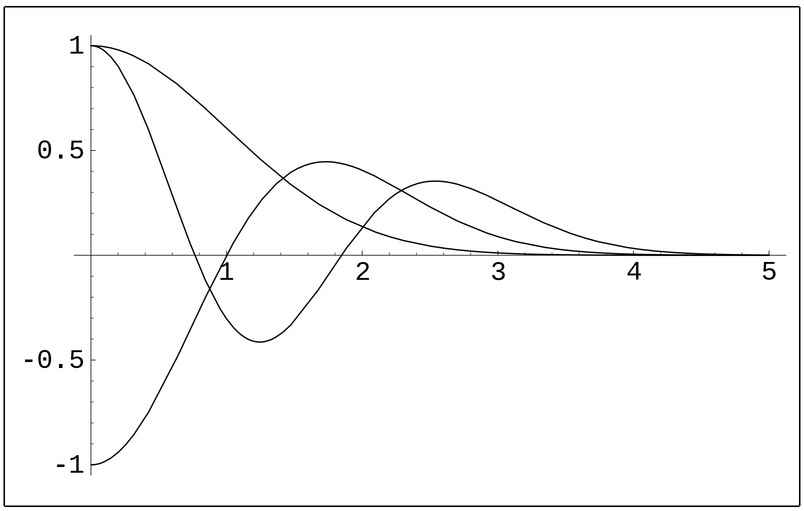

FIGURE 2: *Graphs of*

$$\left.\begin{aligned} F_0(z) &\equiv +e^{-\frac{1}{2}z} \\ F_1(z) &\equiv -e^{-\frac{1}{2}z}\{1 - z\} \\ F_2(z) &\equiv +e^{-\frac{1}{2}z}\{1 - 2z + \tfrac{1}{2}z^2\} \end{aligned}\right\} \quad \textit{vs.} \quad \sqrt{z}$$

$F_n(z)$ crosses the axis n times, with the consequence that each of the excited state functions $P_n(x, p)$ $\{n = 1, 2, \ldots\}$ becomes negative *on concentric elliptical rings drawn on classical phase space. This is the reason for the "quasi-distribution" terminology, and the reason that such distributions do not admit of classical interpretation.*

Venturing now a bit farther afield, we notice that the function

$$f(x) \equiv \frac{1}{\sqrt{2\alpha}}\,\mathrm{sech}(\tfrac{x}{\alpha}) \tag{62}$$

are localized/normalized, and might plausibly be taken to describe the initial design of a wavepacket; $|f(x)|^2$ looks, in fact, very "Gaussian" when plotted (see Figure 3), and $f(x)$ shares with Gaussians the uncommon property of being self-inversive under Fourier transformation:

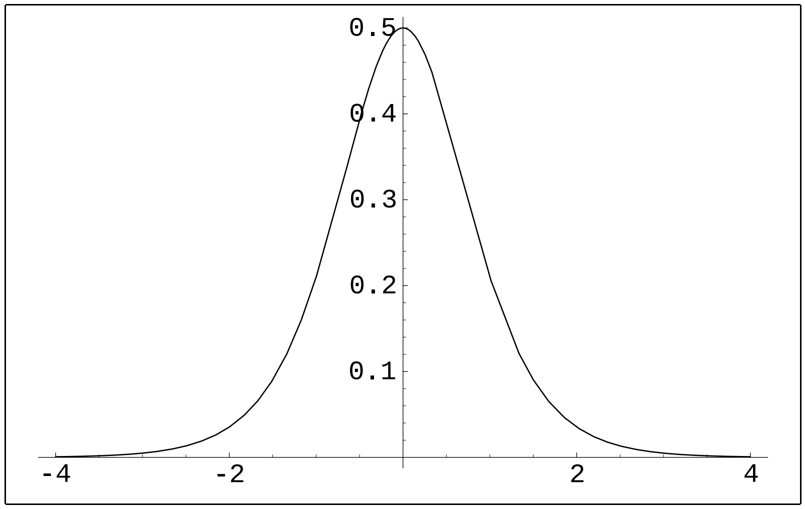

FIGURE 3: *Plot of* $|f(x)|^2 = \frac{1}{2\alpha}\,\text{sech}^2(\frac{x}{\alpha})$ *in the case* $\alpha = 1$.

$$
\begin{aligned}
g(x) \equiv \big[\tfrac{1}{\sigma\sqrt{2\pi}}\big]^{\frac{1}{2}} e^{-\frac{1}{4}(\frac{x}{\sigma})^2} \xrightarrow[\text{Fourier}]{} \tilde{g}(p) &\equiv \tfrac{1}{\sqrt{h}}\int e^{-\frac{i}{\hbar}px} g(x)\,dx\\
&= \big[\tfrac{1}{\lambda\sqrt{2\pi}}\big]^{\frac{1}{2}} e^{-\frac{1}{4}(\frac{p}{\lambda})^2} \quad : \quad \lambda \equiv \hbar/2\sigma\\
&\updownarrow \text{compare}\\
f(x) \equiv \tfrac{1}{\sqrt{2\alpha}}\,\text{sech}(\tfrac{x}{\alpha}) \xrightarrow[\text{Fourier}]{} \tilde{f}(p) &\equiv \tfrac{1}{\sqrt{h}}\int e^{-\frac{i}{\hbar}px} f(x)\,dx\\
&= \tfrac{1}{\sqrt{2\beta}}\,\text{sech}(\tfrac{p}{\beta}) \quad : \quad \beta \equiv \tfrac{4}{\pi}\cdot\hbar/2\alpha
\end{aligned}
$$

Because $|f(x)|^2$ and $|\tilde{f}(p)|^2$ are even functions of their respective arguments, it is immediate that

$$\langle x^{\text{odd}}\rangle = \langle p^{\text{odd}}\rangle = 0$$

but when we look to the even moments it becomes clear that the sech-packet places us near the outer limits of "analytical tactability." After a certain amount of experimentation (many complaints from *Mathematica*) I have, however, come up with this trick:

$$
\begin{aligned}
\int \tfrac{1}{2\alpha}\,\text{sech}^2(\tfrac{x}{\alpha})\cdot\cosh(sx)\,dx &= \langle x^0\rangle + \tfrac{s^2}{2!}\langle x^2\rangle + \tfrac{s^4}{4!}\langle x^4\rangle + \cdots\\
&= \tfrac{\pi s\alpha}{4}\big\{\tan\big[\tfrac{\pi s\alpha}{4}\big] + \cot\big[\tfrac{\pi s\alpha}{4}\big]\big\}\\
&= 1 + \tfrac{s^2}{2!}\tfrac{\pi^2\alpha^2}{12} + \tfrac{s^4}{4!}\tfrac{\pi^4\alpha^4}{240} + \cdots
\end{aligned}
$$

from which we conclude that $\Delta x = \sqrt{\pi^2/12}\alpha$ and $\Delta p = \sqrt{\pi^2/12}\beta$:

$$\Delta x\,\Delta p = \tfrac{\pi^2}{12}\alpha\beta = \tfrac{2\pi}{3}\cdot\tfrac{1}{2}\hbar = 2.0944\cdot\tfrac{1}{2}\hbar \tag{63}$$

We had no reason to expect the sech^2 distribution to be minimally dispersive, and it is not. Evidently the normal distribution with the same variance is

$$\frac{1}{\alpha\sqrt{\pi^3/6}}\exp\Big\{-\tfrac{1}{2}\Big[\frac{x}{\alpha\sqrt{\pi^2/12}}\Big]^2\Big\} \tag{64}$$

but if we sought the normal distribution which "best approximates" the sech^2 distribution it would be natural to seek to

$$\text{minimize}\int\Big[\tfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}(\frac{x}{\sigma})^2}-\tfrac{1}{2\alpha}\,\text{sech}^2\big(\tfrac{x}{\alpha}\big)\Big]^2dx$$

The integral, however, appears to be intractable. This train of thought did, however, lead me to set

$$\int\Big[\tfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}(\frac{x}{\sigma})^2}\Big]^2dx = \tfrac{1}{2\sigma\sqrt{\pi}}$$

equal to

$$\int\Big[\tfrac{1}{2\alpha}\,\text{sech}^2\big(\tfrac{x}{\alpha}\big)\Big]^2dx = \tfrac{1}{3\alpha}$$

and thus (on admittedly the most tenuous of grounds) directed my attention to a normal distribution

$$\frac{1}{3\alpha/\sqrt{2}}\exp\Big\{-\tfrac{1}{2}\Big[\frac{x}{3\alpha/2\sqrt{\pi}}\Big]^2\Big\} \tag{65}$$

does, on the evidence of Figure 4, appear to provide a better fit.

Looking to the associated Wigner function, we (after some elementary manipulation[29]) have

$$P_f(x,p) = \tfrac{2}{\hbar}\tfrac{1}{2\alpha}\,2\int_0^\infty \underbrace{\text{sech}\big(\tfrac{x+\xi}{\alpha}\big)\,\text{sech}\big(\tfrac{x-\xi}{\alpha}\big)}\cos\big(\tfrac{2}{\hbar}p\xi\big)\,d\xi$$
$$= \tfrac{1}{2}\big[\cosh\big(\tfrac{2}{\alpha}x\big)+\cosh\big(\tfrac{2}{\alpha}\xi\big)\big]^{-1}$$

The integral appears to lie beyond *Mathematica*'s capability, but can be discovered in Gradshteyn & Ryzhik,[30] whose reported evaluation (after a correction) entails

$$= \frac{\frac{1}{\hbar}\sin\big(\frac{2xp}{\hbar}\big)}{\sinh\big(\frac{2x}{\alpha}\big)\sinh\big(\frac{\pi\alpha p}{\hbar}\big)} \tag{66}$$

---

[29] The intent of the manipulation is to make manifest the reality of the Wigner function. For the identity, see Abramowitz & Stegun, **4.5.43**.

[30] *Table of Integrals, Series, and Products* (1965), where at **3.983.1** one reads

$$\int_0^\infty \frac{\cos ax\,dx}{b\,\text{ch}\,\beta x + c} = (\text{missing factor of }4?)\frac{\pi\sin\big(\frac{a}{\beta}\,\text{arch}\frac{c}{b}\big)}{\beta\sqrt{c^2-b^2}\,\text{sh}\frac{a\pi}{\beta}} \quad : \quad c>b>0$$

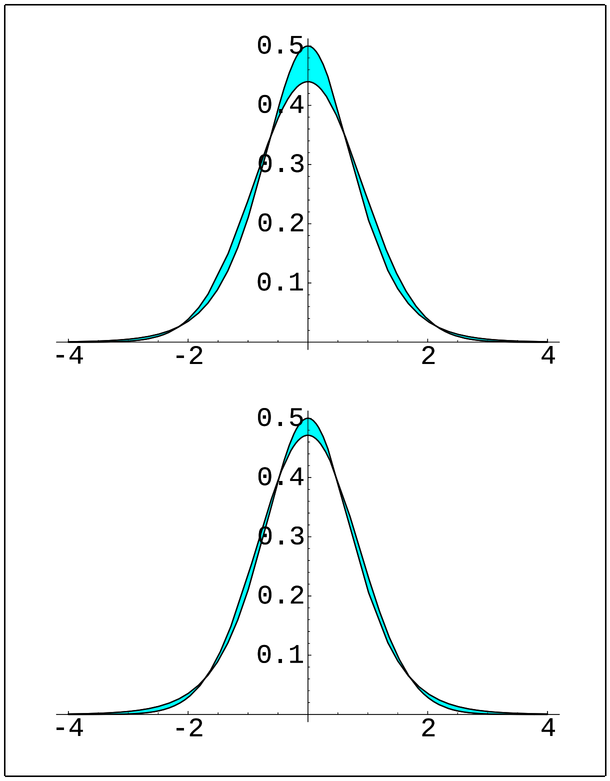

FIGURE 4: *Comparisons of (64)—above—and (65) with* $|f(x)|^2$*, in the case* $\alpha = 1$*. The latter appears to the eye to provide the better fit.*

The accuracy of (66) is supported by the observations that (according to *Mathematica*) $\int\!\!\int P_f(x,p)\,dxdp = 1$ and, as (29.3) requires,

$$|P_f(x,p)| \le P_f(0,0) = \tfrac{2}{h}$$

More convincingly, insertion of (66) into Beck's formula (54) is readily found to give back precisely the $f(x)$ of (62). When written out as functions, $P_f(x,p)$ and its "best fitting" Gaussian counterpart (I work from (52), with $t = v = 0$ and $\lambda = \hbar/2\sigma$)

$$P_g(x,p) = \tfrac{2}{h}\exp\Big\{-\tfrac{1}{2}\big[\tfrac{x}{\sigma}\big]^2 - \tfrac{1}{2}\big[\tfrac{2\sigma p}{\hbar}\big]^2\Big\}\Big|_{\sigma=3\alpha/2\sqrt{\pi}} \tag{67}$$

resemble one another not at all, but when plotted (see Figure 5) their difference is almost imperceptible. They do, however, differ in one important respect: $P_g(x,p)$ is everywhere positive, but $P_f(x,p)$ *vanishes/changes sign* on the hyperbolic contours defined by

$$xp = n\tfrac{h}{4} \quad : \quad n = \pm1, \pm2, \pm3, \ldots$$

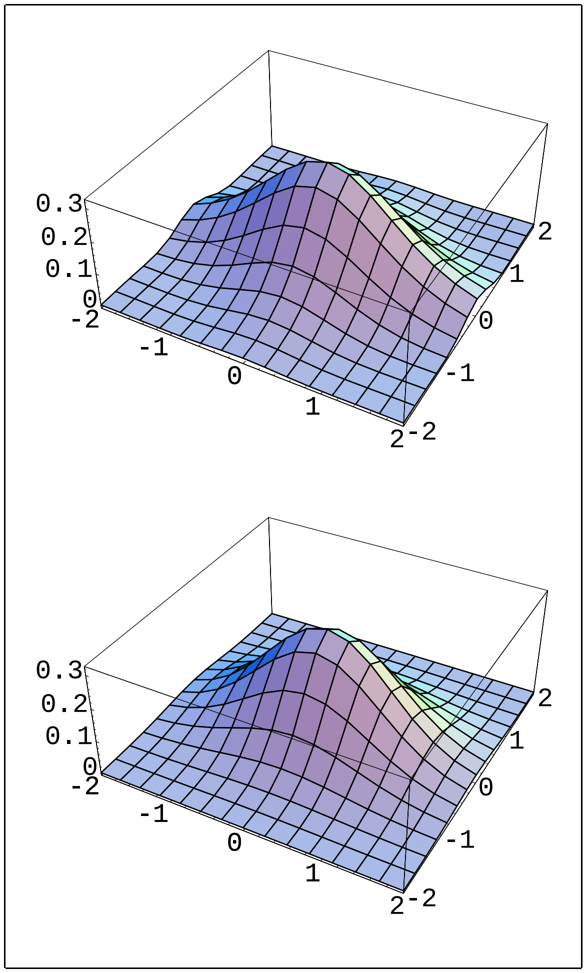

FIGURE 5: *The function $P_f(x,p)$ is plotted above, and its "best Gaussian approximant" $P_g(x,p)$ below. In preparing the figure I have set $\alpha = \hbar = 1$.*

More specifically, we have

$$\left.\begin{array}{lcl} P_f(x,p) > 0 & : & -1 < \frac{4xp}{h} < +1 \\ P_f(x,p) < 0 & : & -2 < \frac{4xp}{h} < -1 \quad \text{or} \quad +1 < \frac{4xp}{h} < +2 \\ P_f(x,p) > 0 & : & -3 < \frac{4xp}{h} < -2 \quad \text{or} \quad +2 < \frac{4xp}{h} < +3 \\ & \vdots & \end{array}\right\} \tag{68}$$

The Wigner function $P_f(x,p)$ is a "quasi-distribution" with emphasis on the quasi, and is classically disallowed, but in the latter connection we notice that the central positive region expands to enclose all of phase space in the classical limit $\hbar \downarrow 0$.

We know on the basis of the general Galilean transform argument advanced in §5 that to "launch" the wavepacket $f(x)$—because to launch *any* wavepacket —we have only to make the adjustment

$$f(x) \xrightarrow[\text{launch}]{} f(x) \cdot e^{\frac{i}{\hbar}mvx} \tag{69.1}$$

which in the phase space formulation of quantum mechanics (see again (53), the equation which, at the beginning of §6, served to *define* $P_f(x,p)$) becomes

$$P_f(x,p) \xrightarrow[\text{launch}]{} P_f(x,p-mv) \tag{69.2}$$

The preceding remark is consistent with the pattern of events foreshadowed already in classical mechanics,[31] where the Hamiltonian theory of Galilean transformations is found to be simpler than the corresponding Lagrangian theory; the later requires that a gauge transformation be built into the definition of a Galilean transformation. That adjustment entails adjustment

$$S \longrightarrow S + \big\{-mvx + \tfrac{1}{2}mv^2t\big\}$$

of the classical action, and it is to this development that both (69.1) and (69.2) can ultimately be traced.

To obtain a description of the (free) dynamical evolution of $\psi(x,0) \equiv f(x)$ we might attempt to proceed from (19), writing

$$\psi(x,0) \xrightarrow[\text{free dynamical}]{} \psi(x,t) = \int \sqrt{\tfrac{m}{iht}}\, e^{\frac{i}{\hbar}\frac{m}{2}\frac{(x-y)^2}{t}} \frac{1}{\sqrt{2\alpha}} \operatorname{sech}\big(\tfrac{y}{\alpha}\big)dy$$

but the integral appears to be intractable. Alternatively, we might pass to the phase space formalism, write

$$\begin{aligned} P_f(x,p;0) \xrightarrow[\text{free dynamical}]{} P_f(x,p;t) &= P_f(x-\tfrac{1}{m}pt,p;0) \\ &= \frac{\frac{1}{\hbar}\sin\big(\frac{2xp}{\hbar}\big)}{\sinh\big(\frac{2x}{\alpha}\big)\sinh\big(\frac{\pi\alpha p}{\hbar}\big)}\Bigg|_{x\to x-\frac{1}{m}pt} \end{aligned} \tag{70}$$

and attempt to recover $\psi(x,t)$ by appeal to Beck's formula (54). But again, the integrals encountered in the final step appear (except at $t=0$) to be intractable; we have encountered a problem in which *the phase space formalism is easier to carry to completion than the standard formalism.*

11. A conjecture—stated and refuted. For a while—intuitively, and on the slender evidence afforded by the harmonic oscillator—it seemed to me plausible to conjecture that

$$P_{\text{ground state}}(x,p) \text{ is, for all systems, everywhere non-negative}$$

[31] See CLASSICAL MECHANICS (1983), p. 252–254.

But we are in position now to show by counterexample that such a conjecture is untenable.

While integrals involving $\operatorname{sech} x$, $\operatorname{sech}^2 x$, etc. tend, as we have seen, to be less tractable than those involving $\exp\{-x^2\}$, the derivative properties of the former functions are in some respects *more* attractive than those of the Gaussian; one has

$$\begin{aligned}\left(\tfrac{d}{dx}\right)^2\sinh\ \ x &= \sinh x - 2\sinh^3 x\\ \left(\tfrac{d}{dx}\right)^2\sinh^2 x &= 4\sinh^2 x - 6\sinh^4 x\\ &\ \,\vdots\end{aligned}$$

which on their right sides exhibit what we might call the property of "non-linear functional closure" (and an absence of radicals). The latter identity lies at the heart of the construction of a famous solitonic solution of the Korteweg-deVries equation $u_t+uu_x+ku_{xxx}=0$,[32] but it is the former which is of present interest; it entails

$$\left(\tfrac{d}{dx}\right)^2\sinh\left(\tfrac{x}{\alpha}\right) = \tfrac{1}{\alpha^2}\Big\{1-2\sinh^2\left(\tfrac{x}{\alpha}\right)\Big\}\sinh\left(\tfrac{x}{\alpha}\right)$$

or again

$$\Big\{-\tfrac{\hbar^2}{2m}\left(\tfrac{d}{dx}\right)^2 - \tfrac{\hbar^2}{m\alpha^2}\sinh^2\left(\tfrac{x}{\alpha}\right)\Big\}\sinh\left(\tfrac{x}{\alpha}\right) = -\tfrac{\hbar^2}{2m\alpha^2}\sinh\left(\tfrac{x}{\alpha}\right)$$

We conclude that (62) describes a localized eigenstate $f(x)$ of the system

$$\begin{aligned}H(x,p) = \tfrac{1}{2m}p^2 + &U(x)\\ &U(x) \equiv -U_{\min}\sinh^2\left(\tfrac{x}{\alpha}\right)\end{aligned}$$

in which the "attractive sech^2-potential" has been "tuned" by setting

$$U_{\min}\alpha^2 = \hbar^2/m$$

The associated eigenvalue is

$$E = -\tfrac{1}{2}U_{\min}$$

Because $f(x)$ is node-free, we conclude that $f(x)$ is in fact the *ground state* of the system in question. But the associated Wigner function $P_f(x,p)$ has already been shown *not* to be everywhere non-negative. My conjecture is refuted.

One remark before I take leave of this topic: Richard Crandall[33] has drawn attention to a wonderful property shared by the quantum potentials

$$U_N(x) \equiv -\tfrac{\hbar^2}{2m\alpha^2}N(N+1)\sinh^2\left(\tfrac{x}{\alpha}\right) \quad : \quad N = 1,2,3,\ldots$$

---

[32] See CLASSICAL FIELD THEORY (1979), p. 55.

[33] R. E. Crandall & B. R. Litt, "Reassembly and time advance in reflectionless scattering," Ann. of Phys. **146**, 458 (1983); R. E. Crandall, "Exact propagator for reflectionless potentials," J. Phys. A **16**, 3005 (1983).

of which we have encountered the leading instance ($N = 1$). Citing Morse & Feshbach,[34] Crandall remarks that $U_N(x)$ admits of exactly N bound states, and that the energies associated with those states can be described

$$E_{N,n} = -\tfrac{\hbar^2}{2m\alpha^2}(N-n)^2 \quad : \quad n = 0, 1, \dots, N-1$$

For the system considered in connection with my ill-fated conjecture there is, on these grounds, a single bound state, which is perforce the ground state: no allusion to a "nodal folk theorem" is in fact required. It was, by the way, Crandall's interest in this problem which stimulated me to look into some of the associated classical mechanics.[35]

12. Fat Gaussians are mixtures. For the purposes of this discussion I adopt a notational refinement: we agree to

write σ and λ when the minimality condition $\sigma\lambda = \frac{1}{2}\hbar$ is in force
write $\boldsymbol{\sigma}$ and $\boldsymbol{\lambda}$ otherwise

We note that it is always possible to *achieve minimality by rescaling*; i.e., to write

$$\boldsymbol{\sigma} = b\sigma \quad\text{and}\quad \boldsymbol{\lambda} = b\lambda \quad : \quad b > 0 \text{ is the "fatness parameter"}$$

Look now again to the familiar bivariate normal distribution function

$$\begin{aligned} P(x,p) &= \tfrac{1}{\boldsymbol{\sigma\lambda}\cdot 2\pi}\exp\Big\{-\tfrac{1}{2}\big[\tfrac{x}{\boldsymbol{\sigma}}\big]^2 - \tfrac{1}{2}\big[\tfrac{p}{\boldsymbol{\lambda}}\big]^2\Big\} \\ &= \tfrac{1}{b^2}\cdot\tfrac{2}{\hbar}\exp\Big\{-\tfrac{1}{2b^2}\big[\tfrac{x}{\sigma}\big]^2 - \tfrac{1}{2b^2}\big[\tfrac{p}{\lambda}\big]^2\Big\} \end{aligned} \tag{70}$$

which will satisfy the quantum mechanically enforced boundedness condition (29.3) if and only if $b \geq 1$ (which we henceforth assume to be the case), but was seen at (59) to be the Wigner transform of a wavefunction $\psi(x)$ only in the case $b = 1$. The question now before us —first posed at the end of §9—is this: What quantum mechanical meaning can be assigned to the "fat Gaussian" distributions which arise from (70) when $b > 1$?

Introduce dimensionless variables

$$\tilde{x} \equiv x/\sigma \quad\text{and}\quad \tilde{p} \equiv p/\lambda$$

and notice that

$$(x/\sigma)^2 + (p/\lambda)^2 = \tilde{x}^2 + \tilde{p}^2 = \text{constant}$$

inscribes an ellipse on the $\{x,p\}$-plane, but a *circle* on the $\{\tilde{x},\tilde{p}\}$-plane. Introduce polar coordinates

$$\begin{aligned} \tilde{x} &= r\cos\theta \\ \tilde{p} &= r\sin\theta \end{aligned}$$

[34] *Methods of Theoretical Physics* (1953), p. 1654.
[35] CLASSICAL MECHANICS (1983), Appendix C.

and notice that (70) can be written

$$P(x,p) = \tfrac{1}{b^2}\cdot\tfrac{2}{h}\exp\left\{-\tfrac{1}{2b^2}r^2\right\}$$

Finally, introduce

$$z \equiv r^2$$

to obtain

$$P(x,p) = \tfrac{1}{b^2}\cdot\tfrac{2}{h}e^{-\frac{1}{2b^2}z} \tag{71}$$

and notice (see again (61)) that one recovers

$$\begin{aligned}&\downarrow\\&= \text{oscillator ground state } \tfrac{2}{h}e^{-\frac{1}{2}z} \text{ at } b\downarrow 1\end{aligned}$$

Evidently $dxdp = \sigma\lambda\cdot d\tilde{x}d\tilde{p} = \frac{\hbar}{2}rdrd\theta = \frac{\hbar}{4}dzd\theta$, in consequence of which we recover

$$\iint P(x,p)\,dxdp = \tfrac{\hbar}{4}\tfrac{1}{b^2}\tfrac{2}{h}\,2\pi\int_0^\infty e^{-\frac{1}{2b^2}z}dz = 1$$

So much by way of preparation.

Look now to the Wigner functions

$$P_n(x,p) \equiv (-)^n\tfrac{2}{h}\,e^{-\frac{1}{2}z}L_n(z) \tag{72}$$

which were at (61) directed to our attention by—of all things!—the quantum theory of oscillators. Notice that

$$\begin{aligned}\iint P_m(x,p)P_n(x,p)\,dxdp &= (-)^{m+n}\left(\tfrac{2}{h}\right)^2\tfrac{\hbar}{4}\,2\pi\int_0^\infty e^{-z}L_m(z)L_n(z)\,dz\\&= \tfrac{1}{h}\delta_{mn}\end{aligned} \tag{73}$$

by the orthogonality of the Laguerre polynomials.[36] The idea now (see again the end of §9) is to write

$$\begin{aligned}P(x,p) &= \sum_n p_nP_n(x,p)\\p_n &= h\iint P(x,p)P_n(x,p)\,dxdp\end{aligned}$$

which, when $P(x,p)$ is given by (71), becomes (after simplifications)

$$\begin{aligned}\text{fat Gaussian} &= \sum_n p_nP_n(x,p)\\p_n &= (-)^n\tfrac{1}{b^2}\int_0^\infty e^{-\frac{1}{2}\left(1+\frac{1}{b^2}\right)z}L_n(z)\,dz\end{aligned} \tag{74}$$

[36] For a nice account of the properties of these polynomials see Chapter 23 of J. Spanier & K. B. Oldham, *An Atlas of Functions* (1987).

The integral is tabulated,[37] and gives

$$p_n = \tfrac{2}{b^2+1}\left[\tfrac{b^2-1}{b^2+1}\right]^n \tag{75}$$

concerning which several elementary remarks (the net effect of which is to inspire confidence in the accuracy of (75)) are in order:

$$\sum_n p_n = 1 \quad : \quad \text{all values of } k$$

$$p_n \text{ are non-negative for all } n \text{ if and only if } b \geq 1$$

$$\text{if } b = 1 \text{ then } p_0 = 1 \text{ and all other } p_n \text{ vanish}$$

The second of the preceding points merits special comment: if $b < 1$ were allowed, then not only would we stand in violation of the uncertainty principle, we would acquire an obligaton to try to make sense of *"negative statistical weights" in the sense of ordinary probability theory*—a notion for which even quantum mechanics makes no provision.

I conclude this discussion with miscellaneous comments of a more general nature:

The issue before us has been *pre*-dynamical: we have been concerned with the quantum mechanical interpretation of Wigner distributions of a certain ("fat Gaussian") design; we have not been concerned with the dynamical motion/flow of such a distribution; no Hamiltonian has been specified. Reference to "oscillators" is relevant only to this extent: the conditions

$$P(x,p) = \text{constant} \qquad \text{and} \qquad H_{\text{osc}}(x,p) = E$$

give rise to identical populations of (elliptical) curves. It would cut closer to the analytical heart of the matter to observe (say) that[38]

$$\sum_{k=0}^{n} \tbinom{n}{k} H_{2k}(x) H_{2n-2k}(y) = (-)^n n! L_n(x^2+y^2)$$

and that the Hermit polynomials $H_n(x) \equiv (-)^n e^{x^2}\left(\frac{d}{dx}\right)^n e^{-x^2}$—made familiar to us by the quantum theory of oscillators—are in all respects very "Gaussian."

[37] See Gradshteyn & Ryzhik, **7.414.5**, where we are informed that

$$\int_0^\infty e^{-az} L_n(z)\,dz = (a-1)^n a^{-n-1} \quad : \quad \Re[\,a\,] > 0$$

[38] See A. Erdélyi *et al*, *Higher Transcendental Functions II* (1953), **10.13.32**. The identity quoted is intended merely to typify the intricately patterned relationships which link the Hermite with Laguerre polynomials.

Straightforward extension of the method described above permits one to accomplish the "spectral analysis" (not just of "fat Gaussians" but) of *any* Wigner function of the specialized form

$$P(x,p) = f(z) \quad \text{with} \quad z = z(x,p) = (x/\sigma)^2 + (p/\lambda)^2$$

But it leaves us still powerless to attack the general problem posed at the end of §9—powerless to discover weights p_n and functions $P_n(x,p)$ which permit one to write

$$P(x,p) = \sum_n p_n P_n(x,p)$$

when the Wigner function $P(x,p)$ is arbitrary.

The question arises: How—physically—would one *prepare* the mixture of states to which a "fat Gaussian" has been found to refer? One thinks naturally of "simultaneous x and p measurements of less-than-optimal precision," but (in the absence of a well worked out phase space formulation of the quantum theory of measurement) I see no way to make that idea precise. I describe an alternative procedure which does in fact hinge upon the quantum *physics of oscillators*. Considering σ (whence also $\lambda = \hbar/2\sigma$) to have been given, select an oscillator in which the mass and frequency have been "tuned" in such a way as to achieve

$$H(x,p) = \tfrac{1}{2m}p^2 + \tfrac{m\omega^2}{2}x^2 = \tfrac{1}{4}\hbar\omega \underbrace{\big\{(x/\sigma)^2 + (p/\lambda)^2\big\}}_{z} \tag{76}$$

Select an oscillator from a thermalized population of such oscillators. It will be in state $|n)$ with probability

$$p_n = \tfrac{1}{Z}e^{-\beta(n+\frac{1}{2})} \qquad : \qquad \beta \equiv \tfrac{\hbar\omega}{kT}$$

and since the partition function

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

we have

$$p_n = \big(1-e^{-\beta}\big)\big[e^{-\beta}\big]^n$$

which if we set

$$e^{-\beta} = \tfrac{b^2-1}{b^2+1} \qquad i.e., \qquad b^2 = \tfrac{1+e^{-\beta}}{1-e^{-\beta}} = \coth\big(\tfrac{1}{2}\beta\big) \tag{77}$$

is found by quick algebra to possess precisely the structure of (75)! That "fat Gaussians" lurk within the thermal physics of a quantum oscillator, and become minimally dispersed ground state Gaussians as $T \downarrow 0$, is a result which I find quite satisfying—made the more so by the following observation: Classically we expect to have $\frac{1}{2}kT = E = \frac{1}{2}m\omega^2(\text{amplitude})^2$, giving

$$(\text{amplitude})^2 = \tfrac{kT}{m\omega^2} \tag{78.1}$$

On the other hand, it was remarked already on p. 28 and follows from (76) that $\sigma^2 = \hbar/2m\omega$, while (77) entails

$$\begin{aligned} b^2 &= \tfrac{2}{\beta} + \tfrac{\beta}{6} - \tfrac{\beta^3}{360} + \cdots \\ &\sim 2\tfrac{kT}{\hbar\omega} \text{ for temperatures of “classical magnitude” } kT \gg \hbar\omega \end{aligned}$$

so we have

$$\boldsymbol{\sigma}^2 = (b\sigma)^2 \sim \tfrac{kT}{m\omega^2} \tag{78.2}$$

The point to notice is that if we accept the identification

$$\boldsymbol{\sigma} \quad\longleftrightarrow\quad \text{amplitude}$$

then (78.1) and (78.2) *say precisely the same thing!*[39]

The idea now is to adjust the temperature T so as to assign the desired value to the “fatness constant” b and then, at time $t = 0$, to “snip the spring.” If $T = 0$ then $b = 1$ and we end up with a Gaussian of minimal dispersion (pure ground state), but if $T > 0$ (which entails $b > 1$) we produce a thermal *mixture* of oscillator states, represented within the phase space formalism by a Wigner function of “fat Gaussian” design. . . which then moves off under the dynamical control of whatever Hamiltonian we have build into our apparatus (i.e., which we abruptly turn on when we snipped the spring; the default Hamiltonian is would be that of a free particle).

13. “Entropy” of a Gaussian wavepacket. It was Boltzmann—working during the decade prior to 1877 toward a kinetic theoretic interpretation of the entropy concept—who first drew the attention of physicists to the importance of the construction

$$H = -\sum p_n \log p_n = -\langle \log p_n \rangle$$

in connection with which it is indispensable to notice that (obviously, else by l’Hôpital’s Rule)

$$\lim_{p\uparrow 1} p \log p = \lim_{p\downarrow 0} p \log p = 0$$

And it was Claude Shannon (1948) who first appreciated that Boltzmann’s construction possesses an abstract significance far deeper, and a utility far more diverse, than that originally contemplated.

---

[39] For related discussion see §30 in L. D. Landau & E. M. Lifshitz, *Statistical Physics* (1958).

Quantum mechanics—which is nothing if not an intensely probabilistic subject—presents an abundance of distribution functions of various types. One might therefore expect importance to attach to expressions of (say) the form

$$H[\psi] \equiv -\int |\psi(x,t)|^2 \log |\psi(x,t)|^2 \, dx \tag{79}$$

and is surprised to discover that the notion thus expressed appears to be—or at least until quite recently have been[40]—absent from the quantum mechanical literature. Within the phase space formalism one is similarly inspired to write

$$H[P] \equiv -\iint P(x,p;t) \log P(x,p;t) \, dx dp \tag{80}$$

which, though absent from the quantum mechanical literature, is formally identical to an equation standard to *classical statistical mechanics*:[41] at thermal equilibrium one expects to have $P(x,p) = \frac{1}{Z} e^{-\frac{1}{kT}E(x,p)}$ which gives

$$H[P] = \tfrac{1}{kT}\langle E \rangle + \log Z$$

and by a celebrated line of argument[42] discovers that the expression on the right admits of the interpretation

$$= \frac{\text{entropy}}{k}$$

It is in allusion to this fact that in modern literature[43] one encounters, even in contexts which have nothing to do with thermodynamics, the terminology

$$\begin{aligned} -\log P(\text{event}) &\equiv \text{“event-information”} \\ H[P] = -\sum_{\text{events}} P \log P = -\langle \log P \rangle &= \text{mean event-information} \\ &\equiv \text{“entropy”} \end{aligned}$$

[40] I will have occasion later to cite papers by C. Adami and others, which began to appear during the mid-1990's.

[41] *Quantum* statistical mechanics is nearly as old as quantum mechanics itself (arguably older, since it was a statistical mechanical problem that inspired Planck), and already by 1927 it had been pointed out by J. von Neumann (in his *Mathematische Begründung der Quantaenmechanik*; see p. 394 in the English translation) that in *thermalized* situations it makes sense to write

$$\boldsymbol{\rho} = \tfrac{1}{Z} e^{-\frac{1}{kT}\mathsf{H}} \quad \text{with} \quad Z = \operatorname{tr} e^{-\frac{1}{kT}\mathsf{H}}$$

and to set

$$\text{thermodynamic entropy} = -k \operatorname{tr} \boldsymbol{\rho} \log \boldsymbol{\rho}$$

and it was to facilitate interpretion of these statements that Wigner was led to the *invention* of the Wigner transform. So when I say "absent from the quantum literature" I should emphasize that I mean *absent in contexts which make no reference to thermal equalibrium*.

[42] See STATISTICAL PHYSICS (1969), Chapter 3, pp. 35–38.

[43] See, for example, "**213.B**, Information Theory" in *Encyclopedic Dictionary of Mathematics* (1993).

$H[P]$ is by nature a real number-valued functional of the distribution $P(\text{event})$.

The relative unimportance of (79) can, I suppose, be accounted for as follows: write $\psi = Ae^{\frac{i}{\hbar}S}$. Then (79) reads $H[\psi] = -\int |A|^2 \log |A|^2\,dx$, from which all reference to the phase factor S has dropped away. But it is within that phase factor—closely related to the classical action—that much of the most characteristically quantum mechanical "good stuff" is known to reside. The dismissive force of that remark evaporates when one turns from (79) to (80), for folded into the functional structure of $P_\psi(x,p;t)$, is *all* the physically significant information conveyed by *both* $A(x,t)$ and $S(x,t)$; this follows from the invertibility of the Wigner transform (Beck's formula).

On another occasion I will look (within the phase space formalism) to some of the more general quantum applications of the entropy concept; here my intention is simply to plant seeds for future harvest, and to examine entropic aspects of the Gaussian wavepackets which are at present my major concern

It serves my present needs to adopt a notational refinement of (1):

$$\begin{aligned} G(x;\sigma) &\equiv e^{-\left\{\frac{1}{2}[\frac{x}{\sigma}]^2+\log(\sigma\sqrt{2\pi})\right\}}\\ &= \text{centered Gaussian distribution with variance } \sigma^2 \end{aligned}$$

Evidently

$$\begin{aligned} H[G] &= \tfrac{1}{2}\big\{1+\log(2\pi)\big\}+\log\sigma\\ &= 1.41894+\log\sigma \end{aligned} \tag{81}$$

For purposes of comparison, consider the flat "particle-in-a-box distribution"

$$b(x;a) \equiv \begin{cases} \frac{1}{2a} & : \quad x^2 \le a^2\\ 0 & : \quad \text{outside that interval} \end{cases}$$

Obviously $\langle x\rangle = 0$, and by quick calculation $\langle x^2\rangle = \frac{1}{3}a^2$, so

$$B(x;\sigma) \equiv b(x;\sigma\sqrt{3}) \equiv \text{centered box distribution with variance } \sigma^2$$

We compute

$$\begin{aligned} H[B] &= \Big[-\int_{-a}^{+a}\tfrac{1}{2a}\log\tfrac{1}{2a}\,dx = \log(2a)\Big]_{a=\sigma\sqrt{3}}\\ &= \log(2\sqrt{3})+\log\sigma\\ &= 1.24245+\log\sigma < H[G] \quad : \quad \text{all values of } \sigma \end{aligned}$$

Or consider this distribution, borrowed from discussion subsequent to (62) in §10:

$$\begin{aligned} F(x;\sigma) &\equiv \Big[\tfrac{1}{2\alpha}\,\mathrm{sech}^2\big(\tfrac{x}{\alpha}\big)\Big]_{\alpha=\sqrt{12/\pi^2}\,\sigma}\\ &\equiv \text{centered sech}^2 \text{ distribution with variance } \sigma^2 \end{aligned}$$

We compute

$$
\begin{aligned}
H[F\,] &= \Big[-\!\int \tfrac{1}{\alpha}\,\mathrm{sech}^2\big(\tfrac{x}{\alpha}\big)\log\mathrm{sech}\big(\tfrac{x}{\alpha}\big)\,dx + \log(2\alpha)\Big]_{\alpha=\sqrt{12/\pi^2}\sigma}\\
&= 2 - 2\log 2 + \log(2\sqrt{12/\pi^2}) + \log\sigma\\
&= 1.40458 + \log\sigma < H[G] \quad : \quad \text{all values of } \sigma
\end{aligned}
$$

The preceding inequalities are illustrative of a *general* inequality, which emerges from the solution of this problem:

$$
\text{maximize } H[p(x)] \text{ subject to the constraints } \begin{cases} \langle x^0\rangle = 1\\ \langle x^1\rangle = 0\\ \langle x^2\rangle = \sigma^2 \end{cases}
$$

Using the method of Lagrange multipliers to manage the constraints, we are led[44] to write $\delta\int\big\{p\log p + a + bx + cx^2\big\}\,dx = 0$ whence

$$
\int \delta p\big\{\log p + (1+a) + bx + cx^2\big\}\,dp = 0
$$

giving

$$
p(x) = e^{-\big\{(1+a)+bx+cx^2\big\}}
$$

and we find ourselves obligated by the constaints to set $c = \frac{1}{2\sigma^2}$, $b = 0$ and $e^{-(1+a)} = \frac{1}{\sigma\sqrt{2\pi}}$; i.e., to reproduce (1). The proposition[45] that

> *Among the distributions $p(x)$ of specified variance,*
> *entropy $H[p(x)]$ is maximized at the Gaussian*

lends the Gaussian an importance which (it seems to me) cuts much deeper than mere “analytical tractability;” it is on these grounds not a “distribution among equals” but by birthright entitled to a little occasional “tyranny.”

Returning with (81) to (13), we see that the standing motion of a Gaussian wavepacket lends the associated entropy a time-dependence which can be described

$$
H[|\psi(x,t)|^2] = \tfrac{1}{2}\big\{1+\log(2\pi)\big\} + \log\sigma(t) \qquad \sigma(t) = \sigma\sqrt{1+(t/\tau)^2} \tag{82.1}
$$

[44] My analytical method is a variant of one devised by Boltzmann to extract the Maxwell distribution from the H-theorem; see D. ter Haar, *Elements of Statistical Mechanics* (1954), p. 20, or Chapter 2, p. 34 in STATISTICAL PHYSICS (1969). Also Appendix 14 in Max Born’s *Natural Philosophy of Cause and Chance* (1949).

[45] Albyn Jones assures me that this result is “reasonably well known,” but could suggest no published source, nor have I attempted to locate one.

and (asymptotically) grows as $\log t$ as the packet becomes spatially dispersed. In the momentum representation (see again (9) and (5)) one on the other hand obtains

$$H[|\varphi(p,t)|^2] = \tfrac{1}{2}\big\{1+\log(2\pi)\big\} + \log\lambda \tag{82.2}$$

which is constant in time.

The phase space formalism leads to conclusions which are relatively more interesting, and made so because they embrace all of the "missing physics;" it looks, after all, not to marginal distributions but to the underlying joint (quasi-) distribution. In the case of a "freely moving launched Gaussian wavepacket" we at (30.2) had

$$P_\psi(x,p;t) = \tfrac{1}{\sigma\lambda 2\pi}\exp\Big\{-\tfrac{1}{2}\big[\tfrac{X}{\sigma}-\theta\tfrac{P}{\lambda}\big]^2 - \tfrac{1}{2}\big[\tfrac{P}{\lambda}\big]^2\Big\}$$

(here as before: $\lambda \equiv \hbar/2\sigma$, $\theta \equiv t/\tau$, $X \equiv x - vt$ and $P \equiv p = mv$) and if we allow ourselves to work from (80) obtain

$$\begin{aligned} H[P_\psi] &= \iint P_\psi(x,p;t)\Big\{\tfrac{1}{2}\big[\tfrac{X}{\sigma}-\theta\tfrac{P}{\lambda}\big]^2 + \tfrac{1}{2}\big[\tfrac{P}{\lambda}\big]^2 + \log(\sigma\lambda 2\pi)\Big\}\,dXdP \\ &= \tfrac{1}{2\sigma^2}\langle X^2\rangle - \tfrac{\theta}{\sigma\lambda}\langle XP\rangle + \tfrac{1+\theta^2}{2\lambda^2}\langle P^2\rangle + \log(\sigma\lambda 2\pi) \\ &= \tfrac{1+\theta^2}{2} - \theta^2 + \tfrac{1+\theta^2}{2} + \log(\sigma\lambda 2\pi) \\ &= 1 + \log(\sigma\lambda 2\pi) \end{aligned} \tag{83}$$

The structure of (83) inspires the following series of comments, which will in turn motivate some adjustments, and open some doors:

I draw attention first to the fact that at (80)—as previously also at (79)—we committed a major breach of etiquette: we allowed ourselves to form the logarithm of an expression which bears physical dimension. The logarithm is uniquely forgiving of such indiscretions, but our *faux pas* should not be allowed to stand uncorrected. Whether one argues from the defintion (53) of the Wigner transform or from $\iint P(x,p)\,dxdp = 1$, it is clear that

$$\dim P_\psi(x,p) = \tfrac{1}{\text{action}}$$

But it is an implication of the familiar boundedness condition (29.3) (i.e., of the Heisenberg uncertainty principle) that

$$\text{footprint of most compact Wigner function} = \tfrac{h}{2}$$

and it becomes therefore natural to write

$$1 = \iint P(x,p)\,dxdp = \iint \tilde P(x,p)\,\tfrac{dxdp}{h/2}$$
$$\tilde P(x,p) \equiv \tfrac{h}{2}P(x,p)$$

This obvious variant of (80) $H[\tilde{P}\,] \equiv -\iint \tilde{P}\log\tilde{P}\,d\Omega$ (I adopt here the fairly standard abbreviation $d\Omega \equiv \frac{2}{h}dxdp$) is free of the defect to which I have drawn attention,[46] and entails

$$H[\tilde{P}\,] = H[P\,] - \log(h/2) \tag{84}$$

Thus are we led, in place of (83), to write

$$\begin{aligned} H[\tilde{P}_{\text{gaussian}}\,] &= 1 + \log(\sigma\lambda 2\pi) - \log(h/2) \\ &= 1 \quad \text{because } \ \sigma\lambda 2\pi = h/2 \end{aligned} \tag{85}$$

To show that *we stand now in the presence of a contraction* I have now to digress a bit: Let $|\psi_n)$ refer to some/any set of orthonormal states, and let $\mathsf{p}_n \equiv |\psi_n)(\psi_n|$ refer to the associated set of associated projection operators. One has

$$\mathsf{p}_n^2 = \mathsf{p}_n \quad ; \quad \mathsf{p}_m\mathsf{p}_n = \mathbf{0} \ \ (m \neq n) \quad ; \quad \sum \mathsf{p}_n = \mathbf{1} \tag{86.1}$$

where the former statements imply trace-wise orthonormality

$$\mathrm{tr}\,\mathsf{p}_m\mathsf{p}_n = \delta_{mn} \tag{86.2}$$

and where the latter says simply that the $\big\{|\psi_n)\big\}$ are complete in whatever space they span, whether infinite or finite dimensional. Let $\sum a_n\mathsf{p}_n$ be some/any linear combination of projection operators, and let $f(\bullet)$ be some/any formal power series in a single variable; it follows formally from (86.1) that[47]

$$f\big(\textstyle\sum a_n\mathsf{p}_n\big) = \sum f(a_n)\,\mathsf{p}_n \tag{87}$$

Let the density matrix $\boldsymbol{\rho} = \sum p_n\mathsf{p}_n$ refer to some/any statistically weighted mixture of such states. It is (formally) an implication of (87) that

$$\boldsymbol{\rho} = e^{\sum(\log p_n)\mathsf{p}_n} \tag{88}$$

which lends meaning *in orthonormal cases* to the expression "$\log\boldsymbol{\rho}$." We find ourselves therefore in position to write

$$\begin{aligned} H[\boldsymbol{\rho}\,] &= -\mathrm{tr}\,(\boldsymbol{\rho}\log\boldsymbol{\rho}) && (89.1) \\ &= -\mathrm{tr}\Big\{\sum_{m,n}(p_m\log p_n)\,\mathsf{p}_m\mathsf{p}_n\Big\} \\ &= -\textstyle\sum p_n\log p_n && (89.2) \end{aligned}$$

[46] It is interesting to note that (79), in the absence of a generally available "natural length," admits of no such adjustment.

[47] The following statement is more familiar as encountered within the theory of diagonal matrices.

We conclude that the entropy of a mixed state has nothing to do with "quantum statistics," but everything to do with the *mixture statistics*. The entropy of a mixture is maximal when[48]

$$p_n = \tfrac{1}{N} \quad \text{with} \quad N \equiv \text{tr}\,\mathbf{1} = \text{dimension of spanned space} \quad : \quad \text{all } n$$

in which case one has

$$H[\boldsymbol{\rho}]_{\max} = \log N$$

The entropy becomes minimal in the "pure case" $\boldsymbol{\rho}^2 = \boldsymbol{\rho}$; i.e.; when one of the p_n is unity and the rest vanish, in which case one has

$$H[\boldsymbol{\rho}]_{\min} = 0 \tag{90}$$

which is contradicted by (85). The force of the contradiction is, however, blunted by the circumstance that (88) becomes strictly *meaningless* in the pure case, and the accuracy of (89.2) is to that extent compromised. Let us look therefore to a case in which none of the p_n vanish—the case of a thermalized oscillator.

We observed already near the end of §12 that in for a thermalized quantum oscillator one has

$$Z = \left[\,2\sinh\tfrac{1}{2}\beta\,\right]^{-1} \quad \text{and} \quad p_n = (1-e^{-\beta})e^{-n\beta}$$

with $\beta \equiv \hbar\omega/kT$. Working from (89.2) we are led therefore to write

$$\begin{aligned} H_{\text{osc}} &= \sum(1-e^{-\beta})e^{-n\beta}\big\{-\log(1-e^{-\beta})+n\beta\big\}\\ &= -\log(1-e^{-\beta}) + (1-e^{-\beta})\big(-\beta\tfrac{d}{d\beta}\big)\sum e^{-n\beta}\\ &= -\log(1-e^{-\beta}) + \beta/(e^{\beta}-1) \end{aligned} \tag{91}$$

which is readily brought to a form

$$= -\log 2\sinh\tfrac{1}{2}\beta + \tfrac{1}{2}\beta\coth\tfrac{1}{2}\beta$$

which a thermodynamicist would have extracted more swiftly from

$$= \big\{1-\beta\tfrac{d}{d\beta}\big\}\log Z$$

We observe that (91) does conform to (90) in the limit $\beta\uparrow\infty$ (i.e., as $T\downarrow 0$). We know,on the other hand, that a thermalized mixture of quantum states gives rise within the phase space formalism to the "fat Gaussian"

$$P(x,p;t) = \tfrac{1}{\boldsymbol{\sigma}\boldsymbol{\lambda}2\pi}\exp\left\{-\tfrac{1}{2}\big[\tfrac{X}{\boldsymbol{\sigma}}-\theta\tfrac{P}{\boldsymbol{\lambda}}\big]^2-\tfrac{1}{2}\big[\tfrac{P}{\boldsymbol{\lambda}}\big]^2\right\}$$

---

[48] For the detailed argument see A. I. Khinchin, *Mathematical Foundations of Information Theory* (1957), p. 4.

and by trivial adjustment (send $\sigma \mapsto \boldsymbol{\sigma} = b\sigma$ and $\lambda \mapsto \boldsymbol{\lambda} = b\lambda$) of the calculation which gave (83) we have

$$\begin{aligned} H[\tilde{P}_{\text{fat gaussian mixture}}] &= 1 + \log(\boldsymbol{\sigma\lambda}\,2\pi) - \log(h/2) \\ &= 1 + \log b^2 \\ &= 1 + \log\coth\tfrac{1}{2}\beta \\ &= 1 + \log\big\{(1+e^{-\beta})/(1-e^{-\beta})\big\} \end{aligned} \tag{92}$$

It is gratifying to observe in reference to (85) that

$$H[\tilde{P}_{\text{fat gaussian mixture}}] \geq H[\tilde{P}_{\text{pure gaussian}}], \text{ with equality in the limit } \beta\uparrow\infty$$

but the limit does *not* conform to (90), and—which is more to the point—the right sides of (91) and (92) are distinct. How did that come to be so?

The answer has to do with some fundamentals of the Weyl correspondence: from

$$P(x,p) \xleftarrow[\text{Weyl}]{} \boldsymbol{\rho}$$

it does *not* follow that

$$P(x,p)\log P(x,p) \xleftarrow[\text{Weyl}]{} \boldsymbol{\rho}\log\boldsymbol{\rho}$$

because more generally

$$A(x,p) \xleftarrow[\text{Weyl}]{} \mathbf{A} \quad \text{does not imply} \quad f\big(A(x,p)\big) \xleftarrow[\text{Weyl}]{} f(\mathbf{A}) \tag{93}$$

That (invariably)

$$\text{Weyl transform of sum} = \text{sum of Weyl transforms} \tag{94.1}$$

but (except in some special cases)

$$\text{Weyl transform of product} \neq \text{product of Weyl transforms} \tag{94.2}$$

is seen from[49]

$$A(x,p) = \iint \Big\{ \tfrac{1}{h}\,\text{tr}\,\mathbf{A}\, e^{-\frac{i}{\hbar}(\alpha\mathbf{p}+\beta\mathbf{x})}\Big\} e^{\frac{i}{\hbar}(\alpha p+\beta x)}\, d\alpha d\beta \tag{95}$$

to follow from the corresponding elementary properties of the trace. The circumstance highlighted at (93) has sometimes been held (by von Neumann and others: see J. R. Shewell, "On the formation of quantum-mechanical operators,"

[49] The following description of the *inverse Weyl transform*—which can, by the way, be made the basis of an alternative derivation of Beck's formula (54)—can be extracted from material presented on pp. 101–103 of QUANTUM MECHANICS (1967), Chapter 3.

AJP **27**, 16 (1959)) to comprise a defect of the Weyl correspondence procedure, though I fail to see why; starting from the classical Hamiltonian, one has

$$H(x,p) \xrightarrow[\text{Weyl}]{\qquad} \mathsf{H} \quad \text{does not imply} \quad e^{-\beta H(x,p)} \xleftarrow[\text{Weyl}]{\qquad} e^{-\beta \mathsf{H}}$$

and recognizes that if it were otherwise then quantum mechanics and classical statistical mechanics would be the same theory! Standardly, one gives meaning to $e^{-\beta\mathsf{H}}$ either by appeal to the spectral resolution of H

$$H = \sum E_n|n)(n| \quad\Longrightarrow\quad e^{-\beta\mathsf{H}} = \sum e^{-\beta E_n}|n)(n|$$

or (in favorable cases) by appeal to the methods of operator algebra, or by appeal to Feynman's sum-over-paths technique. The former procedure, when brought to bear upon $\boldsymbol{\rho}\log\boldsymbol{\rho}$ places one in position to exploit the additive property (94.1) of the Weyl transform; writing

$$\begin{aligned}\boldsymbol{\rho} = \sum p_n|n)(n| \quad\longrightarrow\quad \boldsymbol{\rho}\log\boldsymbol{\rho} &= \sum p_n \log p_n |n)(n| \\ &\Big\downarrow \text{Weyl/Wigner} \\ &= \sum p_n \log p_n P_n(x,p)\end{aligned}$$

we would be led back again to precisely (89.2), and to the conclusion that "the entropy of a mixed state has. . . to do with the mixture statistics."

We conclude that (89) and (80) *speak about different things*: the former alludes (in quantum dress) to the familiar stuff of ordinary statistical mechanics, while the latter presents us with an opportunity to say something new. To emphasize the peculiarly quantum mechanical character of that "something new" I will henceforth write

$$H_q[P] \equiv -\iint P(x,p)\log\tilde{P}(x,p)\,dxdp \tag{96}$$

where the subscript is intended to suggest "quantum." The expression on the right side of (96) is recommended to our attention with the same force, and for the same function-theoretic reasons, that serve invariably—in *all* contexts—to recommend $\langle \log P\rangle$. But I chose my phraseology with care when I said that $H_q[P]$ is "*peculiarly* quantum mechanical;" $P(x,p)$ refers in (96) to a Wigner function (or "quasi-distribution"), and Wigner functions can assume negative values, with the result that

$H_q[P]$ will, in general, be *complex*

From the elementary observation that

$$\tilde{P}(x,p) = \begin{cases} |\tilde{P}(x,p)| & \text{at points where } P(x,p) \geq 0 \\ e^{i\pi}|\tilde{P}(x,p)| & \text{at points where } P(x,p) < 0\end{cases}$$

we conclude that

$$H_q[P] = -\iint P(x,p)\log|\tilde{P}(x,p)|\,dxdp - i\pi\iint_{\text{neg}} P(x,p)\,dxdp \tag{97}$$

where the first $\iint$ ranges over all of phase space, while the second ranges only over the domain on which $P(x,p) < 0$.

To gain some feeling for what (97) has to say in concrete cases, we look again to the Wigner functions (72) of a harmonic oscillator. Borrowing some computational technique from §12, we for the ground state have

$$\begin{aligned} H_q[P_0] &= -\tfrac{1}{2}\int_0^\infty \Big\{e^{-\frac{1}{2}z}\Big\}\Big\{-\tfrac{1}{2}z\Big\}dz - i\pi\,\tfrac{1}{2}\int_{\text{empty}}\Big\{e^{-\frac{1}{2}z}\Big\}dz \\ &= 1 + i\,0 \end{aligned}$$

which simply reproduces (85). For the first excited state (for which the Wigner function becomes negative on the interval $0 \le z < 1$) we have

$$\begin{aligned} H_q[P_1] &= -\tfrac{1}{2}\int_0^1 \Big\{-e^{-\frac{1}{2}z}\big[1-z\big]\Big\}\Big\{-\tfrac{1}{2}z + \log\big[1-z\big]\Big\}dz \\ &\quad -\tfrac{1}{2}\int_1^\infty \Big\{-e^{-\frac{1}{2}z}\big[1-z\big]\Big\}\Big\{-\tfrac{1}{2}z + \log\big[z-1\big]\Big\}dz \\ &\quad - i\pi\,\tfrac{1}{2}\int_0^1 \Big\{-e^{-\frac{1}{2}z}\big[1-z\big]\Big\}dz \\ &= 2.90469 - 0.60935\,i - \tfrac{1}{2}\int_1^\infty \Big\{-e^{-\frac{1}{2}z}\big[1-z\big]\Big\}\log\big[z-1\big]dz \end{aligned}$$

where the numerics have been obtained by evaluation of expressions which *Mathematica* generated with ease, though it found the final integral to be "intractable." A simple change of variable leads, however, to an elementary Laplace transform

$$\begin{aligned} \tfrac{1}{2}\int_1^\infty \Big\{e^{-\frac{1}{2}z}\big[z-1\big]\Big\}\log\big[z-1\big]dz &= \tfrac{1}{2\sqrt{e}}\int_0^\infty e^{-sy}\,y\log y\,dz\Big|_{s=\frac{1}{2}} \\ &= \tfrac{1}{2\sqrt{e}}\Big[\tfrac{1-\gamma-\log s}{s^2}\Big]_{s=\frac{1}{2}} \\ &= 1.35369 \end{aligned}$$

so we have $H_q[P_1] = 1.55100 - 0.60935\,i$, and the emergence of no evident pattern. The evaluation of $H_q[P_2]$ promises to be even more tedious, and I see no way to adapt generating function techniques to the evaluation $H_q[P_n]$. The moral appears to be that one can expect $H_q[P]$-evaluation to be difficult. Nor is this surprising; it is merely a particular manifestation of a pervasive fact: In thermalized situations, entropy can be extracted from the partition function—thermodynamics supplies the general procedure—where all computational difficulties come into convenient focus ("to know Z is to know everything"). But in non-thermalized contexts (i.e., the absence of a partition function) one enjoys no such advantage, and can expect entropy calculations *almost always* to be difficult. It is into just such a context that $H_q[P]$ plunges us, for (in the absence of a successful hidden-variable theory) we certainly do not expect the design of $P_\psi(x,p)$ to have anything to do with any interpretation of the "thermalization" concept.