SCHRÖDINGER'S TRAIN OF THOUGHT

Nicholas Wheeler, Reed College Physics Department
April 2006

Introduction. David Griffiths, in his *Introduction to Quantum Mechanics* (2nd edition, 2005), is content—at equation (1.1) on page 1—to pull (a typical instance of) the Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} + V\Psi$$

out of his hat, and then to proceed directly to book-length discussion of its interpretation and illustrative physical ramifications. I remarked when I wrote that equation on the blackboard for the first time that in a course of my own design I would feel an obligation to try to encapsulate the train of thought that *led* Schrödinger to his equation (1926), but that I was determined on this occasion to adhere rigorously to the text. Later, however, I was approached by several students who asked if I *would* consider interpolating an account of the historical events I had felt constrained to omit. That I attempt to do here. It seems to me a story from which useful lessons can still be drawn.

1. Prior events. Schrödinger cultivated soil that had been prepared by others. Planck was led to write $E = h\nu$ by his successful attempt (1900) to use the then-recently-established principles of statistical mechanics to account for the spectral distribution of thermalized electromagnetic radiation. It was soon appreciated that Planck's energy/frequency relation was relevant to the understanding of optical phenomena that take place far from thermal equilibrium (photoelectric effect: Einstein 1905; atomic radiation: Bohr 1913). By 1916 it had become clear to Einstein that, while light is in some contexts well described as a Maxwellian wave, in other contexts it is more usefully thought of as a massless particle, with energy $E = h\nu$ and momentum $p = h/\lambda$. The mechanical statement $E = cp$[1] and the wave relation $\lambda\nu = c$ become then alternative ways of saying the same thing. Einstein had in effect "invented the photon" (which, however, did not acquire its name until 1926, when it was bestowed by the American chemist, Gilbert Lewis).

In 1923, Louis de Broglie—then a thirty-one-year-old graduate student at the University of Paris—speculated that if light waves are in some respects

[1] This is the relativistic statement $E = c\sqrt{p^2 + (mc)^2}$ in the limit $m \downarrow 0$.

particle-like, perhaps particles are in some respects wave-like, with frequency and wavelength that stand in the relation $(h\nu) = c\sqrt{(h/\lambda)^2 + (mc)^2}$.[2] It was obvious to others that if particles are indeed in some respects "wave-like" then they should exhibit the interference and diffractive effects most characteristic of waves. Electron diffraction was in fact observed by Davisson & Germer at Bell Labs in 1927, but by then Schrödinger had already completed his work.

Schrödinger cannot have been the only one to appreciate that the wave-like properties of particles—if any—must be describable by a wave equation of some sort. Because de Broglie had been led to his idea by Einstein and relativity, so also—and for that reason—were Schrödinger's initial efforts relativistic. He observed that if one looks for planewave solutions

$$\varphi(\boldsymbol{x}, t) = \exp\left\{\tfrac{i}{\hbar}[\,\boldsymbol{p}\cdot\boldsymbol{x} - Et]\right\}$$

of

$$\left\{\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} + (mc/\hbar)^2\right\}\varphi = 0$$

one immediately recovers $(E/c)^2 = \boldsymbol{p}\cdot\boldsymbol{p} + (mc)^2$, but when he attempted to use a modification of the preceding equation[3] to compute the energy spectrum of hydrogen he obtained results that were not in agreement with the measured values.[4]

---

[2] This is more neatly formulated $\omega = c\sqrt{k^2 + \varkappa^2}$, where $\omega = 2\pi\nu$, $k = 2\pi/\lambda$ and $\varkappa = mc/\hbar$. De Broglie had written $mc^2 = h\nu$ and interpreted ν to refer to some kind of "interior vibration." He then set the particle into motion and observed that m increased while ν—owing to time dilation—decreased, so had a sophomoric "paradox" on his hands. It was from an effort to resolve the paradox—not from mere idle word play—that his idea was born. For an English translation of his brief paper, visit http://www.davis-inc.com/physics/. At the end of his paper de Broglie remarked that if a mass pursues a circular orbit of radius r with momentum p, and if we can assume that Einstein's $p = h/\lambda$ pertains not only to light quanta but also to particles, then Bohr's quantization of angular momentum $rp = n\hbar$ can be interpreted as the stipulation that the circumference of the orbit be integrally many wavelengths long: $2\pi r = n\lambda$. The architecture of the Bohr atom could thus be understood in a new way. This modest accomplishment notwithstanding, de Broglie's advisors were so dubious that they sent the text of his work to Einstein for evaluation. Einstein, on the other hand, was so favorably impressed that he appended to a paper of his own—then in press—a note reporting that he had learned of an idea put forward by one Louis de Broglie that he thought merited closer study. It was this note that engaged Schrödinger's attention.

[3] It is known today as the Klein-Gordon equation, and is of fundamental importance in a connection distinct from the one originally contemplated by Schrödinger.

[4] See L. I. Schiff, *Quantum Mechanics* (3rd edition 1968, page 470) for the details, which Schrödinger published in 1926 after his non-relativistic papers were already in print.

So he put the work aside, not—so the story goes—to take it up again until Peter Debye asked him to present a seminar at the ETH (Eidgenössische Technische Hochschule) in Zürich, to which they were both attached at the time. For the occasion, Schrödinger chose to develop a *non-relativistic* version of his earlier work. Schrödinger (1887–1961) was 38 in 1925, only eight years younger than Einstein, five years younger than Born, two years younger than Bohr and—by a wide margin—the "old man" among the founding fathers of the "new quantum theory" (Pauli was 25, Heisenberg was 24, Dirac was 23). Partly as a result of his unusually thorough and well-rounded formal education (he had been a student of Fritz Hasenöhrl, Boltzmann's successor at the University of Vienna) and partly because of his prior research experience, Schrödinger's knowledge of physics appears to have been in many respects much broader than that of his junior colleagues. This fact made it possible for Schrödinger to approach his objective with an erudite circumspection that was, I suspect, quite beyond their reach. Schrödinger was uniquely positioned to do what he did.

Schrödinger remarked that optics came into the world as a theory of rays, which broadened into a wave theory from which ray optics could be recovered as a certain (high-frequency) approximation, and that at the interface between the two theories stood a certain elegant formalism ("theory of the characteristic function") that had been devised by Hamilton—devised *without* reference to wave theory—nearly a century earlier.[5] Schrödinger was aware also that Hamilton himself had noticed (then forgotten, then noticed again) that it is possible to construct a formalism (known today as Hamilton-Jacobi theory) that stands to the classical theory of particle trajectories as the theory of characteristic functions stands to the classical theory of rays. Schrödinger proposed to **complete the optico-mechanical analogy** by constructing a "wave mechanics of particles" that stands to Hamilton-Jacobi theory (whence to the theory of particle trajectories) as optical wave theory stands to characteristic function theory (whence to the theory of optical rays).

2. Review of the situation on the optical side of the analogy. For the purposes at hand it serves to pretend that light is a *scalar* wave phenomenon, described in vacuum by the simple wave equation

$$\nabla^2\varphi(\boldsymbol{x},t) = \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\varphi(\boldsymbol{x},t) \tag{1}$$

It is from the linearity of this equation (*i.e.*, from the availability of a principle of superposition) that optical interference effects arise, in which connection

[5] Hamilton's ideas were developed in a pair of papers that were published under the title "Theory of systems of rays" in 1830 and 1832. The central idea was adapted to mechanics in "On a general method in dynamics" (1834). By the late 19$^{\text{th}}$ Century Hamilton-Jacobi methods were used routinely to study the 3-body problem, the stability of the solar system and similar problems. A participant in that work was H. Bruns, who in 1895 had the bright idea of using those methods to in effect reinvent Hamiltonian optics.

we recall that local field intensity is proportional to the *square* of local field strength.

Separation of the time variable leads to the theory of *monochromatic* wave fields

$$\varphi(\boldsymbol{x},t) = f(\boldsymbol{x})\cdot e^{-i\omega t} \quad \text{with} \quad \nabla^2 f(\boldsymbol{x}) + k^2 f(\boldsymbol{x}) = 0 \quad : \quad k^2 \equiv (\omega/c)^2$$

Monochromatic *plane*waves are produced in cases of the type

$$f(\boldsymbol{x}) \sim e^{ik\,\hat{\boldsymbol{n}}\cdot\boldsymbol{x}}$$

Surfaces of constant phase are in such cases planes normal to the unit vector $\hat{\boldsymbol{n}}$, defined by equations of the form

$$\hat{\boldsymbol{n}}\cdot\boldsymbol{x} = \text{constant}$$

We then have

$$\begin{aligned}\varphi(\boldsymbol{x},t) &\sim e^{i[k\,\hat{\boldsymbol{n}}\cdot\boldsymbol{x} - \omega t]}\\ &\downarrow\\ &\sim e^{i[kx - \omega t]} \quad \text{in the one-dimensional case}\end{aligned}$$

To compute—in the one-dimensional case—the speed with which points of constant phase advance we set $\frac{d}{dt}[kx - \omega t] = 0$ and obtain $v = \omega/k = c$. In the three-dimensional case we note that the advance is along the local normal, so write $\boldsymbol{x} = x\,\hat{\boldsymbol{n}}$ and are led back to the same result.

If the vacuum is replaced by the simplest kind of refractive medium we expect the wave equation (1) to be replaced by

$$\nabla^2\varphi(\boldsymbol{x},t) = \Big[\frac{n(\boldsymbol{x})}{c}\Big]^2\frac{\partial^2}{\partial t^2}\varphi(\boldsymbol{x},t) \tag{2}$$
$$\Big[\frac{n(\boldsymbol{x})}{c}\Big] = \frac{1}{v(\boldsymbol{x})}$$

where $n(\boldsymbol{x})$ refers to the *local index of refraction*. Time-separation is still possible (which is to say: it is still possible to speak of monochromatic wave fields), but leads to the requirement that the space factor satisfy the modified equation

$$\nabla^2 f(\boldsymbol{x}) + k^2[n(\boldsymbol{x})]^2 f(\boldsymbol{x}) = 0 \tag{3}$$

It becomes natural to contemplate solutions of the form $f(\boldsymbol{x}) \sim e^{ikF(\boldsymbol{x})}$. We would then have

$$\varphi(\boldsymbol{x},t) \sim e^{i[kF(\boldsymbol{x}) - \omega t]}$$

and would find that he formerly planar surfaces of constant phase have become —at each instant—a space-filling population of generally curved surfaces of

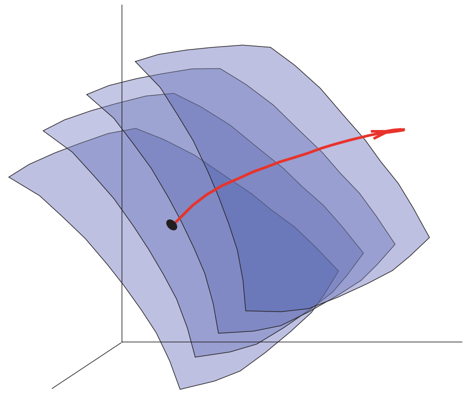

FIGURE 1: *The fundamental image from which both Hamiltonian optics and Hamilton-Jacobi theory derive.*

constant $F(\boldsymbol{x})$. Let $\Sigma_\wp(t)$ denote the surface defined by the equation $[kF(\boldsymbol{x}) - \omega t] = \wp$. Normal to that population of surfaces stands a population of curves, which we identify with a **system of light rays** (in Hamilton's phrase).

Let $\boldsymbol{x}$ refer to some point on $\Sigma_\wp(t)$, and $\mathcal{R}$ to the ray that stands normal to $\Sigma_\wp(t)$ at $\boldsymbol{x}$—that "punctures" $\Sigma_\wp(t)$ at $\boldsymbol{x}$. Time δt passes. The surface moves: $\Sigma_\wp(t) \longmapsto \Sigma_\wp(t+\delta t)$. And so also does the puncture point: $\boldsymbol{x} \longmapsto \boldsymbol{x} + v\hat{\boldsymbol{n}}\delta t$ (here $\hat{\boldsymbol{n}}$ is the unit vector that stands normal to $\Sigma_\wp(t)$—tangent to $\mathcal{R}$—at $\boldsymbol{x}$). From

$$\begin{aligned}\tfrac{d}{dt}[kF(\boldsymbol{x}) - \omega t] &= k\dot{\boldsymbol{x}}\cdot\nabla F - \omega\\ &= kv\hat{\boldsymbol{n}}\cdot\nabla F - \omega = 0\end{aligned}$$

we find that the speed with which the moving surface $\Sigma_\wp(t)$ sweeps past the point $\boldsymbol{x}$ on $\mathcal{R}$—in short: the **phase velocity** at $\boldsymbol{x}$—is given by

$$v(\boldsymbol{x}) = \frac{\omega}{k|\nabla F(\boldsymbol{x})|} = \frac{c}{\sqrt{\nabla F(\boldsymbol{x})\cdot\nabla F(\boldsymbol{x})}} \tag{4}$$

To discover the F-functions that are candidates for inclusion in the preceding discussion, we introduce $e^{ikF(\boldsymbol{x})}$ into (3) and obtain

$$-k^2\nabla F\cdot\nabla F + ik\nabla^2 F + k^2[n(\boldsymbol{x})]^2 = 0$$

It is a fact of optical (also acoustic) experience that the "ray" concept is useful only in situations where frequency is so high—and wavelength so short—as to cause diffractive effects to be negligible. With that fact in mind, we divide by k^2 and proceed to the limit $k \uparrow \infty$, obtaining the **eikonal equation**[6]

$$\nabla F\cdot\nabla F = [n(\boldsymbol{x})]^2 \tag{5}$$

[6] The terminology is Brun's, who called Hamilton's "characteristic function" (or "principal function") the "eikonal," after the Greek for icon, likeness, image.

This non-linear partial differential equation is structurally identical to (4), which can be written

$$[c/v(\boldsymbol{x})]^2 = \nabla F \cdot \nabla F \tag{6}$$

But at (4) we imagined ourselves to be evaluating $v(\boldsymbol{x})$ after $F(\boldsymbol{x})$ has been prescribed, while at (5) it is $n(\boldsymbol{x})$ that has been prescribed, and $F(\boldsymbol{x})$ that is being evaluated. Taken together, the two statements imply that "phase velocity at $\boldsymbol{x}$" has meaning irrespective of which ray $\mathcal{R}$ through $\boldsymbol{x}$ one has in mind, and irrespective also of the specific "system of rays" that one has supposed contains $\mathcal{R}$ as a member—irrespective, that is to say, of the global geometry of the surfaces-of-constant-F from which that system (or "pencil") of rays has been assumed to radiate.

That geometrical optics can be formulated as a theory of surfaces (surfaces of constant $F(\boldsymbol{x})$, with $F(\boldsymbol{x})$ subject to (5)), with rays standing everywhere normal to those surfaces,[7] was first appreciated by Hamilton, rediscovered by Bruns.

The established properties of those surfaces permit one to say sharp things about the associated rays. Let ray $\mathcal{R}$ stand normal to $\Sigma_\wp(0)$ at $\boldsymbol{x}_A$, let $\boldsymbol{x}_B$ mark some down-stream point on $\mathcal{R}$, and let $\boldsymbol{x}(\lambda)$ provide a parametric description of the ray: $\boldsymbol{x}(0) = \boldsymbol{x}_A$ and $\boldsymbol{x}(1) = \boldsymbol{x}_B$. We watch the motion of the puncture point produced by the advancing surface $\Sigma_\wp(t)$. Specifically, we compute the time of flight $\boldsymbol{x}_A \longrightarrow \boldsymbol{x}_B$:

$$\begin{aligned} T[\boldsymbol{x}_A \xrightarrow[\text{along ray}]{\qquad} \boldsymbol{x}_B] &= \int_0^1 \frac{1}{v(\boldsymbol{x}(\lambda))}\sqrt{\dot{\boldsymbol{x}}(\lambda)\cdot\dot{\boldsymbol{x}}(\lambda)}\,d\lambda \quad : \quad \dot{\boldsymbol{x}}(\lambda) \equiv \frac{d\boldsymbol{x}(\lambda)}{d\lambda} \\ &= \tfrac{1}{c}\int n(\boldsymbol{x})\sqrt{\dot{\boldsymbol{x}}\cdot\dot{\boldsymbol{x}}}\,d\lambda \\ &= \frac{\text{“optical length” of the ray segment}}{c} \end{aligned}$$

It follows from the fact that the puncture point advances always normally to $\Sigma_\wp(t)$—which is to say: along the locally shortest path from $\boldsymbol{x}$ on $\Sigma_\wp(t)$ to $\Sigma_\wp(t+\delta t)$—that the time of flight along the ray $\boldsymbol{x}_A \longrightarrow \boldsymbol{x}_B$ is shorter than would be the time of flight along any alternative curve linking those points. Thus have we recovered Fermat's **principle of least time**. Geometrical optics is, however, a theory of curves and surfaces in which time plays actually no role (Fermat worked before it had been established experimentally that it even made sense to speak of a finite "speed of light"): what we are really talking about here is a **principle of shortest optical length**. In any event, from

$$\delta \int n(\boldsymbol{x})\sqrt{\dot{\boldsymbol{x}}\cdot\dot{\boldsymbol{x}}}\,d\lambda = 0$$

it follows that

$$\Big\{\frac{d}{d\lambda}\frac{\partial}{\partial\dot{\boldsymbol{x}}} - \frac{\partial}{\partial\boldsymbol{x}}\Big\} n(\boldsymbol{x})\sqrt{\dot{\boldsymbol{x}}\cdot\dot{\boldsymbol{x}}} = \mathbf{0}$$

[7] The imagery is not at all unfamiliar: think of the "lines of force" that stand everywhere normal to "surfaces of constant potential."

We are led thus to the so-called “**ray equations**”

$$\frac{d}{d\lambda}\Big[n\frac{\dot{\boldsymbol{x}}}{\sqrt{\dot{\boldsymbol{x}}\cdot\dot{\boldsymbol{x}}}}\Big] - \sqrt{\dot{\boldsymbol{x}}\cdot\dot{\boldsymbol{x}}}\,\nabla n = 0$$

which if we adopt arc-length parameterization (where $\sqrt{\dot{\boldsymbol{x}}\cdot\dot{\boldsymbol{x}}}$ reduces to unity) assumes the simpler form

$$\frac{d}{ds}\Big[n(\boldsymbol{x})\frac{d\boldsymbol{x}}{ds}\Big] - \nabla n(\boldsymbol{x}) = \mathbf{0} \tag{7}$$

In homogeneous media (where $\nabla n = \mathbf{0}$) we have simply $d^2\boldsymbol{x}/ds^2 = \mathbf{0}$: all rays become straight lines. Note that (7) describes not a t-parameterized flight $\boldsymbol{x}(t)$ but an s-parameterized curve (ray), with $ds = \sqrt{dx^2 + dy^2 + dz^2}$: to obtain flight one would have to develop s as a function of t, writing $\boldsymbol{x}(s(t))$.

Hamilton’s own line of argument was the reverse of the argument I have presented: he proceeded from the theory of rays (actually, systems of rays) to the equation (5) that governs the structure of “characteristic functions” $W(\boldsymbol{x})$. He therefore had no reason to identify surfaces of constant W with surfaces of constant phase, no reason to consider the possible relevance of a wave equation to the geometrical theory he had constructed.

3. Derivation of the Hamilton-Jacobi equation. To simplify the writing, let us agree to work now in one-dimensional space, and to use subscripts to signify partial derivatives: thus $f_x(x,y) = \partial f(x,y)/\partial x$. At time t let a nice function $S(x,t)$ be defined on configuration space. Interpret

$$p(x,t) = S_x(x,t)$$

to inscribe a curve on phase space. From Hamilton’s canonical equations

$$\begin{aligned}\dot{x} &= +H_p(x,p)\\ \dot{p} &= -H_x(x,p)\end{aligned}$$

it follows that in time δt the points in phase space move

$$\begin{aligned}x &\longmapsto x + H_p(x,p)\delta t\\ p &\longmapsto p - H_x(x,p)\delta t\end{aligned}$$

Our objective is to describe the induced adjustment $S(x,t) \longmapsto S(x,t+\delta t)$. We proceed by noticing that on the one hand

$$S_x(x,t) \longmapsto S_x(x,t) - H_x(x,S_x)\delta t$$

while the expression on the right must be describable also as

$$S_x(x + H_p(x,S_x)\delta t, t+\delta t) = S_x(x,t) + S_{xx}(x,t)H_p(x,S_x)\delta t + S_{xt}(x,t)\delta t$$

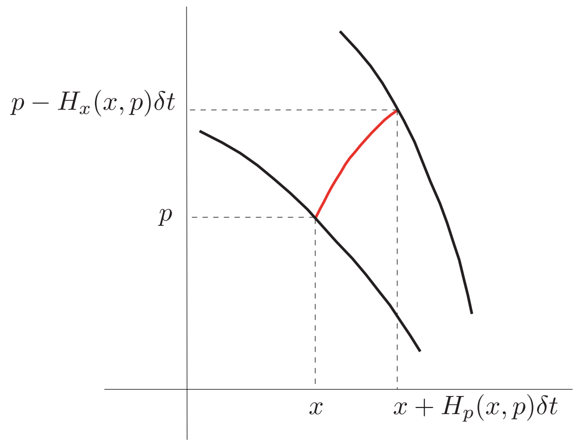

FIGURE 2: *"Canonical drift" of a curve that has been inscribed on phase space. Hamilton-Jacobi theory addresses the drift of curves (surfaces) that have been inscribed on phase space by the equation* $\boldsymbol{p}(t) = \nabla S(\boldsymbol{x}, t)$.

We conclude that

$$H_x(x, S_x) + H_p(x, S_x)S_{xx} + S_{xt} = 0$$

But this can be written $[H(x, S_x) + S_t]_x = 0$, which entails

$$H(x, S_x) + S_t = \text{arbitrary function of } t$$

A physically inconsequential adjustment $S \longmapsto S - \int^t$(arbitrary function) could always be used to extinguish the term on the right, so we discard it: we are left with

$$H(x, S_x) + S_t = 0 \tag{8}$$

which is the celebrated **Hamilton-Jacobi equation**.

If we look for solutions of the separated form $S(x,t) = W(x) + f(t)$ we discover at once that necessarily $f(t) = \text{constant}$, which (partly for dimensional reasons) we will agree to call E. We then have the time-independent Hamiltonian-Jacobi equation

$$H(x, W_x) = E \tag{9}$$

In the most commonly encountered case $H(x,p) = \frac{1}{2m}\boldsymbol{p}\cdot\boldsymbol{p} + U(\boldsymbol{x})$ we have

$$\tfrac{1}{2m}\nabla W \cdot \nabla W + U(\boldsymbol{x}) = E \tag{10}$$

Most typically (as in the case just considered) the Hamiltonian $H(\boldsymbol{x}, \boldsymbol{p})$ depends *non-linearly* upon $\boldsymbol{p}$, so typically (8) and (9)—which in several dimensions read

$$H(\boldsymbol{x}, \nabla S) + S_t = 0 \tag{11.1}$$

and

$$H(\boldsymbol{x}, \nabla W) = E \tag{11.2}$$

with $S = S(\boldsymbol{x}, t)$ and $W = W(\boldsymbol{x})$—are non-linear partial differential equations.

The Newtonian, Lagrangian and canonical formalisms culminate in systems —often quite large *systems—of coupled ordinary* differential equations. It is remarkable that in (11) all that physics has come to rest in a *single partial* differential equation, and noteworthy that, while it is ∇U that enters into the equations of motion supplied by the formalisms just mentioned, $U(\boldsymbol{x})$ enters nakedly/undifferentiatedly into the Hamilton-Jacobi equation. . . as it will also into the Schrödinger equation.

It is upon the foundation provided by the elegant material just presented—a body of theory that was more than ninety years old by the time he decided to accept de Broglie/Einstein's challenge—that Schrödinger erected his "wave mechanics."

4. Completion of the optico-mechanical analogy. Note first that from $p = S_x$ it follows that dimensionally

$$[S] = (\text{length})(\text{momentum}) = \text{action} = [\hbar]$$

Pursuing now in reverse the change of dependent variable that led us from (2) to (5/6), we write

$$\psi(\boldsymbol{x}, t) = \psi_0\, e^{iS(\boldsymbol{x},t)/\hbar} \tag{12}$$

and observe that

$$\begin{aligned} -\frac{\hbar^2}{2m}\nabla^2\psi &= \Big\{\frac{1}{2m}\nabla S\cdot\nabla S - i\frac{\hbar}{2m}\nabla^2 S\Big\}\cdot\psi \\ U\psi &= \Big\{U\Big\}\cdot\psi \\ -i\hbar\partial_t\psi &= \Big\{S_t\Big\}\cdot\psi \end{aligned}$$

from which it follows (by $i(\hbar/2m)\psi\nabla^2 S = (\hbar^2/2m)\psi\nabla^2\log[\psi/\psi_0]$) that

$$-\frac{\hbar^2}{2m}\big[\nabla^2\psi - \psi\nabla^2\log[\psi/\psi_0]\big] + U\psi - i\hbar\psi_t = \Big\{\frac{1}{2m}\nabla S\cdot\nabla S + U + S_t\Big\}\cdot\psi$$

We conclude that the equation

$$-\frac{\hbar^2}{2m}\big[\nabla^2\psi - \psi\nabla^2\log[\psi/\psi_0]\big] + U\psi - i\hbar\psi_t = 0$$

is but a renotated variant of—and entirely equivalent to—the Hamilton-Jacobi equation, an eccentric way of rendering *classical* mechanics. Schrödinger had, however, the wit to <u>abandon the $\psi\nabla^2\log[\psi/\psi_0]$ term</u>. He was left then with a *linear* field equation (we know that linearity in quantum mechanics is the name of the game!)—the time-dependent **Schrödinger equation**

$$-\frac{\hbar^2}{2m}\nabla^2\psi + U\psi = i\hbar\psi_t \tag{13}$$

which is manifestly *not* just old wine in a new bottle.

It is instructive to pursue the preceding argument in reverse: insert[8]

$$\psi(\boldsymbol{x},t) = R(\boldsymbol{x},t)\,e^{iS(\boldsymbol{x},t)/\hbar}$$

into the Schrödinger equation (13) and obtain an equation the real and imaginary parts of which can be written

$$\tfrac{1}{2m}\nabla S\cdot\nabla S + U(\boldsymbol{x}) - \tfrac{\hbar^2}{2m}\frac{\nabla^2 R}{R} + S_t = 0 \tag{14.1}$$

$$\partial_t R^2 + \nabla\cdot\big[\tfrac{1}{m}R^2\nabla S\big] = 0 \tag{14.2}$$

respectively. These coupled equations are—conjointly—entirely equivalent to the Schrödinger equation. The latter possesses the form

$$\tfrac{\partial}{\partial t}(\text{density}) + \nabla\cdot(\text{current}) = 0 \quad : \quad \text{density} = R^2 = \psi^*\psi$$

of a "continuity equation," and evidently describes the local conservation of probability.[9] Equation (14.1) gives back the Hamilton-Jacobi in the classical limit $\hbar \downarrow 0$. Recall that a limiting process entered also into the derivation of the eikonal equation (5) from the wave equation (2).

Curiously, one might look upon the Schrödinger equation as a "linearized Hamilton-Jacobi equation." Even if the quantum world did not exist, it would make computational good sense to adopt the following 3-step program

- linearize
- solve the resulting "Schrödinger equation"
- proceed to the classical limit $\hbar \downarrow 0$

as a way of solving *classical* problems.

If one introduces the time-separated wave function

$$\psi(\boldsymbol{x},t) = R(\boldsymbol{x})e^{i[W(\boldsymbol{x}) - Et]/\hbar}$$

into the Schrödinger equation (16) one obtains

$$\tfrac{1}{2m}\nabla W\cdot\nabla W = E - U(\boldsymbol{x}) + \tfrac{\hbar^2}{2m}\frac{\nabla^2 R}{R} \tag{15.1}$$

$$\nabla\cdot\big[\tfrac{1}{m}R^2\nabla W\big] = 0 \tag{15.2}$$

[8] We use this polar alternative to (12) to insure the reality of $R(\boldsymbol{x},t)$ and $S(\boldsymbol{x},t)$.

[9] Note that Schrödinger's derivation of (13) did not carry with it any interpretive advice. Nor did it have anything to say about the (very unclassical) quantum measurement process: all of that came later (Born, von Neumann).

In the classical limit the former equation decouples, to give back the time-independent Hamilton-Jacobi equation (10):

$$\nabla W \cdot \nabla W = 2m[E - U(\boldsymbol{x})]$$

5. The "rays" traced by classical particles. We now take leave of Schrödinger to trace this story back to point of departure. In recent discussion we have proceeded

$$\text{classical H-J theory} \longrightarrow \text{quantum wave theory}$$

and I propose now to proceed in the reverse direction:

$$\text{classical particle trajectories} \longleftarrow \text{classical H-J theory}$$

It is by way of preparation that we distinguish several "natural velocities" (speeds) that arise from the theory of waves and particles.

Recall that the **phase velocity** of the simple wave $e^{i[kx - \omega(k)t]}$, got by setting $\frac{d}{dt}[kx - \omega(k)t] = 0$, is given by

$$v_{\text{phase}} = \frac{\omega(k)}{k}$$

while the **group velocity**, got by looking to the motion of the envelope when such waves of nearly identical k-values are superimposed,[10] is given by

$$v_{\text{group}} = \frac{d\omega(k)}{dk}$$

If we use de Broglie's relations $E = h\nu = \hbar\omega$ and $p = h/\lambda = \hbar k$ to translate those equations into mechanical variables we obtain

$$v_{\text{phase}} = \frac{E(p)}{p}$$

$$v_{\text{group}} = \frac{dE(p)}{dp}$$

which in the case $E(p) = \frac{1}{2m}p^2 + U$ become

$$v_{\text{phase}}(\boldsymbol{x}) = \frac{E}{\sqrt{2m[E - U(\boldsymbol{x})]}}$$

$$v_{\text{group}}(\boldsymbol{x}) = \tfrac{1}{m}p = \tfrac{1}{m}\sqrt{2m[E - U(\boldsymbol{x})]}$$

Recall also that a mass point m, if moving with conserved energy E in the

[10] See Chapter 6, §5 in my SOPHOMORE NOTES (2005).

presence of a potential $U(\boldsymbol{x})$, necessarily has speed

$$v_{\scriptscriptstyle E}(\boldsymbol{x}) = \sqrt{\tfrac{2}{m}[E - U(\boldsymbol{x})]}$$

when it passes through point $\boldsymbol{x}$, whatever the direction in which it is moving. That

$$v_{\scriptscriptstyle E}(\boldsymbol{x}) = v_{\text{group}}(\boldsymbol{x})$$

conforms to our experience that in quantum mechanics classical particle motion is represented by the motion of wave *packets*.

In "Geometrical mechanics: remarks commemorative of Heinrich Hertz"[11] I show it to be an implication of Newton's $m\ddot{\boldsymbol{y}}(t) = -\nabla U(\boldsymbol{y})$ that one can always write $\boldsymbol{y}(t) = \boldsymbol{x}(s(t))$ where

- $\boldsymbol{x}(s)$ describes the *path* traced by the particle (mechanical analog of a "ray") and s signifies Euclidean arc length ;
- $s(t)$ describes *progress along* that path, and is got by integrating $\dot{s} = v_{\scriptscriptstyle E}(\boldsymbol{x})$.

I am able to show, moreover, that $\boldsymbol{x}(s)$ satisfies a set of equations

$$\frac{d}{ds}\Big[n(\boldsymbol{x})\frac{d\boldsymbol{x}}{ds}\Big] - \nabla n(\boldsymbol{x}) = \mathbf{0} \tag{16}$$

that is structurally identical to the ray equations (7), except that here

$$n(\boldsymbol{x}) = \frac{c}{v_{\text{phase}}(\boldsymbol{x})} = \frac{\sqrt{2mc^2[E - U(\boldsymbol{x})]}}{E} = \frac{\sqrt{2mc^2}}{E}\sqrt{\text{kinetic energy } T}$$

where physically inconsequential factors have been introduced simply to render $n(\boldsymbol{x})$ dimensionless, and to highlight the formal identity of the mechanical equation with its optical counterpart. It follows already from our previous work that (16) can be obtained from[12]

$$\delta_{\scriptscriptstyle E}\int n(\boldsymbol{x}(s))ds \sim \delta_{\scriptscriptstyle E}\int \sqrt{T}\,ds \sim \delta_{\scriptscriptstyle E}\int T\,dt$$

This variational principle, first stated by Jacobi and later by Hertz, is the direct mechanical analog of Fermat's Principle of Least Time. Neither principle has anything to do with time or motion: both have to do with the design of curves, rays, trajectories, "flight paths." The Jacobi-Hertz **principle of least action** states that of all curves that are pursued $\boldsymbol{x}_A \longrightarrow \boldsymbol{x}_B$ with constant energy E, the curve realized by a particle in natural motion will be the one that extremizes the mechanical analog of optical path length.[13]

All of which emerges quite naturally when (with Hamilton) we identify those curves with the curves normal to surfaces of constant $S(\boldsymbol{x}, t) = W(\boldsymbol{x}) - Et$ and observe that those surfaces advance with speed

$$v(\boldsymbol{x}) = \frac{E}{\sqrt{\nabla W\cdot\nabla W}} = \frac{E}{\sqrt{2m[E - U(\boldsymbol{x})]}} = v_{\text{phase}}(\boldsymbol{x})$$

---

[11] Notes for a Reed College Physics Seminar presented 23 February 1994.

[12] The final step here follows from $\dot{s}^2 = \frac{m}{2}T$; *i.e.*, from $ds \sim \sqrt{T}\,dt$.

[13] For detailed discussion see H. Goldstein, C. Poole & J. Safko, *Classical Mechanics* (3rd edition 2002), §8.6.

I note in conclusion that David Bohm has proposed that, after one has worked to obtain the physically interesting solution to the Schrödinger equation, one use that information to evaluate what he calls the “quantum potential”

$$Q(\boldsymbol{x},t) = -\frac{\hbar^2}{2m}\frac{\nabla^2 R}{R}$$

—in which notation (14.1) reads

$$\tfrac{1}{2m}\nabla S\cdot\nabla S + U(\boldsymbol{x}) + Q(\boldsymbol{x},t) + S_t = 0$$

Bohm would have us use that modified Hamilton-Jacobi equation to obtain a “system of rays” (particle trajectories) in precisely the manner that one would employ classically. On this basis he erects what he calls a “causal interpretation of quantum mechanics.” Most physicsts (I among them) consider Bohm’s proposal to be, while not technically in error, profoundly misguided—a complex embellishment of a theory that stands in no need of embellishment. For an excellent account of Bohm’s theory, see P. R. Holland, *The Quantum Theory of Motion*: *An Account of the de Broglie-Bohm Causal Interpretation of Quantum Mechanics* (1993), which one reviewer called “a good book about a bad theory.”

ADDENDUM

Hamilton’s <u>Other</u> Contribution to the Development of Quantum Mechanics

Introduction. Hamilton-Jacobi theory comes in *two* flavors. The first—the “one-point theory,” the theory to which we have thus far restricted our attention, and upon which Schrödinger drew—is concerned with functions $S(\boldsymbol{x},t)$ of a single space-time point that satisfy

$$H(\boldsymbol{x},\nabla S) + S_t = 0$$

and in Schrödinger’s hands gave rise to what might be called “one point wave mechanics”:

$$S(\boldsymbol{x},t) \xrightarrow[\text{Schrödinger quantization}]{} \psi(\boldsymbol{x},t)$$

The second—the two-point theory—is concerned with functions $S(\boldsymbol{x}_1,t_1;\boldsymbol{x}_0,t_0)$ of variables that refer to a *pair* of space-time points, and that satisfy a *pair* of Hamilton-Jacobi equations:

$$\left.\begin{aligned} H(\boldsymbol{x}_1,+\nabla_{\!1}S) + S_{t_1} &= 0\\ H(\boldsymbol{x}_0,-\nabla_{\!0}S) - S_{t_0} &= 0 \end{aligned}\right\} \tag{17}$$

The quantum mechanical relevance of the two-point theory remained generally unrecognized until 1942, when it acquired a central place in what was, in effect, Feynman’s *reinvention* of wave mechanics:

$$S(\boldsymbol{x},t;\boldsymbol{x}_0,t_0) \xrightarrow[\text{Feynman quantization}]{} K(\boldsymbol{x},t;\boldsymbol{x}_0,t_0)$$

Here $K(\boldsymbol{x},t;\boldsymbol{x}_0,t_0)$ is the Green’s function or **propagator** that describes the *dynamical evolution* of Schrödinger's wave function:

$$\psi(\boldsymbol{x},t) = \int K(\boldsymbol{x},t;\boldsymbol{x}_0,t_0)\,\psi(\boldsymbol{x}_0,t_0)\,d\boldsymbol{x}_0$$

The two-point propagator satisfies a *pair* of Schrödinger equations

$$\left.\begin{aligned}\big\{H(\boldsymbol{x}\ \ ,+\tfrac{\hbar}{i}\nabla\ \ ) - i\hbar\partial_t\ \big\}K(\boldsymbol{x},t;\boldsymbol{x}_0,t_0) = 0\\ \big\{H(\boldsymbol{x}_0,-\tfrac{\hbar}{i}\nabla_0) + i\hbar\partial_{t_0}\big\}K(\boldsymbol{x},t;\boldsymbol{x}_0,t_0) = 0\end{aligned}\right\}\tag{18.1}$$

together with the initial condition

$$\lim_{t\downarrow t_0} K(\boldsymbol{x},t;\boldsymbol{x}_0,t_0) = \delta(\boldsymbol{x}_0 - \boldsymbol{x})\tag{18.2}$$

Here I undertake to describe the most salient features of two-point Hamilton-Jacobi theory, and to indicate how that theory relates to Feynman's "sum-over-paths" formulation of non-relativistic quantum mechanics.

5. Two-point Hamilton-Jacobi theory. Let $L(\boldsymbol{x},\dot{\boldsymbol{x}})$ be the Lagrangian of whatever mechanical system we may have in mind, and let the well-behaved function $\boldsymbol{x}(t)$ describe a hypothetical "path" (curve inscribed on $(n+1)$-dimensional spacetime) that connects a specified pair of endpoints: $\boldsymbol{x}(t_0)=\boldsymbol{x}_0$, $\boldsymbol{x}(t_1)=\boldsymbol{x}_1$. To every such path we associate the "action functional"

$$S[\text{path}] \equiv \int_{t_0}^{t_1} L\big(\boldsymbol{x}(t),\dot{\boldsymbol{x}}(t)\big)\,dt$$

Hamilton's principle (which it was Feynman's habit to call the "principle of least action") asserts that the natural or *dynamical path* $(\boldsymbol{x}_1,t_1)\longleftarrow(\boldsymbol{x}_0,t_0)$ is the path[14] that extremizes the action:

$$\delta\int_{t_0}^{t_1} L\big(\boldsymbol{x}(t),\dot{\boldsymbol{x}}(t)\big)\,dt = 0 \quad\Rightarrow\quad \frac{d}{dt}\frac{\partial L}{\partial\dot{\boldsymbol{x}}} - \frac{\partial L}{\partial\boldsymbol{x}} = 0$$

Solutions of the equations of motion—though usually fixed by specificiation of initial data $(\boldsymbol{x}_0,\dot{\boldsymbol{x}}_0)$—can as well be fix by specification of endpoint data $(\boldsymbol{x}_0,\boldsymbol{x}_1)$. Thus does the path *functional* $S\big[(\boldsymbol{x}_1,t_1)\xleftarrow[\text{dynamical path}]{}(\boldsymbol{x}_0,t_0)\big]$ become a *function* of its endpoint coordinates:

$$S[\boldsymbol{x}_{\text{dynamical}}(t)] = S(\boldsymbol{x}_1,t_1;\boldsymbol{x}_0,t_0)$$

We have here made the acquaintance of the **two-point action function**,[15] the

---

[14] Of these there might in fact be several, in which case we would find ourselves talking about "local extrema in the space of paths." It serves my present purposes to ignore that possibility.

[15] . . . which in cases of the type mentioned in the preceding footnote may be multi-valued.

central object in two-point Hamilton-Jacobi theory, the characteristic properties of which I will illustrate by

> **EXAMPLE**: For a free particle one has $L(\boldsymbol{x},\dot{\boldsymbol{x}})=\frac{1}{2}m\dot{\boldsymbol{x}}\cdot\dot{\boldsymbol{x}}$. The equation of motion therefore reads $\ddot{\boldsymbol{x}}=\boldsymbol{0}$ and the dynamical path linking $(\boldsymbol{x}_0,t_0)$ to $(\boldsymbol{x}_1,t_1)$ is
>
> $$\boldsymbol{x}(t)=\boldsymbol{x}_0+\frac{\boldsymbol{x}_1-\boldsymbol{x}_0}{t_1-t_0}(t-t_0)$$
>
> It now follows that
>
> $$\begin{aligned} S_{\text{free particle}}(\boldsymbol{x}_1,t_1;\boldsymbol{x}_0,t_0) &= \int_{t_0}^{t_1}\frac{m}{2}\frac{\boldsymbol{x}_1-\boldsymbol{x}_0}{t_1-t_0}\cdot\frac{\boldsymbol{x}_1-\boldsymbol{x}_0}{t_1-t_0}\,dt \\ &= \frac{m}{2}\frac{(\boldsymbol{x}_1-\boldsymbol{x}_0)\cdot(\boldsymbol{x}_1-\boldsymbol{x}_0)}{t_1-t_0} \end{aligned} \tag{19}$$
>
> Immediately
>
> $$\left.\begin{aligned} \nabla_1 S &= +m\frac{\boldsymbol{x}_1-\boldsymbol{x}_0}{t_1-t_0} \quad\text{and}\quad S_{t_1} = -\frac{m}{2}\frac{(\boldsymbol{x}_1-\boldsymbol{x}_0)\cdot(\boldsymbol{x}_1-\boldsymbol{x}_0)}{(t_1-t_0)^2} \\ \nabla_0 S &= -m\frac{\boldsymbol{x}_1-\boldsymbol{x}_0}{t_1-t_0} \quad\text{and}\quad S_{t_0} = +\frac{m}{2}\frac{(\boldsymbol{x}_1-\boldsymbol{x}_0)\cdot(\boldsymbol{x}_1-\boldsymbol{x}_0)}{(t_1-t_0)^2} \end{aligned}\right\} \tag{20}$$
>
> from which it becomes obvious that
>
> $$\begin{aligned} \tfrac{1}{2m}\nabla_1 S\cdot\nabla_1 S+S_{t_1}&=0 \\ \tfrac{1}{2m}\nabla_0 S\cdot\nabla_0 S-S_{t_0}&=0 \end{aligned}$$
>
> But these are precisely the equations that are supplied by (17) in the present instance: $H(\boldsymbol{x},\boldsymbol{p})=\frac{1}{2m}\boldsymbol{p}\cdot\boldsymbol{p}$.

Though born of the Lagrangian formulation of classical mechanics, the two-point function $S(\boldsymbol{x}_1,t_1;\boldsymbol{x}_0,t_0)$ gains importance primarily from the role it plays in the theory of canonical transformations. Hamilton's canonical equations can be written

$$\begin{aligned} \dot{\boldsymbol{x}}&=-[H(\boldsymbol{x},\boldsymbol{p}),\boldsymbol{x}] \\ \dot{\boldsymbol{p}}&=-[H(\boldsymbol{x},\boldsymbol{p}),\boldsymbol{p}] \end{aligned}$$

where the Poisson bracket is defined $[A,B]\equiv\sum_k\left\{\frac{\partial A}{\partial x_k}\frac{\partial B}{\partial p_k}-\frac{\partial B}{\partial x_k}\frac{\partial A}{\partial p_k}\right\}$. The Hamiltonian acquires thus the role of the generator (in the sense of Sophus Lie) of infinitesimal canonical transformations

$$\begin{aligned} \boldsymbol{x}_0\mapsto\boldsymbol{x}(\delta t)&=\boldsymbol{x}_0-[H,\boldsymbol{x}]_0\delta t \\ \boldsymbol{p}_0\mapsto\boldsymbol{p}(\delta t)&=\boldsymbol{p}_0-[H,\boldsymbol{p}]_0\delta t \end{aligned}$$

that by iteration result finite t-parameterized phase flow:

$$\begin{aligned} \boldsymbol{x}_0&\mapsto\boldsymbol{x}(t;\boldsymbol{x}_0,\boldsymbol{p}_0) \\ \boldsymbol{p}_0&\mapsto\boldsymbol{p}(t;\boldsymbol{x}_0,\boldsymbol{p}_0) \end{aligned}$$

The function $S(\boldsymbol{x},t;\boldsymbol{x}_0,0)$ achieves that same finitistic result by different means: write

$$\begin{aligned}\boldsymbol{p}\ \ &= \boldsymbol{p}\ \ (\boldsymbol{x},t;\boldsymbol{x}_0,t_0) = +\nabla\ \ S(\boldsymbol{x},t;\boldsymbol{x}_0,t_0)\\ \boldsymbol{p}_0 &= \boldsymbol{p}_0(\boldsymbol{x},t;\boldsymbol{x}_0,t_0) = -\nabla_0 S(\boldsymbol{x},t;\boldsymbol{x}_0,t_0)\end{aligned}$$

By functional inversion (often more easily said than done!) of the latter equation obtain

$$\boldsymbol{x} = \boldsymbol{x}(t;\boldsymbol{x}_0,\boldsymbol{p}_0,t_0)$$

Then return with that information to the first equation to obtain

$$\boldsymbol{p} = \boldsymbol{p}(t;\boldsymbol{x}_0,\boldsymbol{p}_0,t_0)$$

We recognize the sequence of operations just described to comprise an instance of Legendre's procedure for promoting derivatives to the status of independent variables. Evidently

- the Hamiltonian $H(\boldsymbol{x},\boldsymbol{p})$ is the **Lie generator** of dynamical flow in phase space, and does its work incrementally;
- the two-point action function $S(\boldsymbol{x},t;\boldsymbol{x}_0,t_0)$ is the **Legendre generator** of that evolving canonical transformation, and does its work wholistically

and the Hamilton-Jacobi equations (17) describe the *relationship between* those generators.

EXAMPLE REVISITED: Pressing the free particle Hamiltonian $H(\boldsymbol{x},\boldsymbol{p}) = \frac{1}{2m}\boldsymbol{p}\cdot\boldsymbol{p}$ into service as a Lie generator, we have

$$\begin{aligned}\boldsymbol{x}(t) &= \boldsymbol{x}_0 - t[H,\boldsymbol{x}]_0 + \tfrac{1}{2}t^2[H,[H,\boldsymbol{x}]]_0 - \cdots\\ &= \boldsymbol{x}_0 + t\tfrac{1}{m}\boldsymbol{p}_0 \qquad (21.1)\\ \boldsymbol{p}(t) &= \boldsymbol{p}_0 - t[H,\boldsymbol{p}]_0 + \tfrac{1}{2}t^2[H,[H,\boldsymbol{p}]]_0 - \cdots\\ &= \boldsymbol{p}_0 \qquad (21.2)\end{aligned}$$

while from (20) we have

$$\begin{aligned}\boldsymbol{p}\ \ &= +\nabla\ \ S = m\frac{\boldsymbol{x}-\boldsymbol{x}_0}{t}\\ \boldsymbol{p}_0 &= -\nabla_0 S = m\frac{\boldsymbol{x}-\boldsymbol{x}_0}{t}\end{aligned}$$

Functional inversion of the latter gives

$$\boldsymbol{x}(t) = \boldsymbol{x}_0 + t\tfrac{1}{m}\boldsymbol{p}_0 \qquad (22.1)$$

which when introduced into the former equation gives back

$$\boldsymbol{p}(t) = \boldsymbol{p}_0 \qquad (22.2)$$

Comparison of (21) with (22) shows that—in this instance, as generally—the two procedures lead to identical descriptions of the dynamical phase flow.

It is obvious—yet instructive to notice—that two-point functions become single-point functions (of a certain type) when the second argument is frozen: $S(\boldsymbol{x},t;\bullet,\bullet) = S(\boldsymbol{x},t)$. Such neutered functions are, however, incapable of generating canonical transformations. I note also that constants of the motion arise within H-theory from conditions of the form $[H,A]=0$, but in S-theory from *symmetries* of the two-point action (Noether's theorem), but will refrain from discussion of how those two notions come to be interconnected.

6. Quantum mechanics according to Feynman. The train of thought that led the 23-year-old Richard Feynman (1918–1988) to his very fruitful reinvention of non-relativistic quantum mechanics can be traced to an occasion when he asked a drinking buddy whether he "had ever encountered a quantum mechanical application of the principle of least action" (his allusion being actually to Hamilton's principle), and was referred to an obscure paper by P.A.M. Dirac.[16]

The objects of central interest to Dirac (as also to Feynman after him) were not "one-point Schrödinger functions" $\psi(\boldsymbol{x},t)$ but the "two-point Schrödinger functions" $K(\boldsymbol{x},t;\boldsymbol{x}_0,t_0)$ that arise in $\boldsymbol{x}$-representation

$$K(\boldsymbol{x},t;\boldsymbol{x}_0,t_0) = (\boldsymbol{x}|\mathsf{U}(t;t_0)|\boldsymbol{x}_0)$$

of the unitary operator that (in the Schrödinger picture) sends $|\psi)_{t_0} \longmapsto |\psi)_t$, and that, as was remarked at (18), become single-point functions (of a certain type) when the second argument is frozen: $K(\boldsymbol{x},t;\bullet,\bullet) = \psi(\boldsymbol{x},t)$.

Necessarily

$$\mathsf{U}(t_1;t_0) = \mathsf{U}(t_1;t)\mathsf{U}(t;t_0) \quad : \quad t_1 \geqslant t \geqslant t_0$$

which in $\boldsymbol{x}$-representation becomes the slightly less obvious **composition rule**

$$K(\boldsymbol{x}_1,t_1;\boldsymbol{x}_0,t_0) = \int K(\boldsymbol{x}_1,t_1;\boldsymbol{x},t)\,d\boldsymbol{x}\,K(\boldsymbol{x},t;\boldsymbol{x}_0,t_0)$$

Feynman—typically—attached to that equation a much more robustly physical interpretation than Dirac, in his sophistication, had reason to do:[17] Feynman considered it to state that

$$\begin{aligned}&\text{probability amplitude of transition } (\boldsymbol{x}_1,t_1) \longleftarrow (\boldsymbol{x}_0,t_0)\\ &\qquad\qquad = \int \big\{\text{amplitude of transition \textit{via} the intermediate point } (\boldsymbol{x},t)\big\}\,d\boldsymbol{x}\end{aligned}$$

---

[16] "The Lagrangian in quantum mechanics," Physikalische Zeitschrift der Sowjetunion **3**, 1933. This beautiful paper is reprinted in J. Schwinger (editor), *Selected Papers on Quantum Mechanics* (1958)—a valuable collection of classic papers known commonly as "the Schwinger collection."

[17] Dirac, working in the Heisenberg picture, looked upon $(\boldsymbol{x}_1,t_1|\,\boldsymbol{x}_0,t_0)$ as an element of a transformation matrix that relates one representation to another.

Dirac noted—and Feynman later made much of the fact— that by repeated application of the composition rule one can resolve $K(\boldsymbol{x}_1,t_1;\boldsymbol{x}_0,t_0)$ into iterated *short-time propagators*

$$K(\boldsymbol{x}_1,t_1;\boldsymbol{x}_0,t_0)=\int\!\cdots\!\int K(\boldsymbol{x}_1,t_1;\boldsymbol{y}_N,t_1-\tau)\,d\boldsymbol{y}_N\cdots d\boldsymbol{y}_1 K(\boldsymbol{y}_1,t_0+\tau;\boldsymbol{x}_0,t_0)$$

Dirac—who considered himself to be simply pointing out a formal parallelism between quantum mechanical transformation theory (in the development of which he had played the leading role) and the classical theory of canonical transformations—was led at this point to remark that

$$K(\boldsymbol{y},t+\tau;\boldsymbol{x},t)\quad\text{“corresponds to”}\quad e^{iS(\boldsymbol{y},t+\tau;\boldsymbol{x},t)/\hbar}$$

and that therefore $K(\boldsymbol{x}_1,t_1;\boldsymbol{x}_0,t_0)$ “corresponds (in the limit $N\uparrow\infty$) to”

$$\begin{aligned}&\int\!\cdots\!\int e^{iS(\boldsymbol{x}_1,t_1;\boldsymbol{y}_N,t_1-\tau)/\hbar}\,d\boldsymbol{y}_N\cdots d\boldsymbol{y}_1 e^{iS(\boldsymbol{y}_1,t_0+\tau;\boldsymbol{x}_0,t_0)/\hbar}\\&\qquad=\int\!\cdots\!\int \exp\Big\{(i/\hbar)\int_{t_0}^{t_1}L(\boldsymbol{x}_1\xleftarrow[\boldsymbol{y}_N,\ldots,\boldsymbol{y}_2,\boldsymbol{y}_1]{}\boldsymbol{x}_0)\,dt\Big\}\,d\boldsymbol{y}_N\cdots d\boldsymbol{y}_1\end{aligned}$$

where $\tau=\frac{t_1-t_0}{N+1}$ and $\boldsymbol{x}_1\xleftarrow[\boldsymbol{y}_N,\ldots,\boldsymbol{y}_2,\boldsymbol{y}_1]{}\boldsymbol{x}_0$ refers to the segmented path $\boldsymbol{x}_1\longleftarrow\boldsymbol{x}_0$ that visits successively the points $\{\boldsymbol{y}_N,\ldots,\boldsymbol{y}_2,\boldsymbol{y}_1\}$. Dirac found it natural, in view of his objective, to point out (without explicit reference to the **method of stationary phase**) that in the classical limit $\hbar\downarrow 0$ the paths contemplated above tend collectively to “buzz themselves to extinction,” the only path immune to this fate being the path (or paths) whose nodal-points $\{\boldsymbol{y}_N,\ldots,\boldsymbol{y}_2,\boldsymbol{y}_1\}$ are so placed as to *extremize*[18]

$$S(\boldsymbol{x}_1\xleftarrow[\boldsymbol{y}_N,\ldots,\boldsymbol{y}_2,\boldsymbol{y}_1]{}\boldsymbol{x}_0)=\sum_{n=0}^{N}S(\boldsymbol{y}_{n+1},t_n+\tau;\boldsymbol{y}_n,t_n)$$

—placed, that is to say, so as to lie on the classical path $(\boldsymbol{x}_1,t_1)\leftarrow(\boldsymbol{x}_0,t_0)$.

Feynman—mystified by Dirac's “corresponds to”—chose tentatively to interpret that phrase to mean “equals” (actually: “equals to within an adjustable factor”) and to see where he was led. Setting

$$K(\boldsymbol{y},t+\tau;\boldsymbol{x},t)=\frac{1}{A(\tau)}\,e^{iS(\boldsymbol{y},t+\tau;\boldsymbol{x},t)/\hbar}\tag{23}$$

his conjecture read

$$\begin{aligned}&K(\boldsymbol{x}_1,t_1;\boldsymbol{x}_0,t_0)\\&\qquad=\lim_{N\uparrow\infty}\int\!\cdots\!\int\prod_{n=0}^{N}\frac{1}{A(\tau)}\,e^{iS(\boldsymbol{y}_{n+1},t_n+\tau;\boldsymbol{y}_n,t_n)/\hbar}\,d\boldsymbol{y}_1 d\boldsymbol{y}_2\cdots d\boldsymbol{y}_N\end{aligned}\tag{24}$$

[18] Here $\boldsymbol{y}_0\equiv\boldsymbol{x}_0$, $\boldsymbol{y}_{N+1}\equiv\boldsymbol{x}_1$ and $t_n=t_0+n\tau$.

Proceeding in the one-dimensional case from the assumption that

$$L(x,\dot{x}) = \tfrac{1}{2}m\dot{x}^2 - V(x)$$

Feynman was able to establish in fairly short order (which is to say: by heavy calculation completed within hours of his first reading of Dirac's paper) that if (24) is used to construct the propagator $K(x,t;x_0,t_0)$ that appears in

$$\psi(x,t) = \int K(x,t;x_0,t_0)\psi(x_0,t_0)\,dx_0$$

then the resulting $\psi(x,t)$ *does in fact satisfy the Schrödinger equation*, and does in fact give back $\psi(x_0,t_0)$ as $t\downarrow t_0$.[19]

> **EXAMPLE**: For a $\boxed{\text{FREE PARTICLE}}$ it was established already at (19) that
>
> $$S(x_1,t_0+\tau;x_0,t_0) = \frac{m(x_1-x_0)^2}{2\tau}$$
>
> so in this instance (23) reads
>
> $$\begin{aligned} K(x_1,t_0+\tau;x_0,t_0) &= \frac{1}{A(\tau)}\,e^{im(x_1-x_0)^2/2\hbar\tau} \\ &= \frac{1}{A(\tau)}\sqrt{\pi\sigma}\cdot\frac{1}{\sqrt{\pi\sigma}}\,e^{-(x_1-x_0)^2/\sigma} \\ &\qquad\sigma \equiv 2i\hbar\tau/m \end{aligned}$$
>
> With Feynman we temporarily complexify $\hbar$
>
> $$\hbar \longmapsto \hbar - i\epsilon \quad : \quad \epsilon > 0$$
>
> so as to render the real part of σ positive. This done, we have
>
> $$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{\pi\sigma}}\,e^{-(x_1-x_0)^2/\sigma}\,dx_1 = 1$$
>
> and recognize the integrand to provide (in the limit $\sigma\downarrow 0$) a Gaussian representation of $\delta(x_1-x_0)$. We therefore have, as was required at (18.2),
>
> $$\lim_{\tau\downarrow 0} K(x_1,t_0+\tau;x_0,t_0) = \delta(x_1-x_0) \tag{25}$$
>
> provided we set
>
> $$A(\tau) = \sqrt{\pi\sigma} = \sqrt{ih\tau/m}$$

---

[19] For proof see pages 29–30 in Chapter 3 of ADVANCED QUANTUM TOPICS (2000).

Observe next that

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{\pi\sigma}}\, e^{-(x_1 - y)^2/\sigma} \frac{1}{\sqrt{\pi\sigma}}\, e^{-(y - x_0)^2/\sigma}\, dy = \frac{1}{\sqrt{\pi 2\sigma}}\, e^{-(x_1 - x_0)^2/2\sigma}$$

and that N-fold iteration of this result would lead Feynman to write

$$\begin{aligned} K(x_1,t_1;x_0,t_0) &= \lim_{N\uparrow\infty} \frac{1}{\sqrt{\pi(N+1)\sigma}}\, e^{-(x_1 - x_0)^2/(N+1)\sigma} \\ &= \frac{1}{\sqrt{ih(t_1 - t_0)/m}}\, e^{(i/\hbar)m(x_1 - x_0)^2/2(t_1 - t_0)} \end{aligned} \tag{26}$$

We verify by quick calculation that the expression on the right does in fact satisfy the free particle Schrödinger equations (18.1)

$$\begin{aligned} \big\{-\tfrac{\hbar^2}{2m}\partial_{x_1}^2 - i\hbar\partial_{t_1}\big\}K(x_1,t_1;x_0,t_0) &= 0 \\ \big\{-\tfrac{\hbar^2}{2m}\partial_{x_0}^2 + i\hbar\partial_{t_0}\big\}K(x_1,t_1;x_0,t_0) &= 0 \end{aligned}$$

EXAMPLE: For an $\boxed{\text{OSCILLATOR}}$—assuming classical motion to be uniform in the short term

$$x(t) = x_0 + \frac{x_1 - x_0}{\tau}\,(t - t_0) \quad : \quad t_0 \leqslant t \leqslant t_0 + \tau$$

—we compute

$$\begin{aligned} &S(x_1, t_0+\tau; x_0, t_0) \\ &\quad= \int_{t_0}^{t_0+\tau} \left\{\tfrac{1}{2}m\Big[\frac{x_1 - x_0}{\tau}\Big]^2 - \tfrac{1}{2}m\omega^2\Big[x_0 + \frac{x_1 - x_0}{\tau}\,(t - t_0)\Big]^2\right\} dt \\ &\quad= \frac{m(x_1 - x_0)^2}{2\tau} - \frac{m\omega^2\tau}{6}\big(x_1^2 + x_1x_0 + x_0^2\big) \end{aligned} \tag{27}$$

A moment's thought shows that to recover (25) we have again to set $A(\tau) = \sqrt{ih\tau/m}$, and so have

$$\begin{aligned} &K(x_1, t_0+\tau; x_0, t_0) \\ &\quad= \frac{1}{\sqrt{ih\tau/m}}\exp\left\{\frac{i}{\hbar}\Big[\frac{m(x_1 - x_0)^2}{2\tau} - \frac{m\omega^2\tau}{6}\big(x_1^2 + x_1x_0 + x_0^2\big)\Big]\right\} \end{aligned} \tag{28}$$

Introduction of this result into Feynman's formula (24) leads (after complexification of $\hbar$) to a multiple integral that can be managed effectively by appeal to the N-dimensional Gaussian integral formula

$$\int\cdots\int_{-\infty}^{+\infty} e^{i\boldsymbol{x}\cdot\boldsymbol{y}} e^{-\frac{1}{2}\boldsymbol{y}\cdot\mathbb{A}\boldsymbol{y}} dy_1 dy_2 \cdots dy_N = \frac{(2\pi)^{N/2}}{\sqrt{\det\mathbb{A}}}\, e^{-\frac{1}{2}\boldsymbol{x}\cdot\mathbb{A}^{-1}\boldsymbol{x}}$$

Taking the result to the limit $N \uparrow \infty$ requires a bit of finesse (for these and other computational details see pages 43–47 in Chapter 3 of my ADVANCED QUANTUM TOPICS (2000)), but leads ultimately to the formula

$$\begin{aligned}K(x_1,t_1;x_0,t_0)& \\ &= \sqrt{\tfrac{m\omega}{ih\sin\omega(t_1-t_0)}}\exp\Big\{\tfrac{i}{\hbar}\tfrac{m\omega}{2\sin\omega(t_1-t_0)}\big[(x_1^2+x_0^2)\cos\omega(t_1-t_0)-2x_1x_0\big]\Big\}\end{aligned} \tag{29}$$

from which, it is gratifying to observe, we recover (26) as $\omega \downarrow 0$. Computation establishes that the Feynman's oscillator propagator (29) satisfies

$$\begin{aligned}\big\{-\tfrac{\hbar^2}{2m}\partial_{x_1}^2+\tfrac{1}{2}m\omega^2x_1^2-i\hbar\partial_{t_1}\big\}K(x_1,t_1;x_0,t_0)&=0\\ \big\{-\tfrac{\hbar^2}{2m}\partial_{x_0}^2+\tfrac{1}{2}m\omega^2x_0^2+i\hbar\partial_{t_0}\big\}K(x_1,t_1;x_0,t_0)&=0\end{aligned}$$

Feynman's central idea, as set forth in his thesis (May 1942) and in the paper which he published (Reviews of Modern Physics, 1948),[20] admits of the picturesque if somewhat imprecise formulation

$$K(\boldsymbol{x},t;\boldsymbol{x}_0,t_0)=\int_{\text{all paths}} e^{\frac{i}{\hbar}\int L(\text{path})\,dt}\,\mathcal{D}[\text{paths}] \tag{30}$$

which he interprets to describe a *sum over the probability amplitudes* that he would associate with each conceivable path $(\boldsymbol{x},t) \longleftarrow (\boldsymbol{x}_0,t_0)$, his contention being that each path contributes with the same weight, but with a phase determined by the classical action $S[\text{path}] = \frac{i}{\hbar}\int L(\text{path})\,dt$. He shows how (30) can be used to provide novel insight into the construction of

- quantum mechanical conservation laws
- commutation relations
- expectation values.

Subsequent work by others established that (30) can be made to retain its accuracy even when the meaning of "all paths" is altered in various ways—the important implication being that Feynman's success cannot be interpreted

[20] The thesis is entitled "The principle of least action in quantum mechanics," and in his abstract Feynman claims to have produced a *generalization* of the standard Heisenberg/Schrödinger/Born theory. In the published version (which is reprinted in the Schwinger collection) the title has been changed (to "Space-time approach to non-relativistic quantum mechanics"—the "space-time" evidently intended to draw attention to the fact that Feynman works throughout in the $\boldsymbol{x}$-representation, and the "non-relativistic" to disabuse readers of the presumption that "space-time" means "spacetime") and that claim is dropped: Feynman claims "no fundamentally new results," aspires only to the "pleasure of recognizing old things from a new point of view."

as having established that quantum particles *do* move $\boldsymbol{x} \longleftarrow \boldsymbol{x}_0$ along diverse nowhere-differentiable paths: it establishes only that they *can be imagined* to do so.

There remains—today as in 1942—no properly constructed general "theory of functional integration" adequate to the work Feynman would have it do. Successful functional integrals appear in every instance to be *Gaussian* integrals, and to derive that characteristic from the fact that $\dot{\boldsymbol{x}}$ enters *quadratically* into the construction of physical Lagrangians. Feynman seems never to have been much bothered by the fact that his physical intuition had carried him to a place beyond the frontier of established mathematics, but...

Near the end of the conclusion to his thesis Feynman expresses his regret that

> *"A point of vagueness is the normalizing factor, A. No rule has been given to determine it for a given action expression. This question is related to the difficult mathematical question as to the conditions under which the limiting process of subdividing the time scale... actually converges."*

"Hamilton's principle," it has often been remarked—in the first instance by Hamilton himself!—is due actually to Lagrange, whose work Hamilton held to be a kind of "mathematical poem." But it was Hamilton who first recognized that $\delta\!\int\! L\,dt = 0$ is but the point of entry into a rich landscape[21] (theory of canonical transformations, Hamilton-Jacobi theory) that emerges as soon as one looks to the *properties* of the functions $S(\boldsymbol{x}_1, t_1; \boldsymbol{x}_0, t_0)$ that extremize the functional $S[\text{path}] = \int L\,dt$. Feynman's objective was, in a way, more restricted than Dirac's: he sought only to discover a natural place within quantum theory for the "principle of least action." Had he not been so quick to dismiss those aspects of Dirac's paper that drew upon the deeper reaches of Hamiltonian physics he might well have resolved the problem identified in the preceding quotation—as, indeed, did others (Pauli 1951, Cécile Morette 1952) soon after him. I turn now to discussion of the relevant details.[22]

Working initially in one dimension, we proceed from

$$\mathsf{U}(t+\tau,t) = \mathsf{I} - \frac{i}{\hbar}\mathsf{H}\tau + \cdots \approx e^{-i\mathsf{H}\tau/\hbar}$$

where H will be considered to have been obtained by xp-ordered substitution into the classical Hamiltonian $H(x,p)$:

$$\mathsf{H} = {}_{\mathsf{x}}\Big[H(x,p)\Big]_{\mathsf{p}} + O(\hbar)$$

[21] Cornelius Lanczos, quoting from the Book of Exodus at the beginning of Chapter 8 in his *Variational Principles of Mechanics* (1949), refers to this landscape as "holy ground." Lanczos provides, by the way, an excellent account of Hamilton-Jacobi theory that is directly relevant to many of the topics discussed in the present essay.

[22] The reader should be aware that my methods, taken from a notebook dated 1971, are somewhat eccentric.

In this notation

$$K(x,t+\tau;x_0,t) = \int (x|\mathbf{U}(t+\tau,t)|p)dp(p|x_0)$$
$$\approx \int \tfrac{1}{h}\, e^{\frac{i}{\hbar}[(x-x_0)p-\tau H(x,p)]}\, dp$$

from which it follows in particular (and not at all surprisingly) that

$$\lim_{\tau\downarrow 0} K(x,t+\tau;x_0,t) = \int \tfrac{1}{h}\, e^{\frac{i}{\hbar}(x-x_0)p}\, dp = \delta(x-x_0) \quad : \quad \text{all } H(x,p)$$

Notice now that $[(x-x_0)p-\tau H(x,p)] = \tau[\frac{x-x_0}{\tau}\,p - H(x,p)]$ is, by Hamilton's principle,[23] stationary at the momentum p_C associated with the classical path $(x,t+\tau) \leftarrow (x_0,t)$, where it assumes the value $S(x,t+\tau;x_0,t)$. With these points in mind, we use the method of stationary phase[24] to obtain

$$\lim_{\hbar\downarrow 0} \int \tfrac{1}{h}\, e^{\frac{i}{\hbar}[(x-x_0)p-\tau H(x,p)]}\, dp \sim \tfrac{1}{h}\Big[\frac{i2\pi\hbar}{-\tau H_{pp}(x,p)}\Big]^{\frac{1}{2}}\, e^{iS(x,t+\tau;x_0,t)/\hbar}\Big|_{p=p_C}$$

Hamilton's canonical equations of motion supply

$$x_0 = x - \tau H_p(x,p)\Big|_{p=p_C} \quad \text{giving} \quad \frac{\partial x_0}{\partial p} = -\tau H_{pp}(x,p)\Big|_{p=p_C}$$

whence

$$\frac{1}{-\tau H_{pp}(x,p)} = \frac{\partial p}{\partial x_0} = \frac{\partial^2 S(x,t+\tau;x_0,t)}{\partial x \partial x_0}$$

Assembling these results, we have

$$K(x,t+\tau;x_0,t) = \sqrt{(i/h)D(x,t+\tau;x_0,t)}\cdot e^{iS(x,t+\tau;x_0,t)/\hbar}$$
$$D(x,t+\tau;x_0,t) \equiv \frac{\partial^2 S(x,t+\tau;x_0,t)}{\partial x \partial x_0}$$

which in several dimensions becomes

$$K(\boldsymbol{x},t+\tau;\boldsymbol{x}_0,t) = \sqrt{(i/h)^n D(\boldsymbol{x},t+\tau;\boldsymbol{x}_0,t)}\cdot e^{iS(\boldsymbol{x},t+\tau;\boldsymbol{x}_0,t)/\hbar} \tag{31}$$
$$D(\boldsymbol{x},t+\tau;\boldsymbol{x}_0,t) \equiv \det\left\|\frac{\partial^2 S(\boldsymbol{x},t+\tau;\boldsymbol{x}_0,t)}{\partial x^j \partial x_0^k}\right\|$$

EXAMPLE: For a FREE PARTICLE we on page 19 had

$$S(x_1,t_0+\tau;x_0,t_0) = \frac{m(x_1-x_0)^2}{2\tau}$$

[23] Recall that $L(x,\dot{x}) = \dot{x}p - H(x,p))$

[24] See ADVANCED QUANTUM TOPICS (2000), Chapter 0, page 47.

which is seen to give $\sqrt{(i/h)D(x,t+\tau;x_0,t)} = \sqrt{-im/h\tau}$, in precise agreement with Feynman's $1/A(\tau) = \sqrt{m/ih\tau}$.

EXAMPLE: For an OSCILLATOR we on page 20 had

$$S(x_1,t_0+\tau;x_0,t_0) = \frac{m(x_1-x_0)^2}{2\tau} - \frac{m\omega^2\tau}{6}\big(x_1^2+x_1x_0+x_0^2\big)$$

which gives $\sqrt{(i/h)D(x,t+\tau;x_0,t)} = \sqrt{-im/h\tau[1+\frac{1}{6}(\omega\tau)^2]}$. This (in leading order) again reproduces Feynman's normalization factor.

It is curious that Feynman himself seems to have been quite uninterested in this work, though it resolves what he recognized to be a defect in his own work. It is neither cited nor exploited in Feynman & Hibbs, *Quantum Mechanics and Path Integrals* (1965).

The material sketched above refers to a property of the *short-time propagator* $K(x_1,t_0+\tau;x_0,t_0)$. Interestingly, it can be considered to be a special consequence of a property of *general propagators in the classical limit*. If we write

$$K(\boldsymbol{x}_1,t_1;\boldsymbol{x}_0,t_0) = R(\boldsymbol{x}_1,t_1;\boldsymbol{x}_0,t_0)\,e^{iS(\boldsymbol{x}_1,t_1;\boldsymbol{x}_0,t_0)/\hbar}$$

then the argument that on page 10 gave (14) gives

$$\begin{aligned}
\tfrac{1}{2m}\nabla_1 S\cdot\nabla_1 S + U(\boldsymbol{x}_1) + O(\hbar^2) + S_{t_1} &= 0 && (32.1)_1\\
\tfrac{1}{2m}\nabla_0 S\cdot\nabla_0 S + U(\boldsymbol{x}_0) + O(\hbar^2) - S_{t_0} &= 0 && (32.1)_0\\
\nabla_1\cdot\big[\big(\tfrac{1}{m}\nabla_1 S\big)R^2\big] + \partial_{t_1}R^2 &= 0 && (32.2)_1\\
\nabla_0\cdot\big[\big(\tfrac{1}{m}\nabla_0 S\big)R^2\big] - \partial_{t_0}R^2 &= 0 && (32.2)_0
\end{aligned}$$

One can show[25] that if the two-point function $S(\boldsymbol{x}_1,t_1;\boldsymbol{x}_0,t_0)$ satisfies the Hamilton-Jacobi equations (32.1)—with $O(\hbar^2)$ terms omitted—then the **Van Vleck determinant**[26]

$$D(\boldsymbol{x}_1,t_1;\boldsymbol{x}_0,t_0) \equiv \det\left\|\frac{\partial^2 S(\boldsymbol{x}_1,t_1;\boldsymbol{x}_0,t_0)}{\partial x_1^j\,\partial x_0^k}\right\|$$

[25] See ADVANCED QUANTUM TOPICS (2000), Chapter 3, page 15. The argument does not presume the Hamiltonian to have the specialized form $H(\boldsymbol{x},\boldsymbol{p}) = \frac{1}{2m}\boldsymbol{p}\cdot\boldsymbol{p} + U(\boldsymbol{x})$ implicit in the way I have rendered equations (32).

[26] John Van Vleck (1899–1980) came upon this determinant in the course of an early study of several-dimensional WKB theory (1928). In conversation with me he once remarked that the idea sprang from a conversation with J. R. Oppenheimer, and that his own contribution had been simply to work out the details. Van Vleck remarked with evident pride that the quantum mechanical dissertation (1922) he wrote at Harvard (where he was the first student of E. C. Kemble) was an American first. He shared the Nobel Prize in 1977.

satisfies the continuity equations (32.2). This, I emphasize, is an entirely classical proposition, its quantum mechanical importance notwithstanding.

We are brought thus to the conclusion that the quantum propagator $K(\boldsymbol{x}_1, t_1; \boldsymbol{x}_0, t_0)$ can in semi-classical approximation be described

$$K_C(\boldsymbol{x}_1, t_1; \boldsymbol{x}_0, t_0) = \sqrt{(i/h)^n D(\boldsymbol{x}_1, t_1; \boldsymbol{x}_0, t_0)} \cdot e^{iS(\boldsymbol{x}_1, t_1; \boldsymbol{x}_0, t_0)/\hbar} \tag{33}$$

This is a fact at which Dirac strongly hinted, and of which Feynman made implicit use, but which was first clearly stated by Pauli.[27] From (33) one recovers (31) in the short-time limit. Notice that $K_C(\boldsymbol{x}_1, t_1; \boldsymbol{x}_0, t_0)$ is an entirely *classical* construction, if one which no classical consideration would motivate us to construct. And that it is by now clear that at the heart of the Feynman formalism lies the fact that

> **Quantum mechanics is briefly classical**: the limiting processes $\tau \downarrow 0$ and $\hbar \downarrow 0$ are equivalent.

EXAMPLE: Classical calculations establish that for an OSCILLATOR one has

$$\left.\begin{aligned} S(x_1, t_1; x_0, t_0) &= \frac{m\omega\big[(x_1^2 + x_0^2)\cos\omega(t_1 - t_0) - 2x_1x_0\big]}{2\sin\omega(t_1 - t_0)} \\ &\Downarrow \\ D(x_1, t_1; x_0, t_0) &= -\frac{m\omega}{\sin\omega(t_1 - t_0)} \end{aligned}\right\} \tag{34}$$

which when introduced into (33) give back (29). It is a remarkable fact that in this case—as for all Hamiltonians that depend at most quadratically upon $\boldsymbol{x}$ and $\boldsymbol{p}$—$K_C(\boldsymbol{x}_1, t_1; \boldsymbol{x}_0, t_0) = K(\boldsymbol{x}_1, t_1; \boldsymbol{x}_0, t_0)$: the semi-classical approximation is exact. Quick calculation confirms that the functions (34) do indeed satisfy the continuity equations (32.2) if one interprets R^2 to mean D. Note finally that if in (34) we set $t_1 = t_0 + \tau$ and expand in powers of τ we recover (27) and obtain

$$D(x_1, t_0 + \tau; x_0, t_0) = -\frac{m}{\tau}\big[1 + \tfrac{1}{6}(\omega\tau)^2 + \tfrac{7}{360}(\omega\tau)^4 - \cdots\big]$$

which in the limit $\tau \downarrow 0$ gives $\sqrt{(i/h)D} = \sqrt{m/ih\tau}$, in precise agreement with Feynman's $1/A(\tau)$: see again (28) on page 20.

---

[27] See Chapter 6 in Pauli's *Selected Topics in Field Quantization*, which is the English translation (1973) some ETH lecture notes dated 1950/51, to which this material is attached as an appendix.

7. Equivalence theorems. Standard quantum mechanics provides a "spectral representation of the propagator"

$$K(\boldsymbol{x}_1,t_1;\boldsymbol{x}_0,t_0)=\sum_n \psi_n(\boldsymbol{x}_1)\psi^*(\boldsymbol{x}_0)\,e^{-\frac{i}{\hbar}E_n(t_1-t_0)}$$

which bears little resemblance to Feynman's

$$=\int_{\text{all paths}} e^{\frac{i}{\hbar}\int L(\text{path})\,dt}\,\mathcal{D}[\text{paths}]$$

The standard construction is "wave theoretic," while Feynman's is "particle theoretic." At issue is the question of how they come to say the same thing, which I will discuss in the contexts provided by specific examples:

EXAMPLE: For a FREE PARTICLE we have on the one hand

$$K(x_1,t_1;x_0,t_0)=\int \tfrac{1}{\sqrt{h}}e^{\frac{i}{\hbar}px_1}\tfrac{1}{\sqrt{h}}e^{-\frac{i}{\hbar}px_0}e^{-\frac{i}{\hbar}\frac{1}{2m}p^2(t_1-t_0)}\,dp$$

while Feynman's procedure was seen at (26) to supply

$$=\sqrt{\frac{m}{ih(t_1-t_0)}}\,\exp\left\{\frac{i}{\hbar}\frac{m}{2}\frac{(x_1-x_0)^2}{t_1\,-\,t_0}\right\}$$

The formal equivalence of those results is a familiar fact of Fourier transform theory:

$$\int_{-\infty}^{+\infty} e^{ikx}e^{-\frac{1}{2}k^2t}\,dt=\sqrt{2\pi/t}\cdot e^{-\frac{1}{2}x^2/t}$$

Note that t appears "upstairs" on the left, but "downstairs" on the right.

EXAMPLE: For a PARTICLE CONSTRAINED TO MOVE FREELY ON A RING of circumference a the spectral representation of the propagator becomes (here $p_n=nh/a$ and $\mathcal{E}=h^2/2ma^2$)

$$\begin{aligned}K(x,t;y,0)&=\sum_{n=-\infty}^{\infty}\tfrac{1}{a}e^{\frac{i}{\hbar}p_n(x-y)}e^{-\frac{i}{\hbar}\mathcal{E}n^2t}\\&=\tfrac{1}{a}\Big\{1+2\sum_{n=1}^{\infty}e^{-\frac{i}{\hbar}\mathcal{E}n^2t}\cos\big[2n\pi\frac{x-y}{a}\big]\Big\}\end{aligned}$$

As it happens, a name and elegant theory attaches to series of that design: the theta function $\vartheta_3(z,\tau)$—an invention of the youthful Jacobi—is defined

$$\begin{aligned}\vartheta_3(z,\tau)&\equiv 1+2\sum_{n=1}^{\infty}q^{n^2}\cos 2nz\quad\text{with}\quad q\equiv e^{i\pi\tau}\\&=\sum_{n=-\infty}^{\infty}e^{i(\pi\tau n^2-2nz)}\end{aligned}$$

In this notation $K(x,t;y,0)=\frac{1}{a}\vartheta(z,\tau)$ with $z=\pi\frac{x-y}{a}$ and $\tau=-\frac{\mathcal{E}t}{\pi\hbar}=-\frac{2\pi\hbar t}{ma^2}$. The Feynman formalism leads, on the other hand, to

$$
\begin{aligned}
K(x,t;y,0) &= \sqrt{\tfrac{m}{iht}}\sum_{n=-\infty}^{\infty}\exp\left\{\tfrac{im}{\hbar 2t}(x+na-y)^2\right\}\\
&= \sqrt{\tfrac{m}{iht}}\exp\left\{\tfrac{im}{\hbar 2t}(x-y)^2\right\}\cdot\sum_{n=-\infty}^{\infty}e^{i(-\pi n^2-2nz)/\tau}\\
&= \tfrac{1}{a}\sqrt{\tau/i}\;e^{z^2/\pi i\tau}\cdot\vartheta(\tfrac{z}{\tau},-\tfrac{1}{\tau})
\end{aligned}
$$

The theory of theta functions supplies, however, a zillion wonderful identities, of which the most celebrated is the "Jacobi theta transformation" ("Jacobi's identity")[28]

$$
\vartheta(z,\tau) = \sqrt{\tau/i}\;e^{z^2/\pi i\tau}\cdot\vartheta(\tfrac{z}{\tau},-\tfrac{1}{\tau})
$$

which is exactly what we need to establish the equivalence in this instance of the spectral propagator (wherein t lives upstairs) and the Feynman propagator (wherein t lives downstairs). Jacobi's identity also mediates between the propagators obtained in all soluable "particle in a box" problems, in one or several dimensions.

EXAMPLE: For an OSCILLATOR we have the spectral representation

$$
K(x,t;y,0) = \sum_{n=0}^{\infty}\psi_n(x)\psi_n^*(y)\,e^{-i\omega(n+\frac{1}{2})t} \tag{34.1}
$$

$$
\psi_n(x) = \left(\tfrac{2m\omega}{h}\right)^{\frac{1}{4}}\tfrac{1}{\sqrt{2^n n!}}e^{-\frac{1}{2}(m\omega/\hbar)x^2}H_n\left(\sqrt{\tfrac{m\omega}{\hbar}}x\right)
$$

while Feynman's method was seen at (29) to supply

$$
= \sqrt{\tfrac{m\omega}{ih\sin\omega t}}\exp\left\{\tfrac{i}{\hbar}\tfrac{m\omega}{2\sin\omega t}\left[(x^2+y^2)\cos\omega t-2xy\right]\right\} \tag{34.2}
$$

Equivalence follows now from "Mehler's formula"[29]

$$
\sum_{n=0}^{\infty}\tfrac{1}{n!}\left(\tfrac{1}{2}\tau\right)^n H_n(x)H_n(y) = \frac{1}{\sqrt{1-\tau^2}}\exp\left\{\frac{2xy\tau-(x^2+y^2)\tau^2}{1-\tau^2}\right\}
$$

---

[28] Richard Bellman—in *A Brief Introduction to Theta Functions* (1961)—has remarked that "[Jacobi's identity] has amazing ramifications in the fields of algebra, number theory, geometry and other parts of mathematics [to which I must add also physics!]. In fact, it is not easy to find another identity of comparable significance."

[29] F. G. Mehler (1866). See A. Erdélyi, *Higher Transcendental Functions, Volume II* (1953), page 194. Full details can be found in "Jacobi's theta transformation & Mehler's formula: their interrelation, and their role in the quantum theory of angular momentum" (2000.)

I know of a few additional cases in which both the spectral representation and Feynman's description of the propagator can be evaluated in exact closed form, and equivalence established by appeal to a "well-known" identity. In each such case, t resides upstairs on one side, and downstairs on the other side of the equality. If the Feynman formulation of quantum mechanics is truly equivalent to the standard formulation, then there must exist *as many such identities as there are quantum systems*! And they must, as a population, provide the subject matter of some kind of over-arching general theory. Which has, so far as I am aware, yet to be devised.

One final comment: It is commonly held to follow from the spectral representation that

$$\lim_{t_1 \downarrow t_0} K(x_1, t_1; x_0, t_0) = \sum_n \psi_n(x_1)\psi_n^*(x_0) = \delta(x_1 - x_0)$$

as was required at (18.2). But the latter equality follows from the *presumption* that the set $\{\psi_n(x)\}$ of eigenfunctions is *complete*. But equivalence theorems of the sort we have been considering place us in position to write

$$\sum_n \psi_n(x_1)\psi_n^*(x_0) = \lim_{t_1 \downarrow t_0} \{\text{Feynman propagator}\}$$

and thus to *prove* completeness. This is, in fact, the essential pattern of all completeness proofs known to me.

8. Conclusion: Hamilton's legacy. What began as a modest attempt to sketch "Schrödinger's train of thought" has expanded to become a more ambitious account of some of the *several* ways in which Hamiltonian mechanics, in *all* of its parts, became foundational to quantum mechanics. Hamilton had by 1834—when he was not yet thirty years old!—brought classical mechanics to its highest state of development. But the mountain top on which he stood was wrapped in a dense fog that forever denied him a view of the glorious landscape that lay just ahead. Dispersal of the fog had to await completion of the work of Maxwell (electromagnetic waves, kinetic beginnings of statistical mechanics) and of those who brought thermodynamics and statistical mechanics to mature perfection, had to await Planck's thermodynamics of light and the wide-ranging insight of Einstein, who built upon those developments. Only then did Bohr, de Broglie, Schrödinger... Feynman become possible, and the quantum world latent in Hamilton's classical work come finally into plain view.

It is interesting to note that some of the developments just mentioned made critical use of ideas original to Hamilton. Statistical mechanics, as formulated by Gibbs, is an exercise in Hamiltonian mechanics. Planck's primitive quantization procedure was recognized by Bohr/Sommerfeld to involve the quantization of areas $\oint p\,dq$ inscribed on Hamiltonian phase space, and when Ehrenfest enlarged upon that insight to propose the quantization of classical "adiabatic invariants" it was pointed out by K. Schwarzschild that those issued most naturally from an elaboration of Hamilton-Jacobi theory that had been devised by C. Delaunay as an aid to his study of the motion of the moon (1846).

Between 1843 and the year of his death (1865) Hamilton gave his attention mainly to the development and promotion of the theory of quaternions. While that effort seems in retrospect to have been in many respects misguided, it did mark the introduction of the *noncommutivity* concept which was destined to become one of the most distinguishing features of quantum theory. And it provided a kind of preview of a set of mathematical relationships that—eighty years later—became fundamental to the quantum theory of spin.