# 3

# FEYNMAN QUANTIZATION

*An introduction to path-integral techniques*

**Introduction.** By 1941 Richard Feynman (1918–88), who—after a distinguished undergraduate career at MIT—had come in 1939 as a graduate student to Princeton, was deeply involved in a collaborative effort with John Wheeler (his thesis advisor) to shake the foundations of field theory. Though motivated by problems fundamental to quantum field theory, as it was then conceived, their work was entirely classical,[1] and it advanced ideas so radical as to resist all then-existing quantization techniques:[2] new insight into the quantization process itself appeared to be called for.

So it was that (at a beer party) Feynman asked Herbert Jehle (formerly a student of Schrödinger in Berlin, now a visitor at Princeton) whether he had ever encountered a quantum mechanical application of the "Principle of Least Action." Jehle directed Feynman's attention to an obscure paper by P. A. M. Dirac[3] and to a brief passage in §32 of Dirac's *Principles of Quantum Mechanics*

[1] John Archibald Wheeler & Richard Phillips Feynman, "Interaction with the absorber as the mechanism of radiation," Reviews of Modern Physics **17**, 157 (1945); "Classical electrodynamics in terms of direct interparticle action," Reviews of Modern Physics **21**, 425 (1949). Those were (respectively) Part III and Part II of a projected series of papers, the other parts of which were never published.

[2] See page 128 in J. Gleick, *Genius: The Life & Science of Richard Feynman* (1992) for a popular account of the historical circumstances.

[3] "The Lagrangian in quantum mechanics," Physicalische Zeitschrift der Sowjetunion **3**, 64 (1933). The paper is reprinted in J. Schwinger, *Selected Papers on Quantum Electrodynamics* (1958). I refer to this henceforth as the "Schwinger Collection."

($2^{\text{nd}}$ edition 1935). Thus did it come about that in May 1942 Feynman defended a dissertation entitled "The principle of least action in quantum mechanics."[4] Publication of the work (as of the field theoretic work that had stimulated its creation) was delayed until Feynman (and Wheeler) had returned to academic life from their wartime participation in the Manhatten Project, and by the time it appeared in the pages of Wheeler/Feynman's favorite journal[5] the title had changed—to "Space-time approach to non-relativistic quantum mechanics"—and the allusions to its original field theoretic reason-for-being had been largely discarded. Feynman (who was then at Cornell) begins his classic paper with these words:

> *It is a curious historical fact that modern quantum mechanics began with two quite different mathematical formulations: the differential equation of Schrödinger, and the matrix algebra of Heisenberg. The two apparently dissimilar approaches were proved to be mathematically equivalent. These two points of view were destined to complement one another and to be ultimately synthesized in Dirac's transformation theory.*
>
> *This paper will describe what is essentially a third formulation of non-relativistic quantum theory. This formulation was suggested by some of Dirac's remarks concerning the relation of classical action to quantum mechanics. A probability amplitude is associated with an entire motion of a particle as a function of time, rather than simply with a position of the particle at a particular time.*
>
> *The formulation is mathematically equivalent to the more usual formulations. There are, therefore, no fundamentally new results. However, there is a pleasure in recognizing old things from a new point of view. Also, there are problems for which the new point of view offers a distinct advantage ...*

Though Pauli lectured luminously on the germ of Feynman's idea (to his students at the ETH in Zürich) already in 1950/51,[6] and Cécile Morette, at about that same time (she was then at the Institute for Advanced Study, and in working contact with both von Neumann and Oppenheimer), attempted to clarify some of the mathematical details (and to extend the range) of a

---

[4] Feynman says "principle of least action" but means *Hamilton's principle*. In classical mechanics the former terminology refers to something quite else: see H. Goldstein, *Classical Mechanics* ($2^{\text{nd}}$ edition 1980) §8–6 or my "Geometrical mechanics: Remarks commemorative of Heinrich Hertz" (1994).

[5] Reviews of Modern Physics **20**, 267 (1948). The paper is reprinted in the Schwinger Collection.

[6] Pauli's lecture notes (in German) circulated widely. They were made available in English translation as *Pauli Lectures on Physics: Volumes 1–6* in 1973. Pauli's remarks concerning the Feynman formalism were presented as an appendix ("Feynman's approach to quantum electrodynamics: the path integral method") to Volume 6: *Selected Topics in Field Quantization*.

formalism that Feynman himself had been content merely to sketch[7]... most physicists were content at first to look upon Feynman's accomplishment in the terms he had presented it: as an amusing restatement of "old things from a new point of view." Some important contributions were made by a few people during the early 1950's,[8] but Feynman's great quantum electrodynamical papers did not appear until 1948–50 and it appears to have been mainly in delayed response to those,[9] and to Feynman's participation in several important workshops and conferences,[10] that the path-integral method entered the mainstream of physics. During the 1960's the previous trickle of papers on the subject became a flood: the method was explored from every angle, applied to every conceivable quantum mechanical problem, appropriated by other branches of physics (statistical mechanics, most notably[11]). In subsequent decades the method became basic first to gauge field theory, and more recently to string theory.

---

[7] "On the definition and approximation of Feynman's path integrals," Phys. Rev. **81**, 848 (1951).

[8] I am thinking especially of Ph. Choquard, "Traitement semi-classique des forces générales dans la représentation de Feynman," Helvetica Physica Acta **28**, 89 (1955); H. J. Groenewold, "Quasi-classical path integrals," Mat. Pys. Medd. Dan. Vid. Selsk. **30**, No.19 (1956); C. W. Kilmister, "A note on summation over Feynman histories," Proc. Camb. Phil. Soc. **54**, 302 (1957). Choquard, by the way, was a student of Pauli, and Groenewold a leading authority on and proponent of the Weyl transform and Wigner/Moyal phase space formalism.

[9] The paper in which "Feynman diagrams" make their first appearance technique bears a title—"Space-time approach to quantum electrodynamics" (Phys. Rev. **76**, 769 (1949))—that alludes explicitly to the path integral paper. In his introductory remarks Feynman reports that "the Lagrangian form of quantum mechanics described in the Reviews of Modern Physics" marks "the genesis of this paper," but that he has attempted to proceed "without relying on the Lagrangian method, because it is not generally familiar." In a footnote Feynman mentions that (to his chagrin?) application of the sum-over-paths technique "to electrodynamics [has been] described in detail [already] by [that same] H. J. Groenewold" in a publication that had appeared a few months earlier.

[10] The famous "Chapel Hill converence"—the proceedings of which were published as *Conference on the role of gravitation in physics* (1957), edited by Cécile M. De Witt (formerly Morette, but now the wife of the mathematical physicist Bryce De Witt)—marked the beginning of the modern era for general relativity, and (though he did not claim expertise in the field) was dominated by the personality of Feynman. There was by then an emerging consensus that "Feynman quantization" was the method of choice for quantizing such otherwise intractable systems as the gravitational field.

[11] See, for example, David Falkoff, "Statistical theory of irreversible processes. Part I. Integral over fluctuation path formulation," Annals of Physics **4**, 325 (1958), which is representative of a vast literature, and is cited here because its author was my friend.

The subject has become so broad that even the authors of books—of which there are by now quite a number[12]—are forced to restrict their focus, to treat only *aspects* of the topic.

It is perhaps not surprising that, once Feynman's home-grown way of thinking about quantum mechanics had lost its radical novelty, occasional instructors (among them me, in 1967–1970) would test the feasibility of using the path integral method to *teach* quantum mechanics . . . and it is certainly not surprising that such experiments should be undertaken at Caltech (to which Feynman had gone from Cornell in 1951). A textbook—*Quantum Mechanics and Path Integrals* by Feynman and Albert Hibbs (one of Feynman's thesis students)—provides an expanded record of that experiment, which Feynman—no slave even to his own ideas—untimately abandoned. On evidence of the text I infer that Feynman himself had paid little or no attention to the work of Pauli, Morette, Groenewold and a growing number of others: by the mid-1960's he seems to have been no longer the best authority on the formalism that he himself had invented. It is a measure of the man (or is it an indicator simply of his assessment of the preparation/interests of his students?) that he seems to have been more interested in the diverse applications than in the theoretical refinement of his seminal idea.

In these few pages I can attempt to review only the bed-rock essentials of the path integral method, but warn the reader that my emphasis and mode of proceeding will at many points be ideosyncratic.

**Point of departure.** Quantum dynamics can (in the Schrödinger picture) be considered to reside in the statement

$$i\hbar\tfrac{\partial}{\partial t}|\psi)_t = \mathbf{H}|\psi)_t \quad : \quad |\psi)_0 \text{ known from initial measurement}$$

If $\mathbf{H}$ is time-independent then we have the integrated statement

$$|\psi)_t = \mathbf{U}(t,0)|\psi)_0 \quad : \quad t \geqslant 0 \tag{1.1}$$

where

$$\mathbf{U}(t,0) \text{ is } \textit{unitary}, \text{ with } \mathbf{U}(0,0) = \mathbf{I} \tag{1.2}$$

and given by

$$\mathbf{U}(t,0) = e^{-\frac{i}{\hbar}\mathbf{H}t} \tag{2}$$

In more general (time-dependent) cases we *still* have (1), but lose (2). The side condition $t \geqslant 0$ is intended to emphasize that the theory is *predictive but not retrodictive*: it has nothing to say about what $|\psi)_t$ may have been doing prior

[12] Some titles pulled from my personal bookshelf: L. S. Schulman, *Techniques and Applications of Path Integration* (1981); R. J. Rivers, *Path Integral Methods in Quantum Field Theory* (1987); C. Grosche, *Path Integrals, Hyperbolic Spaces & Selberg Trace Formulae* (1996); T. Tashiwa, Y. Ohnuki & M. Suzuki, *Path Integral Methods* (1997). There are many others, not to mention the chapters devoted to the subject in books addressed primarily with other matters.

to the projective act that created $|\psi)_0$.

Relative to any given orthonormal basis $\{|a)\}$ the fundamental equation (1.1) acquires the representation[13]

$$(a|\psi)_t = \sum\!\!\!\!\!\!\int \underbrace{(a|\,\mathbf{U}(t,0)|b)}_{\text{“transition amplitude:” } |b)_0 \longrightarrow |a)_t} db(b|\psi)_0 \tag{3}$$

But (1) entails

$$\mathbf{U}(t,0) = \mathbf{U}(t,t_1)\,\mathbf{U}(t_1,0) \quad : \quad t \geqslant t_1 \geqslant 0 \tag{4.1}$$

which acquires the representation

$$(a|\,\mathbf{U}(t,0)|b) = \int (a|\,\mathbf{U}(t,t_1)|a_1)da_1(a_1|U(t_1,0)|b) \tag{4.2}$$

In ordinary probability theory we encounter situations in which it becomes natural to write

$$\begin{aligned} P_{a\leftarrow b} &\equiv \text{probability of going from } b \text{ to } a \\ P_{a\leftarrow c}\cdot P_{c\leftarrow b} &\equiv \text{probability of going from } b \text{ to } a \textit{ via } c \end{aligned}$$

and on the assumption that the various channels $a \leftarrow c \leftarrow b$ are *statistically independent* obtain

$$P_{a\leftarrow b} = \sum_c P_{a\leftarrow c}\cdot P_{c\leftarrow b} \tag{5}$$

The quantum mechanical construction (4.2) is of similar design, except that it involves *probability amplitudes rather than probabilities*; it asserts, moveover, that in quantum mechanical contexts (5) is, in general, *not* valid:

$$\begin{aligned} |(a|\,\mathbf{U}(t,0)|b)|^2 &= \Big|\int (a|\,\mathbf{U}(t,t_1)|a_1)da_1(a_1|U(t_1,0)|b)\Big|^2 \\ &\neq \int |(a|\,\mathbf{U}(t,t_1)|a_1)|^2 da_1|(a_1|U(t_1,0)|b)|^2 \end{aligned}$$

We have touched here on the subtle shift of emphasis that lies at the heart of Feynman's conception of quantum mechanics. Standardly, we are taught to assign probability amplitudes to the *states* of quantum mechanical systems: that's the kind of thing that $\psi(x)$ is. But Feynman is “process oriented,” in the sense that he would have us

> associate probability amplitudes with the alternative independent channels that *gave rise* to the state

and from those *deduce* the amplitudes of states:

$$\text{amplitude of state} = \sum_{\text{channels}} \text{amplitude of contributing channel} \tag{6}$$

[13] As a matter merely of notational convenience we will—having made our point—henceforth assume the basis elements to be continuously indexed, and write simply $\int$ in place of $\sum\!\!\!\!\!\!\int$.

At this point, Feynman—characteristically but less fundamentally—elects to work in the space representation,[14] so (3) becomes (in the one-dimensional case)

$$\psi(x,t) = \int K(x,t;y,0)\psi(y,0)\,dy \tag{8}$$

where

$$K(x,t;y,0) = (x|\mathbf{U}(t,0)|y) \tag{9}$$

is the familiar 2-point Green's function or "propagator." It is a solution of the Schrödinger equation

$$i\hbar\tfrac{\partial}{\partial t}K(x,t;\bullet,\bullet) = \mathbf{H}K(x,t;\bullet,\bullet) \tag{10.1}$$

—distinguished from other solutions by the circumstance that

$$\lim_{t\downarrow 0} K(x,t;y,0) = \delta(y-x) \tag{10.2}$$

If the Hamiltonian is time-independent then $\mathbf{H}\psi_n(x) = E_n\psi_n(x)$ leads to the familiar *spectral construction of the propagator*

$$K(x,t;y,0) = \sum_n e^{-\frac{i}{\hbar}E_n t}\psi_n(x)\psi_n^*(y) \tag{11}$$

which is readily seen to conform to (10). But Feynman, following in the footsteps of Dirac, elects to proceed otherwise:

Partition the time interval $[t,0]$ into $N+1$ sub-intervals of (let us assume) equal duration

$$\tau \equiv \frac{t}{N+1}$$

and, on the basis of (4.2), write

$$\begin{aligned} K(x,t;y,0) &= \int\cdots\iint K(x,t;x_N,t-\tau)dx_N\cdots\,dx_2K(x_2,t,x_1,2\tau)\,dx_1K(x_1,\tau;y,0)\\ &= \int\cdots\iint\prod_{k=0}^{N}K(x_{k+1},k\tau+\tau;x_k,k\tau)\,dx_1dx_2\cdots dx_N \end{aligned} \tag{12}$$

with $x_0 \equiv y$ and $x_{N+1} \equiv x$. For time-independent systems we have the simplification

$$K(x_{k+1},t+\tau;x_k,t) = K(x_{k+1},\tau;x_k,0) \quad : \quad \text{all } t$$

[14] Whence, ultimately, Feynman's title: "*Space-time* formulation of non-relativistic quantum mechanics," where the "non-relativistic" is intended to signal that "space-time" is not, in this instance, to be read as an allusion to special relativity: Feynman does not, at this stage, propose to address the problems characteristic of relativistic quantum mechanics (quantum field theory).

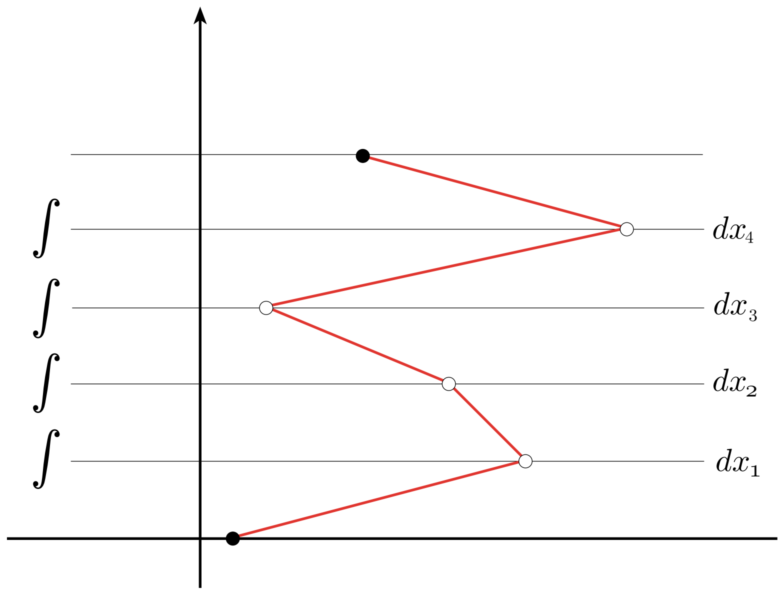



FIGURE 1: *The red sequence of transitions—called a "channel" in the generic language of (6)—is, because Feynman elects to work in the space representation, and the diagram is inscribed on spacetime, more naturally/vividly considered to describe a "path." Infinitely many such, according to (12), contribute independently/additatively to the transition* $(x, t) \longleftarrow (y, 0)$.

The representation (12) may look like a step in the wrong direction, since $N$-fold integration is typically difficult/impossible to do analytically, and can be a challenge even when approached numerically. The saving circumstance is that, as will be shown,

$$K(x, \tau; y, 0) \text{ simplifies as } \tau \text{ becomes small}$$

and the simplification is of such a nature as to lend a striking *interpretation* to the right side of (12).

Dirac had been brought (by discussion of the relationship between classical and quantum mechanical canonical transformations) to the conclusion[3,15] that

$$\text{“}K(x, t; y, 0) \text{ corresponds to } \exp\Big\{\tfrac{i}{\hbar}\int_0^t L\,dt\Big\}\text{”} \tag{13.1}$$

and that therefore

$$\text{“}K(x, \tau; y, 0) \text{ corresponds to } \exp\Big\{\tfrac{i}{\hbar}L\,\tau\Big\}\text{”} \tag{13.2}$$

[15] See also his "On the analogy between classical and quantum mechanics," Review of Modern Physics **17**, 195 (1945), which Feynman cites.

which "suggests that we ought to consider the classical Lagrangian not as a function of coordinates and velocities, but rather as a function of the coordinates at time $t$ and the coordinates at time $t + dt$."

When Feynman, in the presence of Jehle, first read Dirac's little paper (from which he quotes at length in this thesis) he was reportedly baffled by the occurance there of phrases like "corresponds to."[16] He was very soon convinced that "analogous to" could not mean "equals," but found that if he interpreted the phrase to mean "proportional to"—writing

$$K(x, t+\tau; y, t) = \underbrace{(\text{factor independent of } x \text{ and } y)}_{\text{function only of } \tau \text{ in time-independent cases: call it } 1/A(\tau)} \cdot e^{\frac{i}{\hbar}S(x,t+\tau;y,t)} \tag{14.1}$$

with

$$\begin{aligned} S(x, t+\tau; y, t) &= \int_t^{t+dt} L\big(\dot{x}(t'), x(t')\big)\, dt' \\ &\equiv \begin{cases} \text{"dynamical action" of the brief} \\ \text{classical path } (x, t+dt) \leftarrow (y, t) \end{cases} \end{aligned} \tag{14.2}$$

—then the Schrödinger equation fell into his lap! And upon introduction of (14.1) into (12) he obtained a pretty statement which in the time-independent case (to which I restrict myself simply as a notational convenience) reads

$$\begin{aligned} K(x, t; y, 0) &= \int \cdots \iint \exp\Big\{ \tfrac{i}{\hbar} \underbrace{\sum_{k=0}^{N} S(x_{k+1}, \tau; x_k, 0)}_{\text{action of a segmented path } (x,t)\leftarrow(y,0) \text{ with "dynamical" segments}} \Big\} \frac{dx_1}{A}\frac{dx_2}{A}\cdots\frac{dx_N}{A} \\ &\equiv \int_{\text{paths } (x,t)\leftarrow(y,0)} e^{\frac{i}{\hbar}S[\text{path}]}\, \mathcal{D}[\text{path}] \end{aligned} \tag{15}$$

The key circumstances here are supplied by *classical* mechanics

$$\begin{aligned} S[\text{path } (x,t) \leftarrow (y,0)] &= \int_0^t L(\text{path})\, dt' \\ &= \sum_{\text{segments}} \int L(\text{dynamical segment})\, dt' \\ &= \sum_{\text{segments}} S[\text{dynamical segment}] \\ &= \sum_{\text{segments}} S(\text{segmental endpoint}; \text{segmental endpoint}) \end{aligned}$$

[16] For an amusing account of the moment of discovery, see page 129 in Gleick.[2]

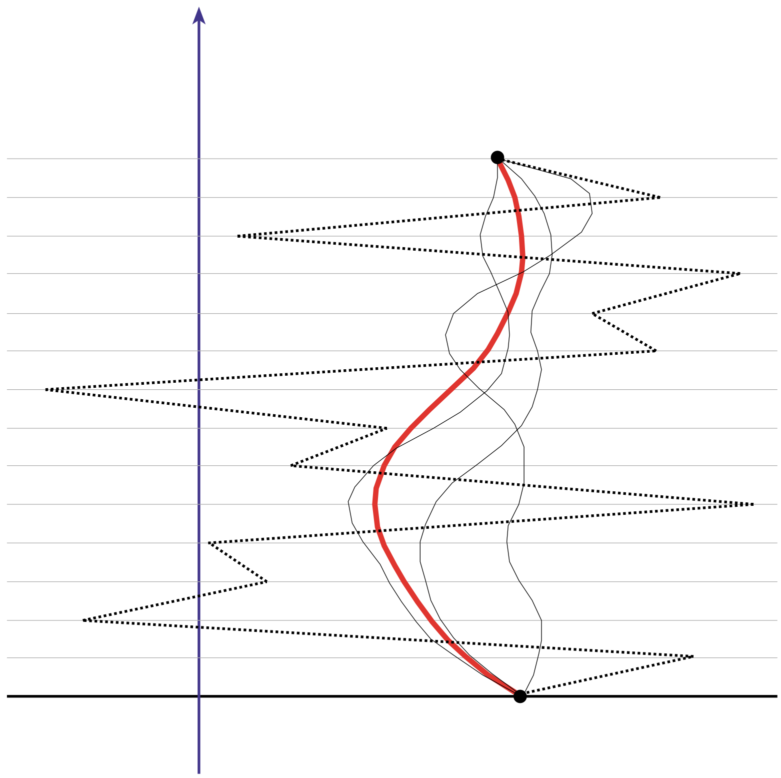


FIGURE 2: *Hamilton's principle asks us to examine the hypothetical paths linking specified endpoints, and to associate classical reality with the (red) path that extremizes the action: $\delta S = 0$. Feynman, on the other hand, is led to view the test paths as "statistically independent channels," and to assign to each a*

$$\text{path amplitude} \sim e^{\frac{i}{\hbar}(\text{classical path action})}$$

*Feynman's paths (represented in the figure by the dashed spline) differ, however, from those contemplated by Hamilton/Lagrange in that almost all are almost everywhere non-differentiable.*

and hinge on the fact that while the action

$$S[x(t)] \equiv \int_0^t L\big(\dot{x}(t'), x(t')\big)\, dt' \text{ is by nature a } \textit{functional}$$

it becomes a 2-point *function* (function of the endpoints) if $x(t)$ is "dynamical" in the sense that it *satisfies the equations of motion*, together with the specified endpoint conditions $x(t_0) = x_0$ and $x(t_1) = x_1$:

$$S[x(t)] = S(x_1, t_1; x_0, t_0) \qquad \text{if} \qquad \delta S[x(t)] = 0$$

Lagrange/Hamilton had contemplated a population of hypothetical "test paths" in order to lend meaning to $\delta S[\text{path}]$, and thus to be placed in position to trace the equations of motion (Lagrange's equations) to an underlying variational principle (Hamilton's Principle). But at (15) Feynman has, in a manner of speaking, declared the test paths to be each as "real" as any other, in the sense that each has quantum mechanical work to do: at the beginning of §4 in the RMP paper[5] we encounter his

> *Postulate II: The paths contribute equally in magnitude, but the phase of their contribution is the classical action (in units of $\hbar$); i.e., the time integral of the Lagrangian taken along the path.*

In his §7 Feynman rehearses Dirac's "very beautiful" argument to the effect that in the limit $\hbar \to 0$ one can expect the contribution of the classical path to predominate. In that sense, *Feynman's Principle gives back Hamilton's Principle in the classical limit*.

Though the point seems never to have bothered Feynman very much,[17] many people have looked upon the "normalization factors" $1/A$ in (15) as—since $A$ must be assigned value on a base-by-case basis—a formal blemish. It is a blemish with which Pauli, in particular, was not content to live. He had the genius to observe[6] that the entirely *classical* object[18]

$$K_C(\boldsymbol{x},\tau;\boldsymbol{y},0) \equiv (ih)^{-n/2}\sqrt{D}\cdot\exp\left\{\tfrac{i}{\hbar}S(\boldsymbol{x},\tau;\boldsymbol{y},0)\right\} \tag{16}$$
$$D \equiv D(\boldsymbol{x},\tau;\boldsymbol{y},0) \equiv (-)^n \det\left\|\frac{\partial^2 S(\boldsymbol{x},\tau;\boldsymbol{y},0)}{\partial x^r \partial y^s}\right\|$$

becomes a solution of the Schrödinger equation when $\tau$ is small—becomes, in fact, not just any old solution, but the "fundamental solution:"

$$\downarrow$$
$$= \delta(\boldsymbol{x}-\boldsymbol{y}) \quad \text{at } \tau = 0$$

That same observation was reported simultaneously by Cécile Morette,[7] who in her footnote 3 acknowledges that she is "greatly indebted" to Léon Van Hove (then also a visitor at the Institute for Advanced Studies); it was Van Hove who directed her attention to papers by P. Jordan (1926) and J. H. Van Vleck (1928) in which (16) had previously appeared. Neither Pauli nor his student Choquard[8] acknowledge any such influence. In the modern literature

[17] See, however, page 33 in Feynman & Hibbs, *Quantum Mechanics & Path Integrals* (1965).

[18] It is—for reasons that will become evident in a moment—convenient for purposes of the present discussion to work in $n$-dimensions; one-dimensional results can be recovered by specialization, but are in themselves too simple to reveal what is going on.

$D$ is usually called the "Van Vleck determinant."[19]

Van Vleck's principal references are to papers reporting the invention—simultaneously and independently by Gregor Wentzel, H. A. Kramers and Léon Brillouin (1926)—of what has become known as the "WKB approximation."[20] Those authors worked in one dimension. Van Vleck's objective was to show that the deeper significance of a certain characteristic detail becomes evident only when one works in higher dimension.

The "semi-classical (or WKB) approximation," in all of its variant forms, proceeds by expansion in powers of $\hbar$. The deep point recognized and exploited

---

[19] John Van Vleck (1899–1980) wrote—under the direction of Edwin Kemble, at Harvard—the first US dissertation treating a quantum mechanical topic, and during the 1920's was a leading player (first at the University of Minnesota, later at the University of Wisconsin) in the application of (old) quantum mechanics to diverse physical problems. Some of his results reportedly contributed to Bohr's development of the Correspondence Principle, and many hinged on deft use of the interplay between classical and quantum mechanical ideas. In 1935 he returned to Harvard as a faculty member and undertook the work (quantum theory of magnetic susceptibility) that led to a Nobel Prize in 1977.

By the late 1950's the relevance to Feynman formalism of the early paper to which I have referred ("The correspondence principle in the statistical interpretation of quantum mechanics," Proceedings of the National Academy of Sciences **14**, 178 (1928)) had become well known. So when a short meeting on the subject took place at Brandeis University (where I was then a graduate student) Van Vleck was asked to speak. He began with the remark that reprints of his papers either disappeared at once or sat on his shelf for years, and that the stack of papers beside him was a nearly complete set of reprints for the paper in question . . . which is how I acquired my treasured copy.

Now it happens that one Mrs. Miner T. (Connie) Patton, who was for many years secretary to the president of Reed College, had earlier in her career been secretary to the physics department at Harvard (her patient secretarial assistance is acknowledged in the preface of Kemble's *Fundamental Principles of Quantum Mechanics* (1937)), and had established a life-long friendship with Van Vleck. So it happened that when, in the early 1970's, Van Vleck came to Portland to pursue (at the Portland Art Museum) his deep interest in Japanese prints (of which he had a large and important collection) he stayed with his old friend, and I had an opportunity to spend some time with him. Ignorant as I then was of the strong classical component in his early work, I asked him how he *came* to write the "Van Vleck determinant paper." He responded that the essential idea was really due to Oppenheimer (then 27, and his junior by five years), suggested to him in conversation at one of the famous Ann Arbor Summer Schools.

[20] See, for example, David Griffiths, *Introduction to Quantum Mechanics* (1994), Chapter 8.

by Pauli is that

*for short times the WKB approximation becomes exact:*

thus was he led to reinvent Van Vleck's "$n$-dimensional WKB theory" and to assign to (15) this sharpened ($A$-independent) meaning:

$$K(\boldsymbol{x},t;\boldsymbol{y},0) = \lim_{N\to\infty}\int\cdots\iint\prod_{k=0}^{N} K_C(\boldsymbol{x}_{k+1},\tau;\boldsymbol{x}_k,0)\,d\boldsymbol{x}_1 d\boldsymbol{x}_2\cdots \boldsymbol{d}x_N \tag{17}$$

where $\boldsymbol{x}_0 \equiv \boldsymbol{y}$, $\boldsymbol{x}_{N+1} \equiv \boldsymbol{x}$, $\tau \equiv t/(N+1)$ and $K_C(\boldsymbol{x}_{k+1},\tau;\boldsymbol{x}_k,0)$ is given by (16).

Taking (15) as his postulated point of departure, Feynman proceeds to demonstrate

- recovery of the Schrödinger equation
- recovery of such fundamental statements as $[\mathbf{x},\mathbf{p}] = i\hbar\mathbf{I}$

and that the path integral method supplies new ways to think about (say) the two-slit experiment and the quantum theory of measurement, new approaches to approximation theory, and much else. But before we take up such matters we must secure some of the claims made in preceding paragraphs.[21]

**Demonstration that quantum mechanics is briefly classical.** Our objective here will be to secure the credentials of the Van Vleck/Pauli construction (16). In an effort to keep simple things simple I look first to systems of the design

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

and then will, by degrees, relax the simplifying assumptions built into that design. We look first to the outlines of Van Vleck's contribution.

The Schrödinger equation reads $-\frac{\hbar^2}{2m}\psi_{xx} + V\psi = i\hbar\psi_t$, and if we assume the wave function $\psi(x,t)$ to have been written in polar form

$$\psi = A\,e^{\frac{i}{\hbar}S}$$

assumes the form

$$\Big\{A\cdot\big[\tfrac{1}{2m}S_x^2 + V + S_t\big] - i\hbar\big[\tfrac{1}{m}(S_xA_x + \tfrac{1}{2}AS_{xx}) + A_t\big] - \hbar^2\tfrac{1}{2m}A_{xx}\Big\}e^{\frac{i}{\hbar}S} = 0$$

Divide by $A\,e^{\frac{i}{\hbar}S}$, obtain a power series in $\hbar$ which in the WKB tradition we

[21] How did Dirac—who had all the essential elements in hand already in 1932—manage *not* to invent the path integral method? The question is frequently posed, and must surely have been asked of Dirac himself, but I am aware of no printed record of his thoughts on the matter. It is my guess that he lacked Feynman's intense but ideosyncratic *motivation* to develop a functional alternative to the standard (Hamiltonian) quantization procedure; that he did not *expect* to achieve more than an interesting but imperfect analogy, so did not seriously try to; that he was preoccupied in 1932 with problems posed by quantum field theory, which made invention of the method too big a bite, even for Dirac.

interpret to mean that in

$$0^{\text{th}}\ \text{order}: \quad \tfrac{1}{2m}S_x^2 + V + S_t = 0 \tag{18.0}$$
$$1^{\text{st}}\ \text{order}: \quad \tfrac{1}{m}\big\{S_x(\log A)_x + \tfrac{1}{2}S_{xx}\big\} + (\log A)_t = 0 \tag{18.1}$$
$$2^{\text{nd}}\ \text{order}: \quad A_{xx}/A = 0 \tag{18.2}$$

These equations are collectively exact. In WKB approximation we agree to abandon the final equation . . . or better: to look upon

$$\frac{1}{A}\frac{\partial^2}{\partial x^2}A \approx 0$$

as a "consistency condition" imposed upon information extracted from the leading pair of equations. At (18.0) we have recovered precisely the Hamilton-Jacobi equation (the equation from which Schrödinger historically *extracted* the equation that bears his name), while multiplying (18.1) by $2A^2$ yields an equation that can be written

$$(\tfrac{1}{m}S_x A^2)_x + (A^2)_t = 0 \tag{19}$$

and has therefore the design of a one-dimensional continuity equation. It has evidently to do with *conservation of probability*, since

$$A^2 = |\psi|^2 = \text{probability density} \tag{20}$$

We lend the "look of generality" to the preceding results by noting that the Hamilton-Jacobi equation can be expressed

$$H(S_x, x) + S_t = 0 \tag{21.0}$$

and that (19) can be written

$$(vA^2)_x + (A^2)_t = 0 \tag{21.1}$$

where

$$v \equiv v(x,t) \equiv \frac{\partial H(p,x)}{\partial p}\Big|_{p\to S_x}$$

is on dimensional grounds a "velocity." For $n$-dimensional systems of the type $H = \frac{1}{2m}\boldsymbol{p}\cdot\boldsymbol{p} + V(\boldsymbol{x})$ we in place of (18.0) and (19) obtain

$$\tfrac{1}{2m}\boldsymbol{\nabla}S\cdot\boldsymbol{\nabla}S + V + S_t = 0 \tag{22.0}$$

and

$$\boldsymbol{\nabla}\cdot(A^2\tfrac{1}{m}\boldsymbol{\nabla}S) + (A^2)_t = 0 \tag{22.1}$$

In the most general classical setting one contemplates Hamiltonians that depend unrestrictedly upon indefinitely many *generalized* coordinates $\boldsymbol{q} \equiv \{q^1, q^2, \dots, q^n\}$ and their conjugate momenta $\boldsymbol{p} \equiv \{p_1, p_2, \dots, p_n\}$; the Hamilton-Jacobi equation then reads

$$H\left(\tfrac{\partial S}{\partial q^1},\ldots,\tfrac{\partial S}{\partial q^n},q^1,\ldots,q^n\right)+\tfrac{\partial S}{\partial t}=0 \tag{23.0}$$

and the associated continuity equation becomes

$$\sum_{i=1}^{n}\frac{\partial}{\partial q^i}\Big[v^iA^2\Big]+\frac{\partial}{\partial t}A^2=0 \tag{23.1}$$

$$v^i\equiv v^i(\boldsymbol{q},t)\equiv\frac{\partial H(\boldsymbol{p},\boldsymbol{q})}{\partial p_i}\Big|_{\boldsymbol{p}\to\boldsymbol{\nabla}S}$$

It is in this general language that I conduct the next phase of this discussion.

Occupying a distinctive place among the solutions of the Hamilton-Jacobi equation are the so-called "fundamental solutions," familiar to us as "dynamical action"

$$S(\boldsymbol{q},t;\boldsymbol{q}_0,t_0)=\int_{t_0}^{t}L\big(\dot{\boldsymbol{q}}(t'),\boldsymbol{q}(t')\big)\,dt'$$

—the action of the *dynamical* path $\boldsymbol{q}(t')$ that links $(\boldsymbol{q}_0,t_0)$ to $(\boldsymbol{q},t)$. The function $S(\boldsymbol{q},t;\boldsymbol{q}_0,t_0)$ is a *two-point* action function: in the leading variables $(\boldsymbol{q}_0,t_0)$ it satisfies the H-J equation (23.0), while in the trailing variables it satisfies the *time-reversed* H-J equation

$$H\left(\tfrac{\partial S}{\partial q_0^1},\ldots,\tfrac{\partial S}{\partial q_0^n},q_0^1,\ldots,q_0^n\right)-\tfrac{\partial S}{\partial t}=0 \tag{23.0}$$

In phase space it is the *Legendre generator* of the $t$-parameterized canonical transformation (dynamical phase flow) the *Lie generator* of which is $H(\boldsymbol{p},\boldsymbol{q})$, and does its work this way: write

$$p_i=+\frac{\partial S(\boldsymbol{q},t;\boldsymbol{q}_0,t_0)}{\partial q^i}\qquad\text{and}\qquad p_{0i}=-\frac{\partial S(\boldsymbol{q},t;\boldsymbol{q}_0,t_0)}{\partial q_0^i} \tag{24}$$

By algebraic inversion of the latter obtain $q^i(t;\boldsymbol{q}_0,\boldsymbol{p}_0,t_0)$, and by insertion into the former obtain $p_i(t;\boldsymbol{q}_0,\boldsymbol{p}_0,t_0)$.

It is the upshot of (what I call) "Van Vleck's theorem" that

$$A^2=D(\boldsymbol{q},t;\boldsymbol{q}_0,t_0)\equiv\left|\frac{\partial\boldsymbol{p}_0(\boldsymbol{q},t;\boldsymbol{q}_0,t_0)}{\partial\boldsymbol{q}}\right|=(-)^n\det\left\|\frac{\partial S(\boldsymbol{q},t;\boldsymbol{q}_0,t_0)}{\partial q^i\partial q_0^j}\right\| \tag{25}$$

satisfies (23.1). The quantum mechanical utility of the theorem should not be allowed to obscure the fact that it is itself entirely classical (though absent from every classical mechanics text known to me). I turn now to the proof of Van Vleck's theorem:[22]

[22] I have been following QUANTUM MECHANICS (1967), Chapter 1, pages 91 *et seq*, where a proof patterned upon Van Vleck's own begins on page 98. Here I present an alternative argument adapted from CLASSICAL MECHANICS (1983), pages 452–456.

Hit the H-J equation (23.0) with $\partial_i \equiv \frac{\partial}{\partial q^i}$ to obtain

$$H_i + v^k S_{ki} + S_{ti} = 0$$

where

$$v^k \equiv \frac{\partial H(\boldsymbol{p},\boldsymbol{q})}{\partial p_k}\Big|_{\boldsymbol{p}\to\boldsymbol{\nabla}S} \quad \text{has been joined now by} \quad H_i \equiv \frac{\partial H(\boldsymbol{p},\boldsymbol{q})}{\partial q_i}\Big|_{\boldsymbol{p}\to\boldsymbol{\nabla}S}$$

Subsequent differentiation by $\partial_j \equiv \frac{\partial}{\partial q_0^i}$ (note the sequestered status of the variables $\boldsymbol{q}_0$, which enter only *via* $S(\boldsymbol{q},t;\boldsymbol{q}_0,t_0)$) gives

$$v_i^k S_{kj} + v^{k\ell} S_{ki} S_{\ell j} + v^k S_{kij} + S_{tij} = 0$$

with

$$v_i^k \equiv \frac{\partial^2 H(\boldsymbol{p},\boldsymbol{q})}{\partial q^i \partial p_k}\Big|_{\boldsymbol{p}\to\boldsymbol{\nabla}S} \qquad \text{and} \qquad v^{k\ell} \equiv \frac{\partial^2 H(\boldsymbol{p},\boldsymbol{q})}{\partial p_k \partial p_\ell}\Big|_{\boldsymbol{p}\to\boldsymbol{\nabla}S}$$

Elementary manipulations now supply

$$(v^k\partial_k + \partial_t)S_{ji} = -S_{jk}\underbrace{(v_i^k + v^{k\ell}S_{\ell i})}_{} \\ = \partial_i v^k \equiv V^k{}_i(\boldsymbol{q},t;\boldsymbol{q}_0,t_0)$$

which in an obvious matrix notation becomes

$$(v^k\partial_k + \partial_t)\,\mathbb{S} = -\,\mathbb{S}\,\mathbb{V}$$

Assume that $D = (-)^n \det\mathbb{S} \neq 0$. Then

$$\mathbb{S}^{-1}(v^k\partial_k + \partial_t)\,\mathbb{S} = -\,\mathbb{V} \tag{26.1}$$

gives (with the abandonment of a lot of potentially useful information)

$$\mathrm{tr}\big\{\mathbb{S}^{-1}(v^k\partial_k + \partial_t)\,\mathbb{S}\big\} = -\partial_k v^k$$

It is, however, a corollary of the elegant identity $\log\det\mathbb{M} = \mathrm{tr}\log\mathbb{M}$ that

$$\mathrm{tr}\big\{\mathbb{M}^{-1}\delta\,\mathbb{M}\big\} = \mathrm{M}^{-1}\delta\,\mathrm{M}$$

where $\mathbb{M}$ is any matrix, $\mathrm{M} \equiv \det\mathbb{M}$ and $\delta$ is any first-order differential operator. We are in position therefore to write $(v^k\partial_k + \partial_t)D = -D\partial_k v^k$, from which Van Vleck's theorem

$$\partial_k(v^k D) + \partial_t D = 0 \tag{26.2}$$

immediately follows. Sweet ...if I do say so myself! It remains, however, to clarify *what Van Vleck's theorem is trying to tell us.*

Solutions $S(\boldsymbol{q},t)$ serve, by $p_i = \partial_i S$, to inscribe moving surfaces

$$p_i = p_i(\boldsymbol{q},t)$$

on $2n$-dimensional phase space. The dynamical flow of phase points is in general described by equations of the form

$$\begin{aligned}\boldsymbol{q}_0 &\longmapsto \boldsymbol{q} = \boldsymbol{q}(t;\boldsymbol{q}_0,\boldsymbol{p}_0)\\ \boldsymbol{p}_0 &\longmapsto \boldsymbol{p} = \boldsymbol{p}(t;\boldsymbol{q}_0,\boldsymbol{p}_0)\end{aligned}$$

but for points resident on the $S$-surface one can use $\boldsymbol{p}_0 = \boldsymbol{\partial} S(\boldsymbol{q}_0,t_0)$ to obtain

$$\boldsymbol{q}_0 \longmapsto \boldsymbol{q} = \boldsymbol{q}(t;\boldsymbol{q}_0,\boldsymbol{\partial} S(\boldsymbol{q}_0,t_0)) \equiv \boldsymbol{q}(t;\boldsymbol{q}_0,t_0)$$

which (note the disappearance of $\boldsymbol{p}_0$) serves to describe a time-dependent map (*i.e.* to install a moving coordinate system) on configuration space. The Jacobian

$$\left|\frac{\partial \boldsymbol{q}}{\partial \boldsymbol{q}_0}\right| \equiv \left|\frac{\partial(q^1,q^2,\ldots,q^n)}{\partial(q_0^1,q_0^2,\ldots,q_0^n)}\right|$$

enters into the description

$$dq^1 dq^2\cdots dq^n = \left|\frac{\partial \boldsymbol{q}}{\partial \boldsymbol{q}_0}\right| dq_0^1 dq_0^2\cdots dq_0^n$$

of the local dilation achieved by the map, and *via*

$$w(\boldsymbol{q},t)dq^1 dq^2\cdots dq^n = w(\boldsymbol{q}_0,t_0)dq_0^1 dq_0^2\cdots dq_0^n$$

informs us that densities written onto configuration space transform "as scalar densities:"

$$w(\boldsymbol{q},t) = \left|\frac{\partial \boldsymbol{q}}{\partial \boldsymbol{q}_0}\right|^{-1}\cdot w(\boldsymbol{q}_0,t_0) \tag{27}$$

But the elementary theory of Jacobians (which historically represent one of the first applications of the theory of "determinants"), taken in combination with (25), supplies

$$\begin{aligned}\left|\frac{\partial \boldsymbol{q}}{\partial \boldsymbol{q}_0}\right|^{-1} &= \left|\frac{\partial \boldsymbol{p}_0}{\partial \boldsymbol{q}_0}\right|^{-1}\cdot\left|\frac{\partial \boldsymbol{p}_0}{\partial \boldsymbol{q}}\right|\\ &= (\text{constant})\cdot(\text{Van Vleck determinant } D)\end{aligned}$$

Evidently $D$ quantifies the changing size of the moving "shadow" (projection onto configuration space) of a "patch" inscribed on the $S$-surface as it "drifts with the dynamical flow" in phase space. State points marked on interior of the original patch are mapped to points interior to the dynamical image of that patch, and the conservation law (26.2) refers to the shadow of that elementary proposition. We can state that Van Vleck's theorem is a projective consequence of Liouville's theorem (*i.e.*, of the incompressibility of Hamiltonian flow), but with this proviso: the measure-preserving sets contemplated by Van Vleck are not Liouville's "blobs" in phase space, but "patches (with area but no volume) inscribed on the surfaces that arise by $\boldsymbol{p} = \partial S/\partial \boldsymbol{q}$ from $S(\boldsymbol{q},t;\boldsymbol{q}_0,t_0)$. The following figures are intended to clarify the situation.

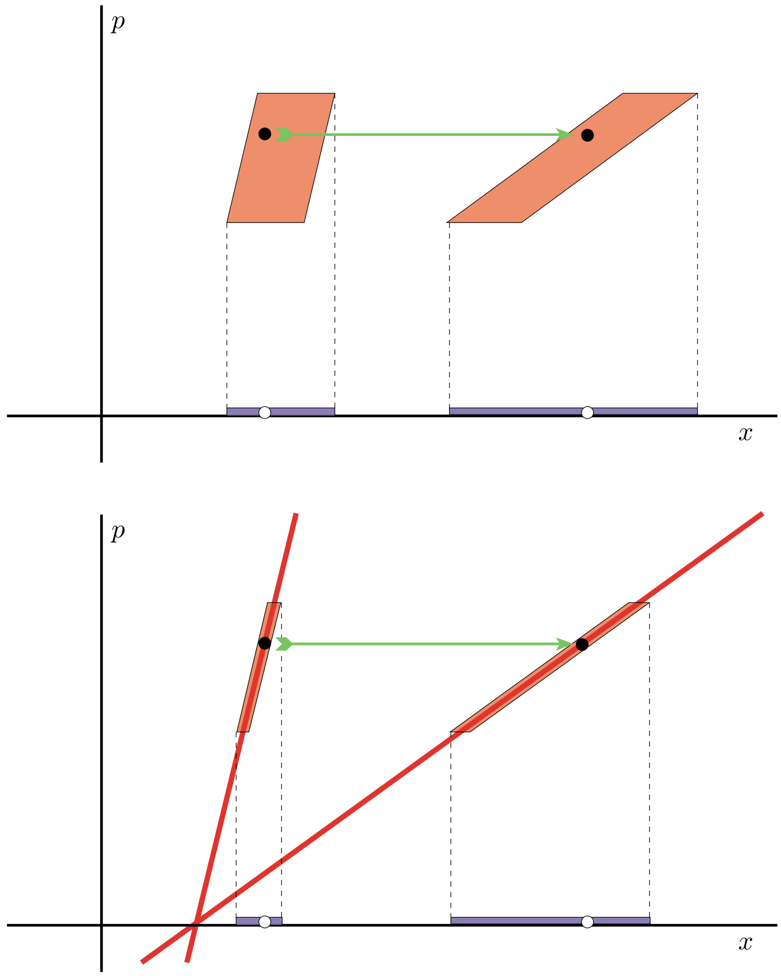



FIGURE 3: *The red parallelogram at upper left in the top figure represents a typical "blob" in the 2-dimensional phase space of a free particle. A single designated state point* • *resides within the blob. The green arrow shows the effect of dynamical phase flow, which by Liouville's theorem is area-preserving. The result of projecting that imagery onto configuration space is shown in blue. The lower figure refers to the related ideas of special interest to Van Vleck. Phase space has become 2-dimensional so "S-surfaces" have become curves—actually straight lines (shown in red), since*

$$S(x,t;x_0,t_0) = \frac{m}{2}\frac{(x-x_0)^2}{t-t_0} \implies p(x;t;x_0,t_0) = m\frac{x-x_0}{t-t_0}$$

*I have attempted to tatto a "patch" on such a surface, to represent the dynamical transport of the patch, and (in blue) to represent the motion of the* projected shadow *of the patch.*

Van Vleck's $D(\boldsymbol{q},t;\boldsymbol{q}_0,t_0)$ describes what in the caption I call the "projected shadow of the patch," and his continuity equation (26.2) is in effect a statement that "a mite riding on the patch will see a locally concerved density of freckles."

Turn now with Pauli to consideration of what Van Vleck has to say when $t$ is small (and called $\tau$ to emphasize the fact). It proves convenient at this point to abandon generalized coordinates $\boldsymbol{q}$ in favor of Cartesian coordinates $\boldsymbol{x}$, and to restrict our attention to systems of the form

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

Additionally, we set $t_0 = 0$ and agree to regard $\boldsymbol{x}_0$ and $\boldsymbol{y}$ as interchangeable symbols. In the short run we expect to have

$$\boldsymbol{x}(t') = \boldsymbol{y} + \tfrac{\boldsymbol{x}-\boldsymbol{y}}{\tau}\,t'$$

giving

$$S(\boldsymbol{x},\tau;\boldsymbol{y},0) = \int_0^\tau \Big\{\tfrac{1}{2}m\tfrac{\boldsymbol{x}-\boldsymbol{y}}{\tau}\cdot\tfrac{\boldsymbol{x}-\boldsymbol{y}}{\tau} - V\big(\boldsymbol{y}+\tfrac{\boldsymbol{x}-\boldsymbol{y}}{\tau}\,t'\big)\Big\}\,dt'$$

which—simply to avoid notational distractions—I prefer to discuss in the one dimensional case: we in that case have

$$S(x,\tau;y,0) = S_0(x,\tau;y,0) - \int_0^\tau V\big(y+\frac{x-y}{\tau}\,t'\big)\,dt' \tag{28}$$

where

$$\begin{aligned} S_0(x,\tau;y,0) &\equiv \frac{m}{2}\frac{(x-y)^2}{\tau} \\ &= \begin{cases}\text{the dynamical action of a}\\ \text{FREE PARTICLE at time } \tau\end{cases} \end{aligned}$$

Interpretation and management of the second term on the right side of (28) is a more delicate matter. It is evident that

$$\lim_{y\to x}\int_0^\tau V\big(y+\frac{x-y}{\tau}\,t'\big)\,dt' = V(x)\,\tau$$

but we have at the moment no special *interest* in setting $y = x$. Differentiation of (28)—recall $p = \partial S/\partial x$ and $F = -\partial V/\partial x$—gives

$$p(x,\tau) - p_0 = \int_0^\tau F\big(y+\frac{x-y}{\tau}\,t'\big)\,\tfrac{t'}{\tau}\,dt'$$

where $p_0 = \partial S_0/\partial x = m(x-y)/\tau$ is the conserved momentum of a particle that moves *freely* from $(y,0)$ to $(x,\tau)$: in other words,

$$\text{change of momentum} = \text{net impulse}$$

The interpretation is nice, but the integral is no less awkward than the one encountered at (28): we cannot expand in powers of $\tau$ because $\tau$ lives in the denominator, and we cannot expand in powers of $(x-y)$ because $(x-y)$ is in general not small. Pauli's plan of attack is to introduce $S_1(x,\tau;y,0)$ by

$$-\int_0^\tau V\big(y+\frac{x-y}{\tau}\,t'\big)\,dt' = -V(x)\,\tau + S_1(x,\tau;y,0)$$

and then, by a fairly intricate function-theoretic argument,[22] to show that under certain weak hypotheses the function $S_1(x,\tau;y,0)$ possesses certain essential properties. I will proceed more simply (but less generally): Assume

$$V(x) = k\cdot(x/a)^n \quad : \quad [a] = \text{length}$$

We can then actually *perform* the integral, and obtain

$$\int_0^\tau V\big(y+\frac{x-y}{\tau}\,t'\big)\,dt' = (k/a^n)\tfrac{1}{n+1}\big(x^n + x^{n-1}y + x^{n-2}y^2 + \cdots + y^n\big)\,\tau$$

Insert

$$\begin{aligned} S(x,\tau;y,0) &= \tfrac{1}{2}m\frac{(x-y)^2}{\tau} \\ &\quad - (k/a^n)\tfrac{1}{n+1}\big(x^n + x^{n-1}y + x^{n-2}y^2 + \cdots + y^n\big)\,\tau \end{aligned} \tag{29}$$

into the Hamilton-Jacobi equation and obtain

$$\begin{aligned} \tfrac{1}{2m}S_x^2 + (k/a^n)x^n + S_\tau &= \begin{cases} \frac{k^2}{2ma^2}\,\frac{1}{2^2}\tau^2 & : \quad n=1 \\ \frac{k^2}{2ma^4}\,\frac{1}{3^2}(2x+y)^2\,\tau^2 & : \quad n=2 \\ \frac{k^2}{2ma^6}\,\frac{1}{4^2}(3x^2+2xy+y^2)^2\,\tau^2 & : \quad n=3 \\ \frac{k^2}{2ma^8}\,\frac{1}{5^2}(4x^3+3x^2y+2xy^2+y^3)^2\,\tau^2 & : \quad n=4 \\ & \ \ \vdots \end{cases} \\ &= 0 + \text{expression of order } O(\tau^2) \end{aligned} \tag{30}$$

Evaluation of the Van Vleck determinant $D = (-)^{\text{dimension}}|\partial^2 S/\partial x\partial y|$ gives

$$\begin{aligned} D(x,\tau;y,0) &= \begin{cases} \frac{m}{\tau} & : \quad n=1 \\ \frac{m}{\tau} + (k/a^2)\frac{1}{3}\,\tau & : \quad n=2 \\ \frac{m}{\tau} + (k/a^3)\frac{1}{4}(2x+2y)\,\tau & : \quad n=3 \\ \frac{m}{\tau} + (k/a^4)\frac{1}{5}(3x^2+4xy+3y^2)\,\tau & : \quad n=4 \\ & \ \ \vdots \end{cases} \\ &= \quad \tfrac{m}{\tau} + \text{expression of order } O(\tau) \end{aligned} \tag{31}$$

[22] See page 169 in the MIT edition of the lecture notes previously cited.[6]

The preceding discussion is readily generalized: Assume the potential can be developed $V(x) = \sum V_n x^n$. Introduce the "superpotential"

$$W(x) \equiv \int_0^x V(x')\,dx' = \sum \tfrac{1}{n+1} V_n x^{n+1}$$

that takes its name from the circumstance that $V(x) = \frac{d}{dx}W(x)$. Use the pretty identity

$$\frac{x^{n+1} - y^{n+1}}{x-y} = x^n + x^{n-1}y^1 + n^{n-2}y^2 + \cdots + x^1 y^{n-1} + y^n$$

to obtain

$$\frac{W(x) - W(y)}{x-y} = \sum_n V_n \frac{x^n + x^{n-1}y^1 + n^{n-2}y^2 + \cdots + x^1 y^{n-1} + y^n}{n+1}$$

The implication is that for potentials of the assumed form we can write

$$\int_0^\tau V\big(y + \frac{x-y}{\tau}\,t'\big)\,dt' = \frac{W(x) - W(y)}{x-y}\,\tau$$

and that (in the "uniform rectilinear approximation") the short-time classical action function can be described

$$S(x,\tau;y,0) = \tfrac{1}{2}m\frac{(x-y)^2}{\tau} - \frac{W(x) - W(y)}{x-y}\,\tau \tag{32}$$

The illustrative equations (30) and (31) become now instances of the more general statements

$$\tfrac{1}{2m}S_x^2 + V(x) + S_\tau = \tfrac{1}{2m}\Big[\frac{W(x)-W(y)}{(x-y)^2} - \frac{V(x)}{x-y}\Big]^2\tau^2 \tag{33.1}$$

and

$$D(x,\tau;y,0) = \tfrac{m}{\tau} - \Big[2\frac{W(x)-W(y)}{(x-y)^3} - \frac{V(x)-V(y)}{(x-y)^2}\Big]\tau \tag{33.2}$$

Insert $K(x,\tau;y,0) \equiv \sqrt{\alpha D}\,e^{\frac{i}{\hbar}S}$ into the Schrödinger equation and (holding the numerical value of $\alpha$ in suspension for the moment) obtain

$$\begin{aligned}
-\tfrac{\hbar^2}{2m}&K_{xx} + (k/a^2)x^n K - i\hbar K_\tau \\
&= \begin{cases} K\cdot \frac{k^2}{2ma^2}\frac{1}{2^2}\,\tau^2 & : \quad n=1 \\ K\cdot \frac{k^2}{2ma^4}\frac{1}{3^2}\Big\{\frac{3ma^2(2x+y)^2+6a^2 i\hbar\tau + k(2x+y)^2\tau^2}{3ma^3+k\tau^2}\Big\}\,\tau^2 & : \quad n=2 \\ & \vdots \end{cases} \\
&= 0 + K\cdot(\text{expression of order } O(\tau^2))
\end{aligned} \tag{34}$$

*Mathematica* supports this general conclusion even if one works from (33), but the expression that then appears on the right is too messy to be usefully written out. Observe finally that in each of the above cases (as also in the general case: work from (32) and (33.2))

$$\lim_{\tau\downarrow 0}\sqrt{\alpha D}\,e^{\frac{i}{\hbar}S} = \sqrt{\alpha m/\tau}\,\exp\Big\{\tfrac{i}{\hbar}\tfrac{m}{2}(x-y)^2/\tau\Big\}$$

Proceeding formally, we write

$$= \sqrt{\alpha m\pi/\beta\tau}\cdot\sqrt{\beta/\pi}e^{-\beta(x-y)^2} \quad\text{with}\quad \beta\equiv m/2i\hbar\tau$$

and observe that if we contrived to assign to $\beta$ a positive real part then we would have $\sqrt{\beta/\pi}e^{-\beta(x-y)^2}\longrightarrow\delta(x-y)$ in the limit $\beta\uparrow\infty$ (which is to say: in the limit $\hbar\tau\downarrow 0$). To achieve $\sqrt{\alpha m\pi/\beta\tau}=1$ we set $\alpha = 1/2\pi i\hbar$ and obtain

$$K(x,\tau;y,0)\equiv\sqrt{D/(ih)^{\text{dimension}}}\,e^{\frac{i}{\hbar}S} = \sqrt{\tfrac{m}{ih\tau}}\,e^{\frac{i}{\hbar}S} \tag{35}$$

The color coding serves here to emphasize that

- the $S(x,\tau;y,0)$ introduced at (29) is an *approximation* to the exact classical action $S(x,\tau;y,0)$, but an approximation so good that it fails only in $O(\tau^2)$ to satisfy the Hamilton-Jacobi equation;
- the $K(x,\tau;y,0)$ is an *approximation* to the $K_C(x,\tau;y,0)$ contemplated at (16), but an approximation so good that it conforms to the prescribed initial condition
$$\lim_{\tau\downarrow 0}K(x,\tau;y,0)=\delta(x-y)$$
and fails only in $O(\tau^2)$ to satisfy the Schrödinger equation.

Pauli[22] asserts on the basis of brief argument, and Choquard[8] works hard to establish in greater detail, the generality of the conclusions to which we have here been led.

It proves instructive to notice that—at least in favorable cases—the conclusions reached above can be recovered directly from the established principles of ordinary quantum mechanics. I discuss now how this can be accomplished. We had occasion already at (97) in Chapter 0 to observe that if—and a mighty big "if" it will turn out to be—we were in position to write

$$e^{\mathbf{H}} = {}_{\mathbf{x}}\big[e^{\mathcal{H}(x,p)}\big]_{\mathbf{p}}$$

then by straightforward application of the "mixed representation trick"[23] we would have

$$\begin{aligned} K(x,t;y,0) &\equiv (x|\mathbf{U}(t,0)|y)\\ &= \tfrac{1}{h}\int\exp\Big\{\tfrac{i}{\hbar}\big[\,p\cdot(x-y)-\mathcal{H}(x,p)\,t\big]\Big\}\,dp \end{aligned}\tag{36}$$

[23] It was an early contribution to the path integral literature (W. Tobocman, "Transition amplitudes as sums over histories," Nuovo Cimento **3**, 1213 (1956)) that brought the power of the "mixed representation trick" first to my attention.

Not quite: the basic idea sound, but it needs to be formulated a bit more carefully . . . as will emerge when we look to specific examples. Look first to the FREE PARTICLE where—trivially—

$$e^{\frac{1}{2m}\mathsf{p}^2} = {}_{\mathsf{x}}\big[e^{\frac{1}{2m}p^2}\big]_{\mathsf{p}} \quad \text{supplies} \quad \mathcal{H}(x,p) = \tfrac{1}{2m}p^2$$

which when introduced into (36) was already seen at (91.5) in Chapter 0 to give

$$K(x,t;y,0) = \sqrt{\tfrac{m}{iht}}\,\exp\Big\{\tfrac{i}{\hbar}S(x,t;y,0)\Big\} \tag{37}$$

$$S(x,t;y,0) = \tfrac{m}{2}\tfrac{(x-y)^2}{t}$$

These familiar results are, as it happens, *exact at all times* and preserve their designs as $t\downarrow\tau$. The example of a PARTICLE IN FREE FALL is in several respects more instructive. An easy application of Zassenhaus' identity[24] gives

$$e^{\frac{1}{2m}\mathsf{p}^2+mg\mathsf{x}} = {}_{\mathsf{x}}\big[e^{mgx+\frac{1}{2m}p^2-\frac{1}{2}i\hbar gp-\frac{1}{6}\hbar^2mg^2}\big]_{\mathsf{p}}$$

but to reach our objective we need this somewhat refined statement (in which I adopt the abbreviation $\beta\equiv i/\hbar$):

$$\begin{aligned}
e^{-\beta(\frac{1}{2m}\mathsf{p}^2+mg\mathsf{x})t} &= {}_{\mathsf{x}}\big[e^{-\beta(mgx+\frac{1}{2m}p^2)t-\frac{1}{2}(-\beta t)^2 i\hbar gp-(-\beta t)^3\frac{1}{6}\hbar^2mg^2}\big]_{\mathsf{p}}\\
&= {}_{\mathsf{x}}\big[e^{-\beta(mgx+\frac{1}{2m}p^2)t-\frac{1}{2}(-\beta t)\,gp\,t+(-\beta t)\frac{1}{6}mg^2t^2}\big]_{\mathsf{p}}\\
&= {}_{\mathsf{x}}\big[e^{-\beta\{mgx+\frac{1}{2m}p^2-\frac{1}{2}gp\,t+\frac{1}{6}mg^2t^2\}t}\big]_{\mathsf{p}}\\
&\equiv {}_{\mathsf{x}}\big[e^{-\beta\mathcal{H}(x,p,t)\,t}\big]_{\mathsf{p}}
\end{aligned}$$

Evidently we should, in general, expect the $\mathcal{H}$ in (36) to depend not only upon $x$ and $p$ but also upon $t$. Returning now with

$$\mathcal{H}(x,p,t) = \tfrac{1}{2m}p^2 + mgx - \tfrac{1}{2}g\,tp + \tfrac{1}{6}mg^2t^2$$

to (36) we confront again a (formal) Gaussian integral, and obtain

$$K(x,t;y,0) = \sqrt{\tfrac{m}{iht}}\,\exp\Big\{\tfrac{i}{\hbar}S(x,t;y,0)\Big\} \tag{38}$$

$$S(x,t;y,0) = \tfrac{m}{2}\tfrac{(x-y)^2}{t} - mg\tfrac{x+y}{2}\,t - \tfrac{1}{24}mg^2t^3$$

in precise agreement with (44.10) in Chapter 2. In discussion of that earlier result we noticed that the $S(x,t;y,0)$ described above is just the dynamical action that arises from the system $L=\frac{1}{2}m\dot{x}^2-mgx$. These gravitational equations are again *exact at all times*. In the limit $t\downarrow\tau$ we can abandon the $t^3$-term; we are led then back to an instance of (29):

$$S(x,\tau;y,0) = \tfrac{m}{2}\tfrac{(x-y)^2}{t} - mg\tfrac{1}{2}(x+y)\,t$$

[24] See again (73.5) in Chapter 0.

My final example—the HARMONIC OSCILLATOR—has, in effect, already been discussed: at (92.2) in Chapter 0 we obtain a result that can be expressed

$$e^{-\beta(\frac{1}{2m}\mathsf{p}^2+\frac{1}{2}m\omega^2\mathsf{x})t} = {}_{\mathsf{x}}\Big[e^{-\beta\left\{\frac{m\omega}{2}\tan\omega t\cdot x^2+(1-\sec\omega t)\cdot xp+\frac{1}{2m\omega}\tan\omega t\cdot p^2+\frac{1}{2}\log\sec\omega t\right\}}\Big]_{\mathsf{p}}$$

Notice that the exponentiated $\{\text{etc.}\}$ does not in this instance present $t$ as a factor: it might therefore seem a bit artifical to write $\{\text{etc.}\} = \mathcal{H}(x,p,t)\,t$, but remains possible in principle, and proves useful in practice. At short times obtain

$$\begin{aligned}\mathcal{H}(x,p,\tau) = \tfrac{1}{2m}p^2 + \tfrac{1}{2}m\omega^2x^2 &+ \tfrac{1}{4}(1-2xp)\omega^2\tau\\ &+ \tfrac{1}{3}\big\{\tfrac{1}{2m}p^2+\tfrac{1}{2}m\omega^2x^2\big\}\omega^2\tau^2\\ &+ \tfrac{1}{24}(1-5xp)\omega^4\tau^3\\ &+ \tfrac{2}{15}\big\{\tfrac{1}{2m}p^2+\tfrac{1}{2}m\omega^2x^2\big\}\omega^4\tau^4+\cdots\end{aligned}$$

Returning with this information to (36) we—by formal Gaussian integration—recover

$$\begin{aligned}K(x,t;y,0) &= \sqrt{\tfrac{m\omega}{ih\sin\omega t}}\,\exp\Big\{\tfrac{i}{\hbar}S(x,t;y,0)\Big\} \\ S(x,t;y,0) &= \tfrac{m\omega}{2\sin\omega t}\big[(x^2+y^2)\cos\omega t - 2xy\big]\end{aligned}\tag{39}$$

The classical significance of this result was discussed already at (0-94/95). In the short-time limit $t\downarrow\tau$ we have

$$\begin{aligned}S(x,\tau;y,0) = \tfrac{m}{2}\tfrac{(x-y)^2}{\tau} &- \tfrac{1}{6}m\omega^2(x^2+xy+y^2)\,\tau\\ &- \tfrac{1}{360}\,m\omega^4(4x^2 + 7xy + 4y^2)\,\tau^3\\ &- \tfrac{1}{15120}m\omega^6(16x^2+31xy+16y^2)\,\tau^5 - O(\tau^7)\end{aligned}$$

where I have highlighted the terms that arise by (29) in "uniform rectilinear approximation."

We have now in hand a short CATALOG OF EXACTLY SOLUABLE CASES,[25] which will serve us well as benchmarks when we look to cases that do *not* admit of exact analysis. And we are in position now to discuss, in concrete detail, the

$$\text{intimate relationship between } \lim_{t\downarrow0}K \text{ and } \lim_{\hbar\downarrow0}K$$

upon which the path integral formalism rests. For a FREE PARTICLE we introduce $\mathcal{H}(x,p,\tau) = \frac{1}{2m}p^2$ into (36)—now written

$$\begin{aligned}K(x,\tau;y,0) &= \tfrac{1}{h}\int\exp\Big\{\tfrac{i}{\hbar}g(p;x,y,\tau)\Big\}\,dp\\ g(p;x,y,\tau) &\equiv p\cdot(x-y) - \mathcal{H}(x,p,\tau)\,\tau\end{aligned}\tag{40}$$

[25] Notice that the Hamiltonian is, in each instance, *quadratic in* $p$ and *depends at most quadratically on* $x$. And that therefore the $\int dp$ is in each instance (formally) Gaussian.

and by direct integration obtain

$$K(x,\tau;y,0) = \sqrt{\tfrac{1}{ih}D}\,\exp\Big\{\tfrac{i}{\hbar}S(x,\tau;y,0)\Big\} \tag{41.1}$$
$$S = \tfrac{m}{2}\tfrac{(x-y)^2}{\tau} \quad\text{and}\quad D = -S_{xy} = \tfrac{m}{\tau}$$

But we can, on the other hand, use Kelvin's "method of stationary phase"[26] to obtain an *asymptotic* evaluation of the integral in the classical limit $\hbar \downarrow 0$: from $g'(\wp) = 0$ obtain $\wp = m(x-y)/\tau$ giving

$$g''(\wp) = -\tau/m \quad\text{and}\quad g(\wp) = \big[\tfrac{1}{2}m\big(\tfrac{x-y}{\tau}\big)^2\big]\tau$$

whence

$$\begin{aligned}\lim_{\hbar\downarrow 0} K(x,\tau;y,0) &= \lim_{\hbar\downarrow 0}\tfrac{1}{h}\int e^{\frac{i}{\hbar}g(p)}\,dp\\ &\sim \tfrac{1}{h}\sqrt{2\pi\hbar/g''(\wp)}\,e^{i\,[\frac{1}{\hbar}g(\wp)+\frac{\pi}{4}]}\\ &= \sqrt{\tfrac{m}{ih\tau}}\,\exp\Big\{\tfrac{i}{\hbar}\big[\tfrac{1}{2}m\big(\tfrac{x-y}{\tau}\big)^2\big]\tau\Big\}\end{aligned} \tag{41.2}$$

The point to notice is that the right sides of (41.1) and (41.2) are *identical*, but the context of the discussion has been too simple to make the point convincingly. Look again therefore to the PARTICLE IN FREE FALL : here[27]

$$g(p;x,y,\tau) \equiv p\cdot(x-y) - \mathcal{H}(x,p,\tau)\,\tau$$
$$\mathcal{H}(x,p,\tau) = \tfrac{1}{2m}p^2 + mgx - \tfrac{1}{2}gp\tau$$

which when introduced into (40) gives

$$K(x,\tau;y,0) = \sqrt{\tfrac{1}{ih}D}\,\exp\Big\{\tfrac{i}{\hbar}S(x,\tau;y,0)\Big\} \tag{42}$$
$$S = \tfrac{m}{2}\tfrac{(x-y)^2}{\tau} - mg\tfrac{x+y}{2}\tau \quad\text{and}\quad D = -S_{xy} = \tfrac{m}{\tau}$$

On the other hand, $g'(\wp) = (x-y) - \frac{1}{m}\wp\tau + \frac{1}{2}g\tau^2 = 0$ supplies

$$\wp = \tfrac{m}{\tau}(x - y + \tfrac{1}{2}g\tau^2)$$

whence

$$g(\wp) = \tfrac{m}{2}\tfrac{(x-y)^2}{\tau} - mg\tfrac{x+y}{2}\tau$$
$$g''(\wp) = -\tau/m$$

[26] See again (101.2) in Chapter 0.

[27] In the following discussion I highlight terms in which some approximation has actually taken place: this *usually but not always* means abandonment of terms of $O(\tau^2)$.

Introducing this information into

$$\begin{aligned}\lim_{\hbar\downarrow 0} K(x,\tau;y,0) &= \lim_{\hbar\downarrow 0} \tfrac{1}{h}\int e^{\frac{i}{\hbar}g(p)}\,dp\\ &\sim \tfrac{1}{h}\sqrt{2\pi\hbar/g''(\wp)}\,e^{i\,[\frac{1}{\hbar}g(\wp)+\frac{\pi}{4}]}\end{aligned}\tag{43}$$

we recover precisely (42). Look finally to the HARMONIC OSCILLATOR, where

$$\begin{aligned}g(p;x,y,\tau) &\equiv p\cdot(x-y)-\mathcal{H}(x,p,\tau)\,\tau\\ \mathcal{H}(x,p,\tau) &= \tfrac{1}{2m}p^2+\tfrac{1}{2}m\omega^2x^2+\tfrac{1}{4}(1-2xp)\omega^2\tau\\ &\quad+\tfrac{1}{3}\big\{\tfrac{1}{2m}p^2+\tfrac{1}{2}m\omega^2x^2\big\}\omega^2\tau^2\end{aligned}$$

Introduction into (40) gives a result of the design (42), except that now

$$S(x,\tau;y,0) = \tfrac{m}{2}\tfrac{(x-y)^2}{\tau}-\tfrac{1}{2}m\omega^2\tfrac{x^2+xy+y^2}{3}\tau \quad\text{and}\quad D=-S_{xy}=\tfrac{m}{\tau}\tag{44}$$

On the other hand, $g'(\wp)=(x-y)-\frac{1}{m}(1+\frac{1}{3}\omega^2\tau^2)\wp\tau+\frac{1}{2}gx\tau^2=0$ supplies

$$\wp = \tfrac{m}{\tau}(1+\tfrac{1}{3}\omega^2\tau^2)^{-1}(x-y+\tfrac{1}{2}gx\tau^2)$$

whence (entrusting the heavy labor to *Mathematica*)

$$\begin{aligned}g(\wp) &= \tfrac{m}{2}\tfrac{(x-y)^2}{\tau}-\tfrac{1}{2}m\omega^2\tfrac{x^2+xy+y^2}{3}\tau\\ g''(\wp) &= -(1+\tfrac{1}{3}\omega^2\tau^2)\cdot\tau/m\end{aligned}$$

which when introduced into (43) give back precisely (42/44). Remarkably (story of a wasted afternoon!), we achieve success in this instance only if we retain the blue term, which is of order $O(\tau^2)$.

It would be easy to argue that analysis of a quantum mechanical system reduces in the end to analysis of its (Green's function or) "propagator"... whose responsibility it is to guide the dynamical motion of the wave function, and into the design of which all spectral information is encoded:[28]

$$K(x,t;y,0)=\sum_n \psi_n(x)e^{-\frac{i}{\hbar}E_nt}\psi_n^*(y)\tag{45}$$

What we have now—in three specific contexts—demonstrated is a fact not at all evident in (45); namely, that "quantum mechanics is briefly classical" in this precise sense:

$$\lim_{\hbar\downarrow 0}K(x,t;y,0)\left\{\begin{array}{l}\text{is a "classical object," but provides an}\\ \text{accurate description of } K(x,t;y,0) \text{ if } t \text{ is small}\end{array}\right.\tag{46}$$

[28] I dismiss as an elegant quibble the observation the behind the scenes lurks the representation-independent object

$$\mathbf{U}(t,0)=\sum_n |n)e^{-\frac{i}{\hbar}E_nt}(n|$$

to which the preceding remark more properly relates.

It is that fact, used in combination with an elementary consequence of this representation

$$K(x,t_2;y,t_0)=\int K(x,t_2;\xi,t_1)\,d\xi\,K(\xi,t_1;y,t_0)$$

of the composition rule

$$\mathsf{U}(t_2,t_0)=\mathsf{U}(t_2,t_1)\mathsf{U}(t_1,t_0)\quad:\quad t_2>t_1>t_0$$

...that lies at the base of the Feynman formalism. The idea (and the source of the "path integral" concept) is to achieve finite quantum propagation by iteration of infinitesimally brief (therefore *classical*) propagation.[29] It is in an effort to secure the credentials of the method that I turn now to exploration of this question: Does (46) pertain generally—generally enough to embrace at least of systems of the form $\mathsf{H}=\frac{1}{2m}\mathsf{p}^2+V(\mathsf{x})$—or is it special to our three examples?

Notice first that if we (*i*) look to the $\tau$-expansion of the logarithm of the Fourier transform

$$\tfrac{1}{h}\int\exp\Big\{\tfrac{i}{\hbar}g(p;x,y,\tau)\Big\}\,dp$$
$$g(p;x,y,\tau)\equiv p\cdot(x-y)-\mathcal{H}(x,p,\tau)\,\tau$$

with

$$\mathcal{H}(x,p,\tau)\equiv\big[\tfrac{1}{2m}p^2+V_0+V_1x+V_2x^2\big]\big[1+\tfrac{2}{3m}V_2\tau^2\big]-\big[\tfrac{1}{2m}(V_1+2V_2x)\big]p\tau$$

or if (which is equivalent, but easier) we (*ii*) look to the $\tau$-expansion of the Legendre transform of $g(p;x,p,\tau)$ ...we are, by either procedure (and with the now indispensable assistance of *Mathematica*), led to

$$S(x,\tau;y,0)=\tfrac{m}{2}\tfrac{(x-y)^2}{\tau}-\big\{V_0+V_1\tfrac{x+y}{2}+V_2\tfrac{x^2+xy+y^2}{3}\big\}\tau$$

from which all three examples can be recovered as special cases. The surprising/disappointing fact, however, is that *there exists no modified* $\mathcal{H}$ *which by either procedure yields*[30]

$$S(x,\tau;y,0)=\tfrac{m}{2}\tfrac{(x-y)^2}{\tau}-\big\{V_0+V_1\tfrac{x+y}{2}+V_2\tfrac{x^2+xy+y^2}{3}+V_3\tfrac{x^3+x^2y+xy^2+y^3}{4}+\cdots\big\}\tau$$

[29] For time-independent systems we have

$$\mathsf{U}(t)=\big[\mathsf{U}(t/N)\big]^N$$

which we take to the limit $N\to\infty$.

[30] The essence of the argument: The physics of the matter stipulates that $p$ enters at most quadratically into the design of $g(p;x,y,\tau)$. Legendre transformation yields therefore an $S(x,\tau;y,0)$ into which $y$ enters at most quadratically. The blue terms lie therefore out of reach.

Means of escape from this impasse are provided by Zassenhaus' formula

$$e^{\mathsf{A}+\mathsf{B}} = e^{\mathsf{A}} e^{\mathsf{B}} e^{\mathsf{C}_2} e^{\mathsf{C}_3} \cdots \quad \text{with} \quad \begin{cases} \mathsf{C}_2 = -\frac{1}{2}[\mathsf{A},\mathsf{B}] \\ \mathsf{C}_3 = \frac{1}{6}[\mathsf{A},[\mathsf{A},\mathsf{B}]] + \frac{1}{3}[\mathsf{A},[\mathsf{A},\mathsf{B}]] \\ \quad\vdots \end{cases}$$

which entails

$$e^{(\mathsf{A}+\mathsf{B})\lambda} = e^{\mathsf{A}\lambda} e^{\mathsf{B}\lambda} e^{\mathsf{C}_2\lambda^2} e^{\mathsf{C}_3\lambda^3} \cdots$$

Use this in combination with $e^{\mathsf{M}\lambda} = \big[e^{\mathsf{M}\frac{\lambda}{N}}\big]^N$ to obtain

$$\begin{aligned} e^{(\mathsf{A}+\mathsf{B})\lambda} &= \big[e^{\mathsf{A}\frac{\lambda}{N}} e^{\mathsf{B}\frac{\lambda}{N}} e^{\mathsf{C}_2(\frac{\lambda}{N})^2} e^{\mathsf{C}_3(\frac{\lambda}{N})^3} \cdots\big]^N \\ &= \big[e^{\mathsf{A}\frac{\lambda}{N}} e^{\mathsf{B}\frac{\lambda}{N}} \big(\mathsf{I} + \text{terms of order } (\tfrac{\lambda}{N})^2\big)\big]^N \\ &\downarrow \\ &= \big[e^{\mathsf{A}\frac{\lambda}{N}} e^{\mathsf{B}\frac{\lambda}{N}}\big]^N \quad \text{as } N \to \infty \end{aligned} \tag{47}$$

which has become known as the "Trotter product formula."[31] One might, on this basis, write

$$\begin{aligned} e^{-\frac{i}{\hbar}(\frac{1}{2m}\mathsf{p}^2+\mathsf{V})t} &= \big[e^{-\frac{i}{\hbar}\frac{1}{2m}\mathsf{p}^2\tau} \cdot e^{-\frac{i}{\hbar}V(\mathsf{x})\tau}\big]^N \quad \text{with } \tau \equiv t/N \\ &= {}_{\mathsf{x}}\big[e^{-\frac{i}{\hbar}H(x,p)\tau}\big]_{\mathsf{p}} \cdot {}_{\mathsf{x}}\big[e^{-\frac{i}{\hbar}H(x,p)\tau}\big]_{\mathsf{p}} \cdots {}_{\mathsf{x}}\big[e^{-\frac{i}{\hbar}H(x,p)\tau}\big]_{\mathsf{p}} \end{aligned}$$

---

[31] The original reference (almost never cited) is Hale F. Trotter,"On the product of semi-groups of operators," Proc. Amer. Math. Soc. **10**, 545 (1959). Trotter wrote (at Princeton) in the formally "mathematical" style that most physicists find off-putting, though he took his motivation from problems having to do with the numerical solution of partial differential equations. More accessible is Masuo Suzuki, "Generalized Trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems," Comm. Math. Phys. **51**, 183 (1976). Suzuki remarks, by the way, that the Zassenhaus $\mathsf{C}$'s can be computed recursively from

$$\begin{aligned} \mathsf{C}_2 &= \tfrac{1}{2!}\Big[\tfrac{\partial}{\partial\lambda}\big(e^{-\mathsf{B}\lambda}e^{-\mathsf{A}\lambda}e^{(\mathsf{A}+\mathsf{B})\lambda}\big)\Big]_{\lambda=0} \\ \mathsf{C}_3 &= \tfrac{1}{3!}\Big[\tfrac{\partial}{\partial\lambda}\big(e^{-\mathsf{C}_2\lambda^2}e^{-\mathsf{B}\lambda}e^{-\mathsf{A}\lambda}e^{(\mathsf{A}+\mathsf{B})\lambda}\big)\Big]_{\lambda=0} \\ \mathsf{C}_4 &= \tfrac{1}{4!}\Big[\tfrac{\partial}{\partial\lambda}\big(e^{-\mathsf{C}_3\lambda^3}e^{-\mathsf{C}_2\lambda^2}e^{-\mathsf{B}\lambda}e^{-\mathsf{A}\lambda}e^{(\mathsf{A}+\mathsf{B})\lambda}\big)\Big]_{\lambda=0} \\ &\vdots \end{aligned}$$

The brief account of Trotter's formula presented as an appendix to Chapter I in Schulman[12] appears to have been adapted from the appendix to E. Nelson, "Feynman integrals & the Schrödinger equation," J. Math. Phys. **5**, 332 (1964).

Hit the final product with

- $(x|$ on the left,
- $|y)$ on the right, and
- insert a copy of $\iint |p')dp'(p'|x')dx'(x'|$ between each of the factors

and we are led back to a statement of the design (12) that marked our point of departure. The implied assertion is that we can <u>abandon the distinction between $\mathcal{H}$ and $H$</u>; *i.e.*, that in place of (40) we can write

$$\begin{aligned} K(x,\tau;y,0) &= \tfrac{1}{h}\int \exp\Big\{\tfrac{i}{\hbar}g(p;x,y,\tau)\Big\}\,dp \\ &\qquad\qquad g(p;x,y,\tau) \equiv p\cdot(x-y) - H(x,p,\tau)\,\tau \\ &= \sqrt{\tfrac{m}{iht}}\,\exp\Big\{\tfrac{i}{\hbar}S(x,\tau;y,0)\Big\} \qquad (48.1) \\ &\qquad\qquad S(x,\tau;y,0) \equiv \tfrac{m}{2}\tfrac{(x-y)^2}{\tau} - V(x)\,\tau \qquad (48.2) \end{aligned}$$

Equations (48), for all their elegant simplicity, might seem like mere wishful thinking, since (compare (33) and (34))

$$\tfrac{1}{2m}S_x^2 + V + S_\tau = \underbrace{\tfrac{1}{2m}V_x\Big\{-2m(x-y) + \tau^2 V_x\Big\}}_{\text{\underline{not} of the form } 0 + O(\tau^2)}$$

and (which is perhaps more to the point)

$$-\tfrac{\hbar^2}{2m}K_{xx} + VK - i\hbar K_\tau = \underbrace{K\cdot\tfrac{1}{2m\sqrt{2\pi}}\Big\{-2m(x-y)V_x + i\hbar V_{xx}\tau + V_x^2\tau^2\Big\}}_{\text{same criticism}}$$

Those criticisms notwithstanding, Feynman proposes to set

$$S_{\text{Feynman}}(x,\tau;y,0) = \begin{cases} \Big\{\tfrac{1}{2}L\big(\tfrac{x-y}{\tau},x\big) + \tfrac{1}{2}L\big(\tfrac{x-y}{\tau},y\big)\Big\}\tau & \text{else} \\ L\big(\tfrac{x-y}{\tau},\tfrac{x+y}{2}\big)\,\tau & \text{else (as above)} \\ L\big(\tfrac{x-y}{\tau},x\big)\,\tau \end{cases} \qquad (49)$$

depending on the specific design of the Lagrangian $L(\dot{x},x)$, the selection to be made case by case, subject to the pragmatic criterion implicit in this question: *Does the selection lead to the "correct" Schrödinger equation*? We cannot object to Feynman's adoption of such a criterion if his objective is simply share the "pleasure in recognizing old things from a new point of view," but so long as it remains in place "Feynman quantization" loses any claim to conceptual autonomy . . . though it was put forward in "the hope that the new point of view will inspire an idea for the modification of present theories, a modification necessary to encompass present experiments."

To gain a better sense of what Feynman's proposals (49) entail, let us back up to (28)

$$S(x,\tau;y,0) = \tfrac{m}{2}\tfrac{(x-y)^2}{\tau} - \int_0^\tau V\big(y + \frac{x-y}{\tau}\,t'\big)\,dt'$$

and observe that simple adjustment of the variable of integration supplies

$$\begin{aligned}&= \tfrac{m}{2}\tfrac{(x-y)^2}{\tau} - \tfrac{\tau}{x-y}\int_y^x V(x')\,dx'\\ &= \tfrac{m}{2}\tfrac{(x-y)^2}{\tau} - V\big(\mu(x,y)\big)\,\tau\end{aligned}$$

where $\mu(x,y)$ marks the point at which the potential assumes (with respect to the interval $[x,y]$ ) its *mean value*. In this notation (33.1) becomes

$$\begin{aligned}&\tfrac{1}{2m}S_x^2 + V(x) + S_\tau \qquad (50)\\ &\quad= \Big\{V(x) - V(\mu) - (x-y)V'(\mu)\mu_x\Big\} + \tfrac{1}{2m}\Big[V'(\mu)\mu_x\Big]^2\cdot\tau^2\end{aligned}$$

where the definition of $\mu(x,y)$ entails

$$\Big\{\text{etc.}\Big\} = V(x) - V(\mu) - (x-y)\cdot\tfrac{d}{dx}\Big[\tfrac{1}{x-y}\int_y^x V(x')\,dx'\Big] = 0$$

Feynman proposes to *abandon* that definition: in its place he would

$$\begin{aligned}&\text{define } \mu(x,y) \text{ by } V(\mu) = \tfrac{1}{2}\big[V(x)+V(y)\big], &&\text{else}\\ &\text{simply } \textit{set} \quad \mu(x,y) = \tfrac{1}{2}(x+y), &&\text{else}\\ &\text{simply } \textit{set} \quad \mu(x,y) = \quad x\end{aligned}$$

and live with the fact that $\big\{\text{etc.}\big\} \neq 0$. That he enjoys any success at all is surprising. Let's see how he does it:

**Origin of the Schrödinger equation, according to Feynman.** Feynman would have us write

$$\begin{aligned}K(x,t+\tau;x_0,t_0) &= K + \tau K_t + \tfrac{1}{2}\tau^2 K_{tt} + \cdots\\ &= \int K(x,t+\tau;y,t)\,dy\,K(y,t;x_0,t_0)\\ &= \int \tfrac{1}{A(\tau)}\exp\Big\{\tfrac{i}{\hbar}\big[\tfrac{m}{2}\tfrac{(x-y)^2}{\tau} - \tau V(\mu) + \cdots\big]\Big\}K(y,t;x_0,t_0)\,dy\\ &= \int \tfrac{1}{A(\tau)}\exp\Big\{\tfrac{i}{\hbar}\big[\tfrac{m}{2}\tfrac{(x-y)^2}{\tau}\big]\Big\}\Big\{1 - \tfrac{i}{\hbar}\tau V(\mu) + \cdots\Big\}K(y,t;x_0,t_0)\,dy\end{aligned}$$

on the basis of which we expect to have

$$\begin{aligned}K &= \Big[\quad\int \tfrac{1}{A(\tau)}\exp\Big\{\tfrac{i}{\hbar}\big[\tfrac{m}{2}\tfrac{(x-y)^2}{\tau}\big]\Big\}\Big\{1 - \tfrac{i}{\hbar}\tau V(\mu) + \cdots\Big]\Big\}K(y,t;x_0,t_0)\,dy\Big]_{\tau\downarrow 0}\\ K_t &= \Big[\tfrac{\partial}{\partial\tau}\int \tfrac{1}{A(\tau)}\exp\Big\{\tfrac{i}{\hbar}\big[\tfrac{m}{2}\tfrac{(x-y)^2}{\tau}\big]\Big\}\Big\{1 - \tfrac{i}{\hbar}\tau V(\mu) + \cdots\Big]\Big\}K(y,t;x_0,t_0)\,dy\Big]_{\tau\downarrow 0}\\ &\vdots\end{aligned}$$

where the blue terms play no role in subsequent analysis, and will henceforth be dropped. The idea now is to exploit the Gaussian representations of the $\delta$-function and its derivatives, as summarized at (100) in Chapter 0. To that end, introduce $\epsilon = \sqrt{i\hbar\tau/m}$ so as to achieve $\frac{i}{\hbar}\frac{m}{2\tau} = -\frac{1}{2\epsilon^2}$. The first of the preceding equations then becomes

$$\begin{aligned} K(x,t;x_0,t_0) &= \Big[\int \tfrac{1}{B(\epsilon)}\exp\Big\{-\tfrac{1}{2}\big[\tfrac{x-y}{\epsilon}\big]^2\Big\}K(y,t;x_0,t_0)\,dy\Big]_{\epsilon\downarrow 0} \\ &= \int \delta(x-y)K(y,t;x_0,t_0)\,dy \end{aligned}$$

provided we set[32]

$$\tfrac{1}{B(\epsilon)} = \tfrac{1}{\sqrt{2\pi}}\big(\tfrac{1}{\epsilon}\big)^1 \quad \text{which entails} \quad \tfrac{1}{A(\tau)} = \sqrt{\tfrac{m}{2\pi i\hbar\tau}}$$

In next higher order we have

$$\begin{aligned} K_t &= \tfrac{i\hbar}{2m\epsilon}\Big[\tfrac{\partial}{\partial\epsilon}\int \tfrac{1}{B(\epsilon)}\exp\Big\{-\tfrac{1}{2}\big[\tfrac{x-y}{\epsilon}\big]^2\Big\}\Big\{1-\tfrac{m}{\hbar^2}V(\mu)\,\epsilon^2\Big\}K(y,t;x_0,t_0)\,dy\Big]_{\epsilon\downarrow 0} \\ &= \tfrac{i\hbar}{2m}\Big[\int \tfrac{1}{\sqrt{2\pi}}\exp\Big\{-\tfrac{1}{2}\big[\tfrac{x-y}{\epsilon}\big]^2\Big\}\Big\{\big(\tfrac{1}{\epsilon}\big)^3\big[\big(\tfrac{x-y}{\epsilon}\big)^2\big(1-\tfrac{2m}{\hbar^2}V(\mu)\,\epsilon^2\big)-1\big] \\ &\qquad\qquad -\big(\tfrac{1}{\epsilon}\big)^1\tfrac{2m}{\hbar^2}V(\mu)\Big\}K(y,t;x_0,t_0)\,dy\Big]_{\epsilon\downarrow 0} \\ &= \tfrac{1}{i\hbar}\int\Big\{-\tfrac{\hbar^2}{2m}\delta''(x-y)+V\big(\mu(x,y)\big)\delta(x-y)\Big\}K(y,t;x_0,t_0)\,dy \end{aligned}$$

Thus do we obtain

$$i\hbar K_t(x,t;x_0,t_0) = \Big\{-\tfrac{\hbar^2}{2m}\tfrac{\partial^2}{\partial x^2}+V\big(\mu(x,x)\big)\Big\}K(x,t;x_0,t_0)$$

which—since

$$\mu(x,x) = x \quad \text{by every plausible construction of } \mu(x,y)$$

—is precisely the Schrödinger equation. Which, as we are in a position now to appreciate, we would have missed had we followed Feynman's initial impulse, which was to set $A = \text{constant}$. Feynman reportedly raced through the argument too fast for Jehle to follow or transcribe. The reader who takes the trouble to schlog through the details ("a small nightmare of Taylor expansions and Gaussian integrals" in the view of Schulman, but not really so bad if carefully managed) is certain to be rewarded by some sense of the excitement which Feynman and Jehle felt on that occasion.[33]

---

[32] We see here how Feynman adjusts his "normalization factors" $\frac{1}{A}$ *after the fact*, in order to make things work out right.

[33] For an informative account of the circumstances surrounding Feynman's own first extraction of the Schrödinger equation from the path-integral idea, see D. Derbes, "Feynman's derivation of the Schrödinger equation," AJP **64**, 881 (1996). The argument presented here—which departs organizationally from Feynman's—was taken from QUANTUM MECHANICS (1967), Chapter 1, page 77.

Feynman based his construction on an approximation to $S(x,\tau;y,0)$ which —though recommended (17 years after the fact!) by Trotter's formula—is so crude that it fails to extinquish the $0^{\text{th}}$-order $\{\text{etc.}\}$-term that appears on the right side of (50). It has become clear that Feynman's surprising success can be attributed to the fact that $\mu(x,y)$ enters into his final equation only as $\mu(x,x) = x$, and $\{\text{etc.}\}$ *does* vanish at $y = x$.

**Extension to more general dynamical systems.** Let $\{x^1, x^2, x^3\}$ refer to an inertial Cartesian frame in physical 3-space, where a particle $m$ moves subject to the conservative forces that arise from the potential $V(\boldsymbol{x})$. To describe the classical motion of the particle we write

$$L(\boldsymbol{x},\dot{\boldsymbol{x}}) = \tfrac{1}{2}m\sum_{k=1}^{3}\dot{x}^k\dot{x}^k - V(\boldsymbol{x})$$

The associated Hamiltonian is

$$H(\boldsymbol{x},\boldsymbol{p}) = \tfrac{1}{2m}\sum_{k=1}^{3}p_k p_k + V(\boldsymbol{x})$$

Working in the uniform rectilinear short-time approximation

$$S(\boldsymbol{x},\tau;\boldsymbol{y},0) = \tfrac{m}{2}\sum_{k=1}^{3}\frac{(x^k - y^k)^2}{\tau} - V(\boldsymbol{\mu}(\boldsymbol{x},\boldsymbol{y}))\,\tau$$

(or even in the crude approximation $\boldsymbol{\mu} = \boldsymbol{x}$), we construct

$$K_C(\boldsymbol{x},\tau;\boldsymbol{y},0) = \sqrt{(\tfrac{1}{ih})^3 D}\,\exp\left\{\tfrac{i}{\hbar}S\right\}$$

and find by the argument already rehearsed that

$$\psi(\boldsymbol{x},t) \equiv \int K_C(\boldsymbol{x},t;\boldsymbol{y},t-\tau)\psi(\boldsymbol{y},t-\tau)d^3y \quad \text{satisfies} \quad \left\{-\tfrac{\hbar}{2m}\nabla^2 + V\right\}\psi = i\hbar\partial_t\psi$$

No sweat, no surprise.

But in 3-dimensional work we often find it convenient to employ *curvilinear coordinates* (which classically, in the presence of holonomic constraints, become "generalized coordinates"). If (moving) curvilinear coordinates are introduced by equations of the form

$$x^k = x^k(q^1,q^2,q^3,t) \quad : \quad k = 1,2,3$$

then the Lagrangian becomes

$$L = \tfrac{1}{2}m\Big\{\sum_{i,j}g_{ij}\dot{q}^i\dot{q}^j + 2\sum_i a_i\dot{q}^i + b\Big\} - U$$

with

$$g_{ij}(q,t) \equiv \sum_k \frac{\partial x^k}{\partial q^i}\frac{\partial x^k}{\partial q^j}\,, \quad a_i(q,t) \equiv \sum_k \frac{\partial x^k}{\partial q^i}\frac{\partial x^k}{\partial t}\,, \quad b(q,t) \equiv \sum_k \frac{\partial x^k}{\partial t}\frac{\partial x^k}{\partial t}$$

and $U(q,t) \equiv V(x(q,t))$. The Hamilton assumes therefore the design

$$\begin{aligned} H(q,p,t) &= \sum_k p_k\dot{q}^k - L(q,\dot{q}) \quad \text{with} \quad p_i = m\sum_j g_{ij}\dot{q}^j + ma_i \\ &= \tfrac{1}{2m}\sum_{i,j} g^{ij}[p_i - ma_i][p_j - ma_j] + (U - \tfrac{1}{2}mb) \end{aligned} \tag{51}$$

If, as is most commonly the case, the $q$-coordinate system is not itself in motion with respect to our inertial frame then we have this simplification:

$$\downarrow$$

$$H(q,p) = \tfrac{1}{2m}\sum_{i,j} g^{ij}(q)p_ip_j + U(q)$$

So far, so good. But when we attempt to make the formal substitutions

$$q^i \mapsto \mathsf{q}^i \quad \text{and} \quad p_i \mapsto \mathsf{p}_i$$

required to construct the corresponding Hamiltonian operator $\mathsf{H}$ we confront (except in cases where $g^{ij}$ and $a_i$ are constants) an *operator ordering ambiguity*, which becomes especially severe if the $g^{ij}(q)$ refer not—as above—to the Euclidean metric structure of physical 3-space but to the metric structure of some curved manifold upon which we are attempting to write quantum mechanics. For this and other reasons—we have lost our former description of $S(q,\tau;q_0,0)$, and possess no theory of Fourier transformations or of Gaussian integration with respect to non-Cartesian coordinates—it would be premature to pursue the path-integral formalism into this particular jungle until we have gained a better sense issues involved . . . and that is an intricate story which I will reserve for another occasion.[34] I will, however, look to a somewhat attenuated instance of some related issues:

To describe—relative to an intertial Cartesian frame—the motion of a charged mass point in the presence of an impressed electromagnetic we write

$$\begin{aligned} L &= \tfrac{1}{2}m\dot{\boldsymbol{x}}\cdot\dot{\boldsymbol{x}} - V(\boldsymbol{x}) + \tfrac{e}{c}\boldsymbol{A}(\boldsymbol{x})\cdot\dot{\boldsymbol{x}} \\ V(\boldsymbol{x}) &\equiv e\phi(\boldsymbol{x}) \end{aligned} \tag{52}$$

where the interesting new feature is the term linear in $\dot{\boldsymbol{x}}$ with variable coefficient. The associated Hamiltonian reads (compare (51))

$$\begin{aligned} H &= \tfrac{1}{2m}[\boldsymbol{p} - \tfrac{e}{c}\boldsymbol{A}]\cdot[\boldsymbol{p} - \tfrac{e}{c}\boldsymbol{A}] + V \\ &= \sum_k \Big\{\tfrac{1}{2m}p_k^2 - \tfrac{e}{mc}p_kA_k + \tfrac{e^2}{2mc^2}A_k^2\Big\} + V \end{aligned}$$

[34] In the meantime, see (for example) Chapter 24 in Schulman.[12]

When we undertake to construct the associated $\mathsf{H}$-operator we confront an ordering problem which, however, in this simple instance "solves itself" in the sense that it is entirely natural to write

$$p_k A_k \mapsto \tfrac{1}{2}\big\{\,\mathsf{p}_k A_k(\mathsf{x}) + A_k(x)\,\mathsf{p}_k\big\}$$

We are led thus to

$$\mathsf{H} = \tfrac{1}{2m}\sum_k [\,\mathsf{p}_k - \tfrac{e}{c}\mathsf{A}_k]\!\cdot\![\,\mathsf{p}_k - \tfrac{e}{c}\mathsf{A}_k] + \mathsf{V}$$

and to the Schrödinger equation

$$\Big\{\tfrac{1}{2m}\sum_k [\,\tfrac{\hbar}{i}\partial_k - \tfrac{e}{c}A_k(x)]\cdot[\,\tfrac{\hbar}{i}\partial_k - \tfrac{e}{c}A_k(x)] + V(x)\Big\}\psi = i\hbar\partial_t\psi \tag{53}$$

This is the equation we undertake now to extract from the sum-over-paths formalism. My argument this time adheres closely to the pattern of Feynman's.[35] As a notational convenience I work in one dimension.

Work from

$$\begin{aligned}\psi(x,t+\tau) &= \psi(x,t) + \tau\,\psi_t(x,t) + \cdots\\ &= \int \underbrace{K_C(x,\tau;y,0)}_{=\sqrt{\frac{1}{2\pi}\beta D}\,e^{-\beta S}}\psi(y,t)\,dy\end{aligned}$$

with $\beta \equiv 1/i\hbar$ and $S(x,\tau;y,0) = \frac{m}{2}\frac{(x-y)^2}{\tau} - V(x)\,\tau + \frac{e}{c}\int_0^\tau A(x)\,\dot{x}\,dt$ (which entails $D = m/\tau$). Using $\int_0^\tau A(x)\,\dot{x}\,dt = \int_y^x A(z)\,dz$, we have

$$= e^{\beta V(x)\tau}\!\int \sqrt{\tfrac{m}{2\pi\tau}\beta}\,\exp\Big\{-\beta\tfrac{m}{2}\tfrac{(y-x)^2}{\tau}\Big\}\,\exp\Big\{-\beta\,\tfrac{e}{c}\!\int_y^x A(z)dz\Big\}\psi(y,t)\,dy$$

Owing to the presence of the red Gaussian (which becomes ever more sharply peaked as $\tau\downarrow 0$) we can interpret $\xi \equiv y - x$ to be small, and on that basis can write

$$\begin{aligned}\int_y^x A(z)dz &= (x-y)\cdot\frac{A(x)+A(y)}{2} \quad \text{by the trapazoidal rule}\\ &= -\tfrac{1}{2}\xi\cdot\big[A(x)+A(x+\xi)\big]\\ &= -A(x)\,\xi - \tfrac{1}{2}A_x(x)\,\xi^2 + \cdots\end{aligned}$$

$$\psi(y,t) = \psi(x,t) + \psi_x(x,t)\,\xi + \tfrac{1}{2}\psi_{xx}(x,t)\,\xi^2 + \cdots$$

---

[35] See Chapter 4 in Schulman.[12] I must confess that I have been unable to make my own former line of argument (which used Gaussian representations of the derivatives of the $\delta$ function) work in the present context.

Putting the pieces together, we have

$$\begin{aligned}\psi(x,t)+\tau\psi_t(x,t)+\cdots=\int&\sqrt{\tfrac{m}{2\pi\tau}\beta}\exp\Big\{-\beta\tfrac{m}{2}\tfrac{\xi^2}{\tau}\Big\}\\ &\cdot\Big[1+\beta V(x)\tau+\cdots\Big]\\ &\cdot\Big[1+\quad\big(\beta\tfrac{e}{c}A(x)\xi+\beta\tfrac{1}{2}\tfrac{e}{c}A_x(x)\xi^2\big)\\ &\qquad+\tfrac{1}{2!}\big(\beta\tfrac{e}{c}A(x)\xi+\qquad\qquad\big)^2+\cdots\Big]\\ &\cdot\Big[\psi(x,t)+\psi_x(x,t)\xi+\tfrac{1}{2}\psi_{xx}(x,t)\xi^2+\cdots\Big]d\xi\end{aligned}$$

But

$$\begin{aligned}\int_{-\infty}^{+\infty}\sqrt{\tfrac{m}{2\pi\tau}\beta}\exp\Big\{-\beta\tfrac{m}{2}\tfrac{\xi^2}{\tau}\Big\}\xi^0d\xi&=1\\ \int_{-\infty}^{+\infty}\sqrt{\tfrac{m}{2\pi\tau}\beta}\exp\Big\{-\beta\tfrac{m}{2}\tfrac{\xi^2}{\tau}\Big\}\xi^1d\xi&=0\\ \int_{-\infty}^{+\infty}\sqrt{\tfrac{m}{2\pi\tau}\beta}\exp\Big\{-\beta\tfrac{m}{2}\tfrac{\xi^2}{\tau}\Big\}\xi^2d\xi&=\tfrac{1}{\beta m}\tau\end{aligned}$$

so in order $O(\tau^0)$ we have the reassuring triviality $\psi(x,t)=\psi(x,t)$ while in $O(\tau^1)$ we have

$$\psi_t=\beta\Big\{\tfrac{1}{2m\beta^2}\psi_{xx}+\tfrac{e}{mc\beta}A\psi_x+\tfrac{e}{2mc\beta}A_x\psi+\tfrac{e^2}{2mc^2}A^2\psi+V\psi\Big\}$$

which can be written

$$\Big\{\tfrac{1}{2m}\big[\tfrac{\hbar}{i}\partial_x-\tfrac{e}{c}A\big]^2+V\Big\}\psi=i\hbar\partial_t\psi$$

The 3-dimensional argument proceeds in exactly the same way.

Several comments are now in order: *Mathematica* informs us that

$$\int_{-\infty}^{+\infty}\sqrt{\tfrac{m}{2\pi\tau}\beta}\exp\Big\{-\beta\tfrac{m}{2}\tfrac{\xi^2}{\tau}\Big\}\xi^nd\xi=[1+(-1)^n]\tfrac{1}{\sqrt{2\pi}}2^{\frac{1}{2}(n-1)}\Gamma\big(\tfrac{n+1}{2}\big)\big(\tfrac{1}{\beta m}\tau\big)^{\frac{1}{2}n}$$

of which we have made use especially of the case $n=2$. The general point is that *Gaussian integration serves to convert power series in $\xi^2$ into power series in $\tau$.* Secondly, we are in position now to understand the casual "or, if it proves more convenient" with which Feynman asserts the effective equivalence of

$$\begin{aligned}\text{TRAPAZOIDAL RULE}\quad:&\quad\int_0^\tau V(x(t))\,dt\approx\tfrac{1}{2}\big[V(x)+V(y)\big]\tau\\ \text{MIDPOINT RULE}\quad:&\quad\int_0^\tau V(x(t))\,dt\approx\big[V(\tfrac{x+y}{2})\big]\tau\end{aligned}$$

for the power series that result from setting $y=x+\xi$ differ only in $O(\xi^2)$.

**Gauge transformations, compensating fields, impressed magnetic fields and the Aharonov-Bohm effect.** I digress now to discuss some of the remarkable ramifications of the seemingly innocuous adjustment

$$L_0 \quad\longrightarrow\quad L \equiv L_0 + \tfrac{e}{c}\boldsymbol{A}(\boldsymbol{x})\boldsymbol{\cdot}\dot{\boldsymbol{x}} \tag{54}$$

where $$L_0 \equiv \tfrac{1}{2}m\dot{\boldsymbol{x}}\boldsymbol{\cdot}\dot{\boldsymbol{x}} - V(\boldsymbol{x})$$

and where $\frac{e}{c}$ is the physically-motivated name given to the coupling constant that describes the strength of the new $\dot{\boldsymbol{x}}$-linear term.

In one dimension (54) becomes

$$L_0 \;\longrightarrow\; L = \underbrace{\tfrac{1}{2}m\dot{x}^2 - V(x)}_{L_0} + \tfrac{e}{c}A(x)\,\dot{x}$$

and we notice that the new term is a *gauge* term

$$\tfrac{e}{c}A(x)\,\dot{x} = \tfrac{d}{dt}\Big\{\tfrac{e}{c}\int^x A(z)\,dz\Big\} \tag{55}$$

so contributes nothing to the classical motion. Its inclusion does, however, entail modification

$$p \equiv m\dot{x} \;\longrightarrow\; p \equiv m\dot{x} + \tfrac{e}{c}A(x) \tag{56}$$

of the equation that defines the "conjugate momentum," and therefore induces an adjustment in the design of the Hamiltonian:

$$H_0 = \tfrac{1}{2m}p^2 + V(x) \;\longrightarrow\; H = \tfrac{1}{2m}\big[p - \tfrac{e}{c}A(x)\big]^2 + V(x) \tag{57}$$

The dynamical action responds additively to gauge transformation

$$S_0(x,t;y,0) \;\longrightarrow\; S(x,t;y,0) = S_0(x,t;y,0) + \tfrac{e}{c}\int_y^x A(z)\,dz \tag{58}$$

so the semi-classical propagator $K_0 \equiv \sqrt{\frac{i}{\hbar}\partial^2 S_0/\partial x\partial y}\cdot e^{\frac{i}{\hbar}S_0}$ acquires a factor:

$$K_0 \;\longrightarrow\; K = K_0\cdot\exp\Big\{i\tfrac{e}{\hbar c}\int_y^x A(z)\,dz\Big\} \tag{59}$$

The equations

$$\Big\{\tfrac{1}{2m}\big(\tfrac{\hbar}{i}\partial_x\big)^2 + V(x)\Big\}K_0 = i\hbar\partial_t K_0$$

and

$$\Big\{\tfrac{1}{2m}[\tfrac{\hbar}{i}\partial_x - \tfrac{e}{c}A(x)]^2 + V(x)\Big\}K \;= i\hbar\partial_t K$$

echo the relationship between $H_0$ and $H$, and their equivalence is an immediate consequence of the "shift rule"

$$\tfrac{\hbar}{i}\partial_x \exp\Big\{-i\tfrac{e}{\hbar c}\int_y^x A(z)\,dz\Big\}\bullet = \exp\Big\{-i\tfrac{e}{\hbar c}\int_y^x A(z)\,dz\Big\}[\tfrac{\hbar}{i}\partial_x - \tfrac{e}{c}A(x)]^2\bullet$$

We have been brought into contact here with an idea that lies at the base of the theory of compensating (or "gauge") fields: if $\psi_0$ satisfies

$$\Big\{\tfrac{1}{2m}\big(\tfrac{\hbar}{i}\partial_x\big)^2 + V(x)\Big\}\psi_0 = i\hbar\partial_t\psi_0 \tag{59}$$

then so also does $\psi \equiv e^{i(e/\hbar c)\chi}\cdot\psi_0$ if $\chi$ is any (real) constant. The design of (59) is, in this sense, "gauge invariant." Gauge invariance is, however, lost if we allow $\chi$ to become $x$-dependent, for as we have seen

$$\psi = e^{i(e/\hbar c)\chi(x)}\cdot\psi_0$$

satisfies

$$\Big\{\tfrac{1}{2m}\big[\tfrac{\hbar}{i}\partial_x - \tfrac{e}{c}\chi'(x)\big]^2 + V(x)\Big\}\psi = i\hbar\partial_t\psi$$

The idea is—in place of (59)—to write

$$\Big\{\tfrac{1}{2m}\big[\tfrac{\hbar}{i}\partial_x - \tfrac{e}{c}\underbrace{A_0(x)}_{\text{"compensating field"}}\big]^2 + V(x)\Big\}\psi_0 = i\hbar\partial_t\psi_0 \tag{60}$$

and to assign to "gauge transformation" this expanded meaning:

$$\left.\begin{aligned} \psi_0 &\longrightarrow \psi = e^{i(e/\hbar c)\chi(x)}\cdot\psi_0 \\ A_0 &\longrightarrow A = A_0 + \chi'(x) \end{aligned}\right\} \tag{61}$$

Then (60) *is* gauge invariant in the sense that under (61) it goes over into an equation of the same design:

$$\Big\{\tfrac{1}{2m}\big[\tfrac{\hbar}{i}\partial_x - \tfrac{e}{c}A(x)\big]^2 + V(x)\Big\}\psi = i\hbar\partial_t\psi$$

Pushed only a little farther, the idea leads spontaneously to the "invention" of Maxwellian electrodynamics.

But when we write $L = \frac{1}{2}m\dot{\boldsymbol{x}}\cdot\dot{\boldsymbol{x}} - V(\boldsymbol{x}) + \frac{e}{c}\boldsymbol{A}(\boldsymbol{x})\cdot\dot{\boldsymbol{x}}$ or

$$\Big\{\tfrac{1}{2m}\big[\tfrac{\hbar}{i}\boldsymbol{\nabla} - \tfrac{e}{c}\boldsymbol{A}(\boldsymbol{x})\big]^2 + V(\boldsymbol{x})\Big\}\psi = i\hbar\partial_t\psi \tag{62}$$

we imagine ourselves to be coming *from* electrodynamics, where $\boldsymbol{A}\rightarrow\boldsymbol{A}+\boldsymbol{\nabla}\chi$ is already in place (was inherited from $\boldsymbol{B} = \boldsymbol{\nabla}\times\boldsymbol{A}$); the gauge invariance of (61) is achieved by adopting this quantum mechanical enlargement

$$\left.\begin{aligned} \psi_0 &\longrightarrow \psi = e^{i(e/\hbar c)\chi(x)}\cdot\psi_0 \\ \boldsymbol{A}_0 &\longrightarrow \boldsymbol{A} = \boldsymbol{A}_0 + \boldsymbol{\nabla}\chi(x) \end{aligned}\right\} \tag{63}$$

of the classical notion of an "electromagnetic gauge transformation."[36] When we compare (61) with (63) we see that the locus of the novelty has switched places.

Some aspects of my present subject are a bit slippery, and it is to get a firmer classical/quantum mechanical grip upon them that I look now to this concrete

EXAMPLE: CHARGED PARTICLE IN HOMOGENEOUS MAGNETIC FIELD Let us take the vector potential $\boldsymbol{A}$ to be given by

$$\boldsymbol{A} = \tfrac{1}{2}B\begin{pmatrix} -y \\ +x \\ 0 \end{pmatrix} \quad : \quad \text{then} \quad \boldsymbol{B} = \boldsymbol{\nabla}\times\boldsymbol{A} = \begin{pmatrix} 0 \\ 0 \\ B \end{pmatrix}$$

describes a homogeneous magnetic field parallel to the $z$-axis. The Lagrangian (52/54) has, in the assumed absence of a potential $V$, become

$$L = \tfrac{1}{2}m\Big\{(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + 2\omega(-y\dot{x} + x\dot{y})\Big\} \qquad (64)$$
$$\omega \equiv Be/2mc$$

and the equations of motion read

$$\begin{aligned} \ddot{x} - 2\omega\dot{y} &= 0 \\ \ddot{y} + 2\omega\dot{x} &= 0 \\ \ddot{z} &= 0 \end{aligned}$$

*Mathematica* informs us that the solution which links $\{\boldsymbol{x}_1, t_1\} \leftarrow \{\boldsymbol{x}_0, 0\}$ can be described

$$\left.\begin{aligned} x(t) &= x_0 + \tfrac{1}{2}\big\{ + (x_1 - x_0) + (y_1 - y_0)\cot\omega t_1\big\}(1 - \cos 2\omega t) \\ &\qquad + \tfrac{1}{2}\big\{ - (y_1 - y_0) + (x_1 - x_0)\cot\omega t_1\big\}\sin 2\omega t \\ y(t) &= y_0 + \tfrac{1}{2}\big\{ + (y_1 - y_0) - (x_1 - x_0)\cot\omega t_1\big\}(1 - \cos 2\omega t) \\ &\qquad + \tfrac{1}{2}\big\{ + (x_1 - x_0) + (y_1 - y_0)\cot\omega t_1\big\}\sin 2\omega t \\ z(t) &= z_0 + \big\{(z_1 - z_0)/t_1\big\}t \end{aligned}\right\} \qquad (65)$$

The $z$-motion is an uninteresting unaccelerated drift: I excise it from the discussion by setting $z_0 = z_1 = 0$. To render the remaining equations more discussably transparent I place the endpoints in "standard position"

$$\boldsymbol{x}_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{x}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

---

[36] A similar enlargment is required to fit the classical concept of a "Galilean boost"

$$\boldsymbol{x}_0 \;\longrightarrow\; \boldsymbol{x} = \boldsymbol{x}_0 - \boldsymbol{v}t$$

into the non-relativistic quantum mechanical framework. For a good discussion of the point see F. A. Kaempffer, *Concepts in Quantum Mechanics* (1965), Appendix 7: "If Galileo had known quantum mechanics."

Then

$$x(t) = \tfrac{1}{2}\cot\omega t_1 \cdot (1 - \cos 2\omega t) - \tfrac{1}{2}\sin 2\omega t$$
$$y(t) = \tfrac{1}{2}(1 - \cos 2\omega t) + \tfrac{1}{2}\cot\omega t_1 \cdot \sin 2\omega t$$

Solve for $\cos\omega t$ and $\sin\omega t$. Use $\cos^2\omega t + \sin^2\omega t = 1$ to obtain the orbital equation

$$\big(x - \tfrac{1}{2}\cot\omega t_1\big)^2 + \big(y - \tfrac{1}{2}\big)^2 = \big(\tfrac{1}{2\sin\omega t_1}\big)^2$$

The orbit is a circle, centered at

$$x_\odot(t_1) = \tfrac{\cos\omega t_1}{2\sin\omega t_1} \quad \text{and} \quad y_\odot = \tfrac{1}{2}$$

with radius

$$R(t_1) = \big|\tfrac{1}{2\sin\omega t_1}\big|$$

The radius is a periodic function of the arrival time

$$R(t_1) = R(t_1 + \tau) \quad : \quad \tau \equiv 2\pi/\omega$$

and becomes infinite when $t_1 = n\tau : n = 0, 1, 2, \ldots$ The particle advances clockwise around the circle with constant angular velocity $\Omega \equiv 2\omega$. Its linear speed is $v = \Omega R$, so we have

$$\begin{aligned}
\text{conserved angular momentum } \mathcal{L} &= 2m\omega R^2 \\
&= \tfrac{m\omega}{2\sin^2\omega t_1} \\
&= \tfrac{2m\omega}{\pi}(\text{orbital area}) \\
&= \tfrac{2m\omega}{\pi B}(\text{orbital flux}) \\
\text{conserved energy } E &= \tfrac{1}{2}m(2\omega R)^2 \\
&= \omega\mathcal{L}
\end{aligned}$$

Both are periodic functions of $t_1$, and both become infinite at $t_1 = n\tau$. When, with the assistance of *Mathematica*, we feed (65) into $S = \int L\,dt$ we obtain

$$S(\boldsymbol{x}, t; \boldsymbol{x}_0, 0) = \tfrac{1}{2}m\omega\Big\{\cot\omega t\big[(x - x_0)^2 + (y - y_0)^2\big] + 2(x_0 y - y_0 x)\Big\} + \tfrac{1}{2}m\frac{(z - z_0)^2}{t} \tag{66}$$

where the subscripts have now been dropped from $x_1$ and $t_1$. The Hamiltonian $H = \boldsymbol{p}\cdot\dot{\boldsymbol{x}} - L$ latent in (64) is

$$H = \tfrac{1}{2m}\Big\{\big[p_x + m\omega y\big]^2 + \big[p_y - m\omega x\big]^2 + p_z^2\Big\}$$

so the Hamilton-Jacobi equation reads

$$\tfrac{1}{2m}\Big\{\big[S_x + m\omega y\big]^2 + \big[S_y - m\omega x\big]^2 + S_z^2\Big\} + S_t = 0$$

of which, as a calculation confirms, the $S$ described above is in fact a solution. Equation (66) agrees precisely (except for a misprinted sign) with the result quoted on page 167 of Pauli.[6] That Pauli was even aware of the result is a little bit surprising, since the *derivation* of (66) requires some fairly heavy calculation—duck soup for *Mathematica*, but heroic if done with paper and pencil; Pauli, however, "knew everything"—especially things having to do with the semi-classical physics of magnetically perturbed quantum systems, in which there was, for experimental reasons, a high level of interest during the first quarter of the 20$^{\text{th}}$ Century.

In the limit $t \downarrow \tau$ (66) becomes

$$\begin{aligned} S(\boldsymbol{x},\tau;\boldsymbol{x}_0,0) = \tfrac{m}{2}\tfrac{(x-x_0)^2+(y-y_0)^2+(z-z_0)^2}{\tau} &+ m\omega(x_0y-y_0x) \\ &- \tfrac{1}{6}m\omega^2\big[(x-x_0)^2+(y-y_0)^2\big]\tau+\cdots \end{aligned}$$

Observe that the term of $O(\tau^0)$ is correctly reproduced by Feynman's "midpoint rule:"

$$m\omega\Big[-\tfrac{y+y_0}{2}(x-x_0)+\tfrac{x+x_0}{2}(y-y_0)\Big] = m\omega(x_0y-y_0x)$$

So much for the classical physics of a charged particle in a homogeneous magnetic field. We note in passing that in the Old Quantum Theory one would set

$$\text{angular momentum} = \hbar\cdot(\text{integer})$$

and that this would have entailed

$$\begin{aligned} \text{energy} &= \hbar\omega\cdot(\text{integer}) \\ \text{orbital flux} &= \pi\frac{\hbar c}{e}\cdot(\text{integer}) \\ &= \frac{hc}{e}\cdot\tfrac{\text{integer}}{2} \end{aligned}$$

Turning now to the quantum theory of the system, we construct the Van Vleck determinant

$$D = (-)^2\begin{vmatrix} -m\omega\cot\omega t & -m\omega \\ m\omega & -m\omega\cot\omega t \end{vmatrix} = \big(\tfrac{m\,\omega}{\sin\omega t}\big)^2$$

assemble Pauli's semi-classical propagator

$$\begin{aligned} K_{\mathsf{c}}(\boldsymbol{x},t;\boldsymbol{x}_0,0) &= \sqrt{\big(\tfrac{1}{i\hbar}\big)^2 D}\,\exp\Big\{\tfrac{i}{\hbar}S\Big\} \\ &= \tfrac{m\,\omega}{i\,h\sin\omega t}\exp\Big\{\tfrac{i}{\hbar}\tfrac{1}{2}m\,\omega\Big\{\cot\omega t\big[(x-x_0)^2+(y-y_0)^2\big]+2(x_0y-y_0x)\Big\}\Big\} \end{aligned} \tag{67}$$

and, with *Mathematica*'s assistance, confirm Pauli's observation that $K_C$ is in fact and *exact* solution of the Schrödinger equation

$$\tfrac{1}{2m}\Big\{\big[\tfrac{\hbar}{i}\partial_x+m\omega y\big]^2+\big[\tfrac{\hbar}{i}\partial_y-m\omega x\big]^2\Big\}\psi = i\hbar\partial_t\psi \tag{68}$$

. . . nor are we particularly surprised by this development: we expect to enjoy similar success when the components of $\boldsymbol{A}$ are *arbitrary* linear functions of $\{x, y, z\}$, even in the presence of a potential $V$ that depends arbitrarily—but at most quadratically—on those variables. More transparently,

$$\lim_{t\downarrow\tau} K_{\mathsf{c}}(\boldsymbol{x}, t; \boldsymbol{x}_0, 0) = \tfrac{m}{ih\tau}\exp\Big\{\tfrac{i}{\hbar}\tfrac{m}{2}\tfrac{(x-x_0)^2+(y-y_0)^2}{\tau}\Big\}$$
$$\downarrow$$
$$= \delta(x-x_0)\delta(y-y_0) \quad \text{in Gaussian representation}$$

We are now assured that all the spectral properties (eigenvalues/eigenfunctions) of the system are encrypted into the design of the right side of (67), even though that expression is assembled from classical components . . . but have yet to consider how such information might be extracted.

Bringing $z$ back into play contributes an $\{x, y\}$-independent additive term to $S$, and therefore a multiplicative factor to the propagator:

$$K_{\mathsf{c}}(\boldsymbol{x}, t; \boldsymbol{x}_0, 0) \;\longrightarrow\; K_{\mathsf{c}}(\boldsymbol{x}, t; \boldsymbol{x}_0, 0)\cdot\sqrt{\tfrac{m}{iht}}\exp\Big\{\tfrac{i}{\hbar}\tfrac{m}{2}\tfrac{(z-z_0)^2}{t}\Big\}$$

Observe finally (and relatedly) that when the vector potential is subjected to a gauge transformation $\boldsymbol{A} \to \boldsymbol{A} + \boldsymbol{\nabla}\chi$ the action responds

$$S(\boldsymbol{x}, t; \boldsymbol{x}_0, 0) \;\longrightarrow\; S(\boldsymbol{x}, t; \boldsymbol{x}_0, 0) + \tfrac{e}{c}\int_0^t \dot{\boldsymbol{x}}\boldsymbol{\cdot}\boldsymbol{\nabla}\chi\, dt' = S + \tfrac{e}{c}\int_{\boldsymbol{x}_0}^{\boldsymbol{x}} \boldsymbol{\nabla}\chi(\boldsymbol{\xi})\boldsymbol{\cdot} d\boldsymbol{\xi}$$
$$= S + \tfrac{e}{c}\big[\chi(\boldsymbol{x}) - \chi(\boldsymbol{x}_0)\big]$$

so the propagator responds

$$K_{\mathsf{c}}(\boldsymbol{x}, t; \boldsymbol{x}_0, 0) \;\longrightarrow\; e^{i(e/\hbar c)\chi(\boldsymbol{x})}\cdot K_{\mathsf{c}}(\boldsymbol{x}, t; \boldsymbol{x}_0, 0)\cdot e^{i(e/\hbar c)\chi(\boldsymbol{x}_0)} \tag{69}$$

But this is precisely the rule to which we are led when we bring to the spectral representation

$$K(\boldsymbol{x}, t; \boldsymbol{x}_0, 0) = \sum_n \psi_n(\boldsymbol{x})e^{-\frac{i}{\hbar}E_n t}\psi_n^*(\boldsymbol{x}_0)$$

the conclusion

$$\psi \;\longrightarrow\; e^{i(e/\hbar c)\chi(\boldsymbol{x})}\cdot\psi$$

to which we were led at (63). This final remark is, of course, not specific to the example that has recently concerned us . . . and here ends the discussion of that example.

I turn finally to discussion of a topic which illustrates the "pleasure of recognizing old things from a new point of view" and serves very nicely to demonstrate that there "are problems for which the new point of view offers a distinct advantage." Feynman (see again his Postulate II on page 10) would have us ascribe to each "path" $(\boldsymbol{x}, t) \leftarrow (\boldsymbol{x}_0, t_0)$ a

$$\text{path amplitude} = \tfrac{1}{A}e^{\frac{i}{\hbar}\int_{\text{path}} L(\dot{\boldsymbol{x}}, \boldsymbol{x})\, dt}$$

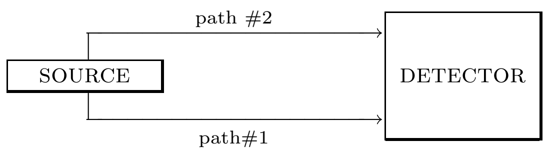
We saw at (54) how to "turn on a magnetic field," and see now that such a physical act entails a multiplicative adjustment of the probability amplitude ascribed to each individual path:

$$\text{path amplitude} \longrightarrow \text{path amplitude} \cdot e^{i(e/\hbar c)\int_{\text{path}} \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\cdot}\dot{\boldsymbol{x}}\,dt}$$

In any reasonable world (though not in a world where paths are allowed to become "almost nowhere differentiable," and thus to deny meaning to $\dot{\boldsymbol{x}}$) we expect—alternatively but more usefully—to be able to write

$$\text{path amplitude} \longrightarrow \text{path amplitude} \cdot e^{i(e/\hbar c)\int_{\text{path}} \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\cdot}d\boldsymbol{x}}$$

In 1959, Y. Aharonov & D. Bohm[37] considered experimental designs in which becomes natural to say of a charged particle that it gets from source to detector by one or the other of only two alternative paths:




The effect of the magnetic field, under such circumstances, is to introduce a

$$\text{phase difference} = (e/\hbar c)\left\{\int_{\text{path }\#1} - \int_{\text{path }\#2}\right\} = (e/\hbar c)\oint \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\cdot}d\boldsymbol{x}$$

*even if* $\boldsymbol{B} = \boldsymbol{\nabla}\times\boldsymbol{A} = \boldsymbol{0}$ *along the entire course of both paths*; the quantum particle is responsive to "remote" magnetic fields . . . for this simple reason: by Stokes' theorem

$$\begin{aligned} &= (e/\hbar c)\iint \boldsymbol{\nabla}\times\boldsymbol{A}\boldsymbol{\cdot}d\boldsymbol{\sigma} \\ &= (e/\hbar c)\cdot(\text{enveloped magnetic flux}) \end{aligned}$$

so it is important only that the path-bounded loop *envelop* magnetic flux, not that $\boldsymbol{B} \neq \boldsymbol{0}$ on the loop itself. Classically, a magnetic field announces its presence to a charged particle through the gauge-invariant construction

$$\boldsymbol{F} = (e/c)\,\dot{\boldsymbol{x}}\times\boldsymbol{B} = (e/c)\,\dot{\boldsymbol{x}}\times(\boldsymbol{\nabla}\times\boldsymbol{A})$$

while Aharonov & Bohm have isolated a quantum phenomenon in which the relevant construction is

$$\oint \boldsymbol{A}(\boldsymbol{x})\boldsymbol{\cdot}d\boldsymbol{x} \quad : \quad \text{gauge invariant by} \quad \oint \boldsymbol{\nabla}\chi\boldsymbol{\cdot}d\boldsymbol{x} = 0$$

[37] "Significance of electromagnetic potentials in quantum theory," Phys. Rev. **115**, 485.

In neither case is $\boldsymbol{A}$ itself observable; the two theories sense the vector potential in distinct but equally gauge-invariant ways. And if classical physics assigns no importance to $\oint \boldsymbol{A}(\boldsymbol{x})\cdot d\boldsymbol{x}$ it does assign high importance to the closely related construct

$$\begin{aligned}\text{“electromotive force”} &\equiv \oint \boldsymbol{E}(\boldsymbol{x})\cdot d\boldsymbol{x}\\ &= \iint \boldsymbol{\nabla}\times\boldsymbol{E}\cdot d\boldsymbol{\sigma}\\ &= -\tfrac{\partial}{\partial t}\iint \boldsymbol{B}\cdot d\boldsymbol{\sigma}\\ &= -\tfrac{\partial}{\partial t}(\text{enveloped magnetic flux})\end{aligned}$$

which provides yet a second mechanism by which magnetic fields acquire “remote” consequences. Aharonov & Bohm gained analytical sharpness by looking to a very simple two-path situation: Feynman, in less specialized situations, might write

$$\begin{aligned}K(\boldsymbol{x},t;\boldsymbol{x}_0,t_0) &= \tfrac{1}{A}\sum_{\text{paths}} e^{\frac{i}{\hbar}S[\text{path}]}\\ &\downarrow\\ &= \tfrac{1}{A}\sum_{\text{paths}} e^{\frac{i}{\hbar}S[\text{path}]}\cdot e^{i(e/\hbar c)\int_{\text{path}}\boldsymbol{A}(\boldsymbol{x})\cdot d\boldsymbol{x}}\end{aligned}$$

to describe the effect of “turning on a magnetic field.” The gauge transformation $\boldsymbol{A}\rightarrow\boldsymbol{A}+\boldsymbol{\nabla}\chi$ contributes additively to each $\int_{\text{path}}\boldsymbol{A}(\boldsymbol{x})\cdot d\boldsymbol{x}$ a path-independent function of the endpoints . . . which leaks out of the summation process to give us back precisely (69).

Some historical remarks: Aharonov and Bohm, who in 1959 were at the University of Bristol, were in some respects anticipated by W. Eherberg & R. E. Siday.[38] But they were using classical methods to study a quantum mechanical problem (electron optics), and seem to have been at pains to argue the *absence* of an AB effect. Aharonov & Bohm doe not allude in their brief paper to the path integral formalism (much less to its singular aptness), nor do Feynman & Hibbs allude to the AB effect; Feynman does, however, provide a luminous discussion of the AB effect in §15–5 of *The Feynman Lectures on Physics: Volume II* (1964). But his ostensible subject there is electrodynamics, not quantum mechanics, and he makes no reference to the path integral method. In 1984 Michael Berry—also at the University of Bristol—published the paper[39]

[38] “The refractive index in electron optics and the principles of dynamics,” Proc. Phys. Soc. London **B62**, 8 (1949). I have heard reports that, at an even earlier date, N. van Kampen—then a visitor at Columbia—assigned what was to become the Aharanov-Bohm effect to his quantum students as a homework problem!

[39] “Quantal phase factors accompanying adiabatic changes,” Proc. Roy. Soc. London **A392**, 45 (1984).

that launched the theory of "geometrical phase." He points out already in that first paper that the AB effect provides an instance of geometrical phase, but makes no reference to the Feynman formalism. The classic papers (by Aharonov & Bohm, Berry and many others) in this twin field[40] are reproduced in A. Shapere & F. Wilczek, *Geometric Phases in Physics* (1989), but I find in that collection only one paper[41] that makes explicit use of path integral methods. There seems to be an unaccountable hole in the literature.

**Sitting down and actually *doing* a path integral.** Look, for illustrative purposes, to the system of paramount importance to Feynman himself—the oscillator

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

which at $\omega \downarrow 0$ becomes the free particle. The classical action, in short time approximation, was found at (44) to be given by

$$S(x,\tau;y,0) = \tfrac{m}{2}\tfrac{(x-y)^2}{\tau} - \tfrac{1}{2}m\omega^2\tfrac{x^2+xy+y^2}{3}\tau$$

so our assignment is to evaluate

$$K(x,t;x_0,0) = \lim_{N\uparrow\infty}\left(\tfrac{m}{ih\tau}\right)^{N+1}\!\!\int\!\cdots\!\int \exp\Big\{\tfrac{im}{2\hbar\tau}\sum_{k=1}^{N+1}\Big[(x_k - x_{k-1})^2 - \tfrac{1}{3}(\omega\tau)^2(x_k^2 + x_k x_{k-1} + x_{k-1}^2)\Big]\Big\}\,dx_1\cdots dx_N \tag{70}$$

with $x_{N+1} = x$ and $\tau = t/(N+1)$. We already know many things about $K$: that it satisfies

$$-\tfrac{\hbar^2}{2m}K_{xx} + \tfrac{1}{2}m\omega^2 K = i\hbar K_t$$

and is given in fact by (39), which can be obtained by a great variety of means. To that list we want now to add another entry: recovery of (39) by execution of Feynman's program (70). That in itself can be accomplished in many ways. I have set things up with the intention to exploit the $n$-dimensional Gaussian integral formula[42]

$$\int_{-\infty}^{+\infty}\!\!\cdots\!\int_{-\infty}^{+\infty} e^{i\boldsymbol{y}\cdot\boldsymbol{x}}e^{-\frac{1}{2}\boldsymbol{x}\cdot\mathbb{A}\boldsymbol{x}}\,dx_1\cdots dx_n = \frac{(2\pi)^{\frac{n}{2}}}{\sqrt{\det\mathbb{A}}}e^{-\frac{1}{2}\boldsymbol{y}\cdot\mathbb{A}^{-1}\boldsymbol{y}} \tag{71}$$

---

[40] For a good introduction to the essentials of both of its parts see D. Griffiths, *Introduction to Quantum Mechanics* (1994), §10.2.

[41] Hiroshhi Kuratsuji & Shinji Iida, "Effective action for adiabatic process: dynamical meaning of Berry and Simon's phase," Prog. Theo. Phys. **74**, 439 (1985).

[42] For the remarkably simple proof, see (for example) Harald Cramér, *Mathematical Methods of Statistics* (1946), pages 99 and 118–121.

where it is assumed that $\mathbb{A}$ is real and symmetric, and that its eigenvalues are all positive (though we—on the usual grounds—will allow ourselves to relax the latter assumption). As a first step, we satisfy ourselves that

$$\exp\Big\{\text{etc.}\Big\} = e^{-\frac{1}{2}\beta(1-\alpha)}(x^2 + x_0^2)\cdot e^{\frac{1}{2}\beta(2+\alpha)}(x_0x_1 + x_{N-1}x)$$
$$\cdot\exp\Big\{-\tfrac{1}{2}\beta(2+\alpha)\Big[2\tfrac{1-\alpha}{2+\alpha}\sum_{k=1}^{N-1}x_k^2 - \sum_{k=1}^{N-2}x_kx_{k+1}\Big]\Big\}$$

with $\beta \equiv -\frac{im}{\hbar\tau}$ and $\alpha \equiv \frac{1}{3}(\omega\tau)^2$. Use $2\frac{1-\alpha}{2+\alpha} = 2 - 3\alpha + \cdots = 2 - (\omega\tau)^2 \equiv a$, drop the $\alpha$'s on grounds that they can make no contribution in the limit $\tau\downarrow 0$ and obtain

$$= e^{-\frac{1}{2}\beta}(x^2 + x_0^2)\cdot e^{i\boldsymbol{y}\cdot\boldsymbol{x}}e^{-\frac{1}{2}\boldsymbol{x}\cdot\mathbb{A}\boldsymbol{x}}$$

with

$$\boldsymbol{x} \equiv \begin{pmatrix} x_1\\ x_2\\ x_3\\ \vdots\\ x_{N-1}\\ x_N\end{pmatrix},\quad \boldsymbol{y} \equiv \begin{pmatrix} -i\beta x_0\\ 0\\ 0\\ \vdots\\ 0\\ -i\beta x_N\end{pmatrix},\quad \mathbb{A} \equiv \beta\begin{pmatrix} a & -1 & 0 & & & \\ -1 & a & -1 & & & \\ & -1 & a & & & \\ & & & \ddots & & \\ & & & & a & -1\\ & & & & -1 & a\end{pmatrix}$$

Notice that $a$ is the only surviving repository of the $\omega$ that serves to distinguish one oscillator from another, and oscillators in general from free particles. It follows now from (70) by (71) that

$$K(x,t;x_0,0) = \lim_{N\uparrow\infty}\left(\tfrac{m}{ih\tau}\right)^{\frac{N+1}{2}}\frac{(2\pi)^{\frac{N}{2}}}{\sqrt{\det\mathbb{A}}}\,e^{-\frac{1}{2}\beta}(x^2 + x_0^2)\cdot e^{-\frac{1}{2}\boldsymbol{y}\cdot\mathbb{A}^{-1}\boldsymbol{y}} \tag{72}$$

where the meanings of $\tau$, $\boldsymbol{y}$ and $\mathbb{A}$ are all $N$-dependent.

We look first to the asymptotic evaluation of the determinant. Let

$$D_n(N) \equiv \begin{cases}\text{determinant of the } n\times n \text{ submatrix that stands in the}\\ \text{upper left corner of the } N\times N \text{ matrix } \mathbb{A}\end{cases}$$

Expansion on the top row gives the recursion relation

$$D_n = a\beta D_{n-1} - \beta^2 D_{n-2} \tag{73}$$

which—if (formally) we set

$$D_0 = 0$$

and use

$$D_1 = \beta a$$

—is readily seen to reproduce the results of direct calculation. The idea now (for which I am indebted to Elliott Montroll[43]) is to notice that (73) can be written

$$\mathcal{D}_n - 2\mathcal{D}_{n-1} + \mathcal{D}_{n-2} = -(\omega\tau)^2\,\mathcal{D}_{n-1}$$

or again

$$\frac{1}{\tau}\left[\frac{\mathcal{D}_n - \mathcal{D}_{n-1}}{\tau} - \frac{\mathcal{D}_{n-1} - \mathcal{D}_{n-2}}{\tau}\right] = -\omega^2\,\mathcal{D}_{n-1}$$

with $\mathcal{D}_n \equiv \omega\tau\, D_n/\beta^n$. Now associate the $\mathcal{D}_n$ $\{n = 0, 1, 2, \ldots, N\}$ with the values assumed by a continuous function $\mathcal{D}(\vartheta)$ at discrete values of its argument

$$\mathcal{D}_n = \mathcal{D}(n\tau)$$

and notice that the preceding difference equation becomes asymptotically the differential equation

$$\tfrac{d^2}{d\vartheta^2}\,\mathcal{D} = -\omega^2\,\mathcal{D}$$

which we want to solve subject to the initial conditions

$$\begin{aligned} \mathcal{D}\,(0) &= 0 \quad &&: \quad \text{from } \mathcal{D}_1 = \omega\tau a \to 0 && \text{as } \tau \downarrow 0 \\ \mathcal{D}'(0) &= \omega \quad &&: \quad \text{from } [\mathcal{D}_2 - \mathcal{D}_1]/\tau = \omega[(a^2-1) - a] = \omega && \text{as } \tau \downarrow 0 \end{aligned}$$

Immediately $\mathcal{D}(\vartheta) = \sin\omega\vartheta$, so for large $N$ we have (since $n = N$ entails $\vartheta = t$)

$$\det \mathbb{A} = D_N = \tfrac{1}{\omega\tau}\beta^N\mathcal{D}_N \to \tfrac{1}{\omega\tau}\beta^N \sin\omega t \tag{74.1}$$

Turning now to the asymptotic evaluation of $\boldsymbol{y}\cdot\mathbb{B}\,\boldsymbol{y}$ with $\mathbb{B} \equiv \mathbb{A}^{-1}$, the sparse design of $\boldsymbol{y}$ entails

$$\boldsymbol{y}\cdot\mathbb{B}\,\boldsymbol{y} = B_{11}y_1^2 + (B_{1N} + B_{N1})y_1 y_N + B_{NN}y_N^2$$

so we have actually to obtain only the four corner elements of $\mathbb{B}$, and the simple design of $\mathbb{A}$ makes those quite easy to compute: we find

$$\begin{aligned} B_{11} &= B_{NN} = \frac{D_{N-1}}{D_N} = \frac{\mathcal{D}_{N-1}}{\beta\,\mathcal{D}_N} \\ B_{1N} &= B_{N1} = \beta^{N-1}\frac{1}{D_N} = \frac{\omega\tau}{\beta\,\mathcal{D}_N} \end{aligned}$$

---

[43] "Markoff chains, Wiener integrals, and quantum theory," Comm. Pure & Appl. Math. **5**, 415 (1952). See especially page 432. Montroll was highly esteemed for his computational virtuosity, and (together with Julian Schwinger) exerted a formative influence on my own early development. The paper here in question sprang from an "abortive attempt (1946–47) to develop a discrete space-time quantum theory"—an effort in which Feynman's publication[5] caused Montroll to lose interest.

giving

$$\begin{aligned}\boldsymbol{y}\cdot\mathbb{B}\,\boldsymbol{y} &= \frac{\mathcal{D}_{N-1}}{\beta\,\mathcal{D}_N}\big[(-i\beta x_0)^2 + (-i\beta x)^2\big] + 2\frac{\omega\tau}{\beta\,\mathcal{D}_N}(-i\beta x_0)(-i\beta x)\\ &= -\beta\left[(x^2+x_0^2)\frac{\mathcal{D}_{N-1}}{\mathcal{D}_N} + 2xx_0\frac{\omega\tau}{\mathcal{D}_N}\right]\end{aligned}$$

The exponential factors in (72) now combine to give

$$\begin{aligned}e^{-\frac{1}{2}\beta(x^2+x_0^2)}\cdot e^{-\frac{1}{2}\boldsymbol{y}\cdot\mathbb{A}^{-1}\boldsymbol{y}} &= \exp\left\{\tfrac{1}{2}\tfrac{m}{i\hbar\tau}\left[(x^2+x_0^2)\left\{\frac{\mathcal{D}_{N-1}}{\mathcal{D}_N}-1\right\} + 2xx_0\frac{\omega\tau}{\mathcal{D}_N}\right]\right\}\\ &= \exp\left\{\tfrac{i}{\hbar}\frac{m}{2\mathcal{D}_N}\left[(x^2+x_0^2)\frac{\mathcal{D}_N-\mathcal{D}_{N-1}}{\tau} - 2\omega x x_0\right]\right\}\end{aligned}$$

But asymptotically

$$\begin{aligned}\mathcal{D}_N &\to \mathcal{D}\,(t) = \sin\omega t\\ \frac{\mathcal{D}_N-\mathcal{D}_{N-1}}{\tau} &\to \mathcal{D}'(t) = \omega\cos\omega t\end{aligned}$$

so we have

$$e^{-\frac{1}{2}\beta(x^2+x_0^2)}\cdot e^{-\frac{1}{2}\boldsymbol{y}\cdot\mathbb{A}^{-1}\boldsymbol{y}} \to \exp\left\{\tfrac{i}{\hbar}\tfrac{m\omega}{2\sin\omega t}\big[(x^2+x_0^2)\cos\omega t - 2xx_0\big]\right\} \tag{74.2}$$

Returning with (74) to (72) we obtain at last

$$\begin{aligned}K(x,t;x_0,0) &= \left[\lim_{N\uparrow\infty}\big(\tfrac{m}{ih\tau}\big)^{\frac{N+1}{2}}(2\pi\tfrac{i\hbar\tau}{m})^{\frac{N}{2}}\sqrt{\tfrac{\omega\tau}{\sin\omega t}}\right]\cdot\exp\Big\{\text{etc.}\Big\}\\ &= \sqrt{\tfrac{m\omega}{ih\sin\omega t}}\exp\left\{\tfrac{i}{\hbar}\tfrac{m\omega}{2\sin\omega t}\big[(x^2+x_0^2)\cos\omega t - 2xx_0\big]\right\}\end{aligned} \tag{75}$$

—in precise agreement with (39). At (70) I tacitly assigned to Feynman's "normalization factor" $A$ its Pauli valuation $\sqrt{ih\tau/m}$. Had we (with Feynman) left the value of $A$ in suspension then we would, just above, have confronted

$$\left[\lim_{N\uparrow\infty}\big(\tfrac{1}{A}\big)^{N+1}(2\pi\tfrac{i\hbar\tau}{m})^{\frac{N}{2}}\sqrt{\tfrac{\omega\tau}{\sin\omega t}}\right] \quad : \quad \text{exists } only \text{ if } A = \sqrt{ih\tau/m}$$

This result is remarkable in several respects. We started with a short time *approximation* to the classical oscillator action, but ended up with a formula in which the *exact* oscillator action appears in the exponent. Equation (75) can be written

$$\begin{aligned}K &= \sqrt{\tfrac{1}{ih}\tfrac{\partial^2 S}{\partial x\partial x_0}}\,e^{\frac{i}{\hbar}S}\\ S &= \tfrac{m\omega}{2\sin\omega t}\big[(x^2+x_0^2)\cos\omega t - 2xx_0\big]\end{aligned}$$

It is, in short, of precisely Pauli/Van Vleck's design, though the time interval $t$ is now not infinitsimal but finite/arbitrary. Which is to say: now that all the tedious summing-over-paths lies behind us we were left with a result to which, in effect, only a single path—the *classical* path—contributes. That magical state of affairs persists, of course, when we proceed to the *free particle limit* $\omega\downarrow 0$, where the result now in hand returns the familiar result

$$\downarrow \\ = \sqrt{\tfrac{m}{iht}}\,\exp\Big\{\tfrac{i}{\hbar}\tfrac{m}{2t}\big[x^2-x_0^2\,\big]\Big\} \tag{76}$$

Had we "turned off the spring" at the outset, the sum-over-paths would have been easier . . . but only marginally.

Stand back and squint at our accomplishment: we have been engaged in an exercise in *multivariable Gaussian integration . . . carried ultimately to the limit in which the number of variables is allowed to become infinite*. It has been an argument of sufficiently many parts that it can be organized in a great many ways. The argument presented here was taken from my QUANTUM MECHANICS (1967) Chapter One, pages 42–48, and owes much to Montroll.[44]

We expect to enjoy similar success whenever the classical action can, at short times, be satisfactorally approximated by a function which depends at most *quadratically* on its arguments. Gaussian integration will in all such cases be the name of the game. For many applications—for example: to quantum field theory—that turns out to be not so severe a limitation as one might suppose.[45]

We cannot expect to be able—by *any* method—to construct *exact* descriptions of the propagator $K(x,t;x_0,t_0)$ except in a relatively limited number of textbookish cases, and are therefore not surprised to discover that we are frequently *unable to evaluate* the Feynman path integral. After all, we are more often than not unable to evaluate—except numerically—the ordinary integrals of functions of a single variable: integration is hard. We expect generally to have to make do with approximation schemes, of one design or another. One important recommendation of the Feynman method is that it presents old problems in quite a new light—a light that invites the invention of novel approximation methods.

That said, it can be reported that path integration is a field in which great strides were taken during the last quarter of the $20^{\text{th}}$ Century. A leading figure in this effort has been Christian Grosche, at the Institut für Theoretische Physik, Universität Hamburg. His *Path Integrals, Hyperbolic Spaces, & Selberg Trace Formulae* (1996) cites more than 500 contributions to the field, and provides a vivid sense of the range and level of sophistication that has recently

---

[44] In the old notes just cited I describe also several alternative lines of attack. See also Chapter 6 in Schulman,[12] and the references cited by him.

[45] Michio Kaku, in Chapter 8 of his *Quantum Field Theory* (1993), begins his survey of our subject with the claim that "the path integral approach has many advantages over the other [quantization] techniques," and proceeds to list seven of its distinctive virtues.

been achieved. In "How to solve path integrals in quantum mechanics,"[46] which provides an excellent survey of the present state of the field, Grosche claims it to be "no exaggeration to say that we are able to solve today essentially all path integrals in quantum mechanics which correspond to problems for which the corresponding Schrödinger problem can be solved exactly." But it is in more abstract areas that lie farther afield (supersymmetric string theory, quantum gravity), where "solving the Schrödinger equation" is not even the point at issue, that the power of Feynman's method becomes most pronounced.

**Summing over *what* paths?** Feynman writes

$$K(x,t;x_0,0) = (\text{normalization factor})\cdot\sum_{\text{paths}} e^{\frac{i}{\hbar}S[\text{path from }(x_0,0)\text{ to }(x,t)]}$$

but to lend concrete meaning to the picturesque expression on the right he would have us write the appropriate variant of (70).[47] To describe the class of paths he has in mind, the class implicit in (70), he draws something like the $N$-node spline curves shown in Figure 1 and Figure 2, which we are to imagine in the limit $N\uparrow\infty$. It appears to have been Feynman's view (see the following figure) that his mathematical discovery was indicative—not in "as if" terms, but literally—of an underlying physical fact: that particles *really do* trace almost nowhere differentiable fractile-like curves in spacetime. And that the concept of "velocity" survives, therefore, with only a statistical meaning. The latter conclusion is not, in itself, radical: the theory of Brownian motion leads, in its idealized formalism, to a similar conclusion, and so did Dirac's notion of "zitterbewegung." I argue here that Feynman's purported view (which seems to me to reflect a naive realism unworthy of the man, and is supported by no direct physical evidence) is untenable on these grounds: *summation over distinct classes of paths leads (at least in some instances) to identical conclusions.* This I demonstrate by example.

Look again to the one-dimensional oscillator

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

for which the dynamical path $(x_1,t_1)\leftarrow(x_0,t_0)$ is known to be described by[48]

$$x_{\mathsf{c}}(t) = \left[\frac{x_0\sin\omega t_1 - x_1\sin\omega t_0}{\sin\omega(t_1-t_0)}\right]\cos\omega t - \left[\frac{x_0\cos\omega t_1 - x_1\cos\omega t_0}{\sin\omega(t_1-t_0)}\right]\sin\omega t$$

Write

$$\begin{aligned} x(t) &= x_{\mathsf{c}}(t) + \lambda a(t) \qquad (77)\\ &\qquad\qquad a(t)\equiv\sum_{n=1}^\infty a_n\sin\Big\{n\pi\frac{t-t_0}{t_1-t_0}\Big\}\\ &\equiv x(t;a_1,a_2,\ldots)\end{aligned}$$

[46] J. Math. Phys. **36**, 2354 (1995), written jointly with F. Steiner.
[47] . . . which as it stands is specific to the harmonic oscillator.
[48] See QUANTUM MECHANICS (1967), Chapter I, page 22. Or simply *verify* that indeed $\ddot{x}+\omega^2x=0$ and $x(t_0)=x_0$, $x(t_1)=x_1$.

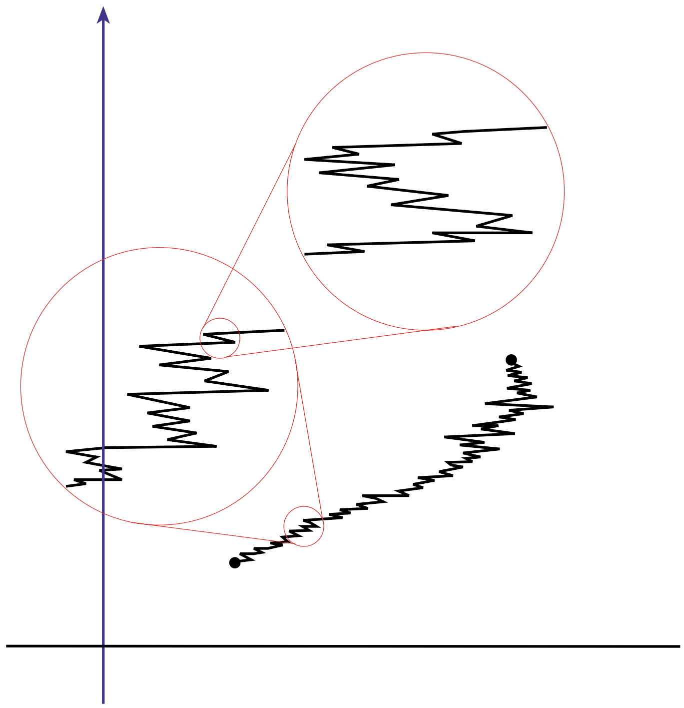


FIGURE 4: *Such a figure appears on page 177 of Feynman & Hibbs'* Quantum Mechanics and Path Integrals (1965). *The accompanying text suggests that Feynman entertained the view that particles really do move about tracing nowhere-differentiable fractal-like trajectories —this well before Benoit Mandelbrot introduced "fractal" into the vocabulary of the physical sciences. I argue here that Feynman's view (if, indeed, he held such a view: certainly he allowed himself to speak as though he did) is mathematically untenable and physically unjustified.*

to describe the elements of what we will now take to be the "class of admissible paths." The $a_n$ serve to address the individual elements (paths) in the "space of paths," while $\lambda$ is a formal device that will in a moment be set equal to unity. The point to notice is that $a(t_0) = a(t_1) = 0$; i.e., that

$$\left.\begin{aligned} x(t_0;\boldsymbol{a}) &= x_0 \\ x(t_1;\boldsymbol{a}) &= x_1 \end{aligned}\right\} \quad \text{for all } \boldsymbol{a}$$

Now insert $x_c + \lambda a$ into $S = \int L\,dt$ and obtain

$$S[x(t;\boldsymbol{a})] = S_0 + \lambda S_1 + \lambda^2 S_2$$

where

$$S_0 \equiv \tfrac{1}{2}m\int_{t_0}^{t_1}\big\{\dot{x}_\text{c}^2 - \omega^2 x_\text{c}^2\big\}\,dt = \tfrac{m\omega}{2\sin\omega(t_1 - t_0)}\big[(x_1^2 + x_0^2)\cos\omega(t_1 - t_0) - 2x_1x_0\big]$$

is already very well known, where

$$\begin{aligned} S_1 &\equiv m\int_{t_0}^{t_1}\big\{\dot{x}_\text{c}\dot{a} - \omega^2 x_\text{c} a\big\}\,dt \\ &= m\int_{t_0}^{t_1}\big\{-\ddot{x}_\text{c} - \omega^2 x_\text{c}\big\}a\,dt + \dot{x}_\text{c}a\Big|_{t_0}^{t_1} \quad \text{after integration by parts} \\ &= 0 \quad \text{because } \begin{cases} x_\text{c}(t) \text{ satisfies the equation of motion, and} \\ a(t) \text{ vanishes at times } t_0 \text{ and } t_1 \end{cases} \end{aligned}$$

and where a calculation that exploits

$$\int_0^\pi \cos m\xi \cos n\xi\,d\xi = \int_0^\pi \sin m\xi \sin n\xi\,d\xi = \tfrac{\pi}{2}\delta_{mn}$$

supplies

$$\begin{aligned} S_2 &\equiv \tfrac{1}{2}m\int_{t_0}^{t_1}\big\{\dot{a}^2 - \omega^2 a^2\big\}\,dt \\ &= \tfrac{m\pi}{2T}\sum_{n=1}^\infty n^2\Big[1 - \big(\tfrac{\omega T}{n\pi}\big)^2\Big]a_n^2 \quad \text{with} \quad T \equiv t_1 - t_0 \end{aligned}$$

Set $\lambda = 1$ and agree to

$$\text{interpret} \sum_\text{paths} \text{ to mean } \lim_{N\uparrow\infty}\int\cdots\int (w_1 da_1)\cdots(w_N da_N)$$

where the $w_n$ are "weight functions" that will presently be assigned meaning in such a way as to make things work out right. We now have

$$\begin{aligned} K(x,t;x_0,t_0) &= (\text{normalization factor})\cdot e^{\frac{i}{\hbar}S_0(x,t;x_0,t_0)} \\ &\qquad \cdot \lim_{N\uparrow\infty}\int\cdots\int \exp\Big\{\tfrac{i}{\hbar}\tfrac{m\pi}{2T}\sum_{n=1}^N n^2\Big[1 - \big(\tfrac{\omega T}{n\pi}\big)^2\Big]a_n^2\Big\}w_1da_1w_2da_2\cdots \\ &= (\text{etc.})\cdot \lim_{N\uparrow\infty}\prod_{n=1}^N w_n\int \exp\Big\{-\tfrac{m\pi}{2i\hbar T}n^2\Big[1 - \big(\tfrac{\omega T}{n\pi}\big)^2\Big]a^2\Big\}da \\ &= (\text{etc.})\cdot \lim_{N\uparrow\infty}\prod_{n=1}^N w_n\Big\{\tfrac{m}{2i\hbar T}n^2\Big[1 - \big(\tfrac{\omega T}{n\pi}\big)^2\Big]\Big\}^{-\frac{1}{2}} \end{aligned}$$

But a classic identity informs us (so does *Mathematica*) that

$$\prod_{n=1}^\infty\Big[1 - \big(\tfrac{z}{n\pi}\big)^2\Big] = \frac{\sin z}{z}$$

With that fact in mind, we set $w_n(T) \equiv \sqrt{\frac{m}{2i\hbar T}}\,n$ and obtain

$$\begin{aligned} &= (\text{etc.})\cdot\sqrt{\tfrac{\omega T}{\sin\omega T}} \\ &= (\text{normalization factor})\sqrt{\tfrac{ihT}{m}}\cdot\sqrt{\tfrac{m\omega}{ih\sin\omega(t_1 - t_0)}}\,e^{\frac{i}{\hbar}S_0(x,t;x_0,t_0)} \end{aligned}$$

We have now only to set

$$(\text{normalization factor}) = \sqrt{\tfrac{m}{ihT}}$$

to recover the familiar exact propagator for the harmonic oscillator.

The point is that we have here summed over a class of paths distinct from that contemplated by Feynman—paths which (though they may exhibit discontinuities and points of non-differentiability in the limit $N \uparrow \infty$) are for all finite $N$ *everywhere* differentiable. And to the extent that

$$\sum_{\text{paths}} \text{ is independent of the precise definition of “path space”}$$

every particular path space loses any claim to “objective physical significance.”

The point has been developed in an especially sharp and revealing way by C. W. Kilmister,[49] whose pretty argument I now sketch. Retain (77) but require of the functions $a(t)$ *only* that $a(t_0) = a(t_1) = 0$. Develop $L(x, \dot{x})$ in powers of $\lambda$:

$$\begin{aligned} L(x_{\mathsf{c}} + \lambda a, \dot{x}_{\mathsf{c}} + \lambda \dot{a}) &= \exp\Big\{\lambda\Big(a\frac{\partial}{\partial x_{\mathsf{c}}} + \dot{a}\frac{\partial}{\partial \dot{x}_{\mathsf{c}}}\Big)\Big\} L(x_{\mathsf{c}}, \dot{x}_{\mathsf{c}}) \\ &= \sum_{k=0}^{\infty} \lambda^k L_k(x_{\mathsf{c}}, \dot{x}_{\mathsf{c}}, a, \dot{a}) \end{aligned}$$

Then

$$\begin{aligned} S[x(t)] &= \sum_{k=1}^{\infty} \lambda^k S_k[x_1, t_1; x_0, t_0; a(t)] \\ &\qquad S_k[x_1, t_1; x_0, t_0; a(t)] \equiv \int_{t_0}^{t_1} L_k(x_{\mathsf{c}}, \dot{x}_{\mathsf{c}}, a, \dot{a})\, dt \end{aligned}$$

Now set $\lambda = 1$ and notice in particular that

$$\begin{aligned} S_0 &= S[x_{\mathsf{c}}(t)] && : \quad \text{the classical action} \\ S_1 &= \int_{t_0}^{t_1} \Big(a\frac{\partial L}{\partial x_{\mathsf{c}}} + \dot{a}\frac{\partial L}{\partial \dot{x}_{\mathsf{c}}}\Big) dt \\ &= \int_{t_0}^{t_1} a\Big(\frac{\partial L}{\partial x_{\mathsf{c}}} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}_{\mathsf{c}}}\Big) dt - \frac{\partial L}{\partial \dot{x}_{\mathsf{c}}} a\Big|_{t_0}^{t_1} = 0 && : \quad \text{Hamilton's principle} \end{aligned}$$

Feynman would now have us write

$$\begin{aligned} K(x_1, t_1; x_0, t_0) = &\quad \tfrac{1}{A}\sum_{\text{paths}} \exp\Big\{\tfrac{i}{\hbar}\sum_{k=0} S_k[x_1, t_1; x_0, t_0; a(t)]\Big\} \\ &= \tfrac{1}{A} e^{\frac{i}{\hbar}S_0(x, t; x_0, t_0)} \cdot \sum_{\text{paths}} \exp\Big\{\tfrac{i}{\hbar}\sum_{k=2} S_k[x_1, t_1; x_0, t_0; a(t)]\Big\} \end{aligned} \tag{78}$$

Thus far our results are quite general.

---

[49] “A note on summation over Feynman histories,” Proc. Camb. Phil. Soc. **54**, 302 (1958).

Look now to systems of the specialized design

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

Quick calculation then supplies

$$\begin{aligned}
L_0 &= \tfrac{1}{2}m\dot{x}_{\mathsf{c}}^2 - V(x_{\mathsf{c}})\\
L_1 &= \text{need not be computed}\\
L_2 &= \tfrac{1}{2}m\,\dot{a}^2 - \tfrac{1}{2}V''(x_{\mathsf{c}})\,a^2\\
&\;\;\vdots\\
L_k &= \qquad\quad - \tfrac{1}{k!}V^{(k)}(x_{\mathsf{c}})\,a^k \quad : \quad k \geqslant 3
\end{aligned}$$

and (78) becomes

$$\begin{aligned}
K = \tfrac{1}{A}e^{\frac{i}{\hbar}S_0}\cdot\sum_{\text{paths}} &\exp\Big\{\tfrac{i}{\hbar}\int_{t_0}^{t_1}\big(\tfrac{1}{2}m\,\dot{a}^2 - \tfrac{1}{2}V''(x_{\mathsf{c}})\,a^2\big)\,dt\Big\}\\
&\cdot\exp\Big\{-\tfrac{i}{\hbar}\sum_{k=3}\int_{t_0}^{t_1}\tfrac{1}{k!}V^{(k)}(x_{\mathsf{c}})\,a^k\,dt\Big\}
\end{aligned}\tag{79}$$

Impose upon $L(x,\dot{x})$ the further restriction that (as in all non-magnetic cases successfully treated thus far) $V(x)$ depends at most *quadratically* upon its arguments: under such circumstances we may as well write

$$V(x) = mgx + \tfrac{1}{2}m\omega^2x^2$$

Then $V''(x_{\mathsf{c}}) = m\omega^2$ and $V^{(k)}(x_{\mathsf{c}}) = 0$: the red factor in (79) becomes $x_{\mathsf{c}}$-independent, and we have

$$K(x_1,t_1;x_0,t_0) = \tfrac{1}{A}\underbrace{\sum_{\text{paths}}\exp\Big\{\tfrac{i}{\hbar}\int_{t_0}^{t_1}\tfrac{1}{2}m\big(\dot{a}^2 - \omega^2a^2\big)\,dt\Big\}}\cdot e^{\frac{i}{\hbar}S_0(x_1,t_1;x_0,t_0)}$$

This will be necessarily of the form $f(t_1-t_0)$ <u>however</u> the path integral is defined!

Feynman would tune the interpretation of $A(t_1-t_0)$ in such a way as to achieve

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

From the remark clipped to our final result Kilmister draws this moral: *The systems that supplied the Feynman formalism with its classic successes are too simple to distinguish one interpretation of "sum over paths" from another.*

The point at issue would assume importance if the Feynman formalism were pressed into service as an autonymous quantization procedure—if distinct and equivalent interpretations of

$$\sum_{\text{paths}} \quad : \quad \text{How? Over what class of paths?}$$

were available and no Schrödinger equation were available to serve as arbitor.[50]

**Predominance of the classical path.** In §7 of the 1948 RevModPhys paper, under the head "Discussion of the wave equation: the classical limit," Feynman writes "... Dirac's remarks were the starting point for the present development. The points he makes concerning the passage to the classical limit $\hbar \to 0$ are very beautiful, and I may perhaps be excused for reviewing them here." His review—anticipated at page 31 in the dissertation—is skeletal (as was Dirac's), but is spelled out greater didactic detail in §2–3 of Feynman & Hibbs. The essential idea is simple enough: write

$$\cos\left\{\tfrac{1}{\hbar}(S - S_{\mathsf{c}})^2\right\} = \text{real part of } \exp\left\{\tfrac{i}{\hbar}(S - S_{\mathsf{c}})^2\right\}$$

to model the effect of ranging over a population of paths $x(t)$ that lie in the immediate neighborhood of the classical path $x_{\mathsf{c}}(t)$. As $\hbar \downarrow 0$ the oscillations, except in the shrinking immediate neighborhood of $x_{\mathsf{c}}(t)$, become more and more densely spaced (see the following figure), and the integrated effect of such paths is to cancel each other out. On the other hand, paths in the *immediate* neighborhood of $x_{\mathsf{c}}(t)$ contribute coherently to the sum-over-paths, since—by Hamilton's principle—the classical path resides at an extremum of the action functional $S[x(t)]$. As Feynman & Hibbs sum up (no pun intended) the situation: "... no path really needs to be considered [in the classical limit] if the neighboring path has a different action, for the paths in the neighborhood [then] cancel out [its] contribution [to the path integral].

But in each of the cases that in preceding discussion yielded successfully to detailed analysis it emerged that *only* the classical path survived the path integration process, even though we did *not* take $\hbar \downarrow 0$; *i.e.*, that $K_{\mathsf{c}}$ was in fact *exact*. I propose to consider how such a state of affairs comes about.

Let $t_0 < t_1 < t_2$ and agree, for the moment, to work in one dimension. Though it is fundamental that

$$K(x_2, t_2; x_0, t_0) = \int K(x_2, t_2; x_1, t_1)\, dx_1 K(x_1, t_1; x_0, t_0) \tag{80}$$

---

[50] The preceding material was adapted from QUANTUM MECHANICS (1967), Chapter I, pages 55–64. A simplified account of Kilmister's paper can be found in §3.5 of his *Hamiltonian Dynamics* (1964). For related material see H. Davies, "Summation over Feynman histories: the free particle and the harmonic oscillator," Proc. Camb. Phil. Soc. **53**, 651 (1957) and especially S. G. Brush, "Functional integrals & statistical physics," Rev. Mod. Phys. **33**, 79 (1961), which provides an exhaustive guide to the older literature, and in §3 reviews several alternative "Methods for calculating functional integrals."

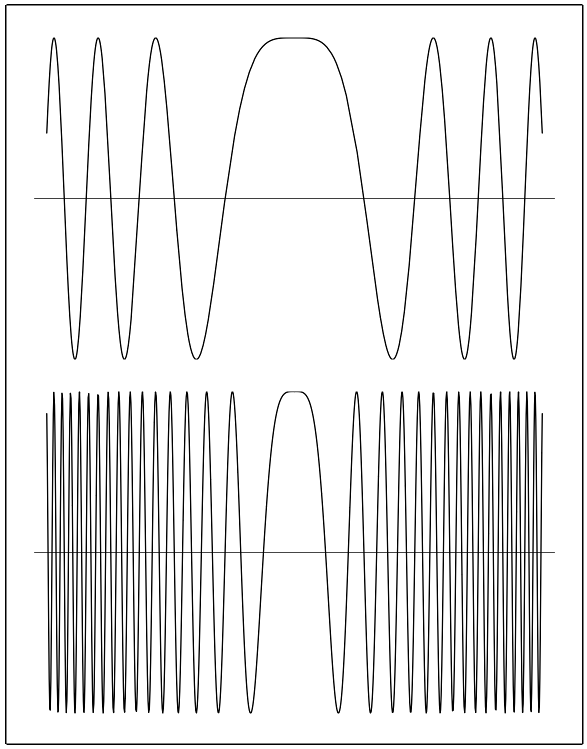


FIGURE 5: *Graphs of* $\cos\left\{\frac{1}{\hbar}(S-S_{\mathsf{c}})^2\right\}$ *with—reading from top to bottom—decreasing values of* $\hbar$. *The observation that only points in the immediate neighborhood of* $S_{\mathsf{c}}$ *contribute asymptotically to the value of* $\int f(S)\cos\left\{\frac{1}{\hbar}(S-S_{\mathsf{c}})^2\right\}dS$ *was attributed by Feynman to Dirac but is in fact ancient: it lies at the heart of all techniques addressed to the "asymptotic evaluation of integrals."*

we recognize that the Pauli/Van Vleck replacement $K\to K_{\mathsf{c}}$ yields a statement

$$K_{\mathsf{c}}(x_2,t_2;x_0,t_0)=\int K_{\mathsf{c}}(x_2,t_2;x_1,t_1)\,dx_1 K_{\mathsf{c}}(x_1,t_1;x_0,t_0) \tag{81}$$

that is typically *not* valid. It becomes, however, *approximately* valid

- as the time intervals become short, or alternatively
- as $\hbar\downarrow 0$

but is exactly valid in some special cases. Look, for example, to the simplest such special case, the FREE PARTICLE, where it is known that $K_{\mathsf{c}}=K$; under such a circumstance (81) is exact because *indistinguishable* from (80). It is instructive, nonetheless, to bring a magnifying glass to the details. The right side of (81) reads

$$\sqrt{\tfrac{1}{i\hbar}D(t_2-t_1)}\sqrt{\tfrac{1}{i\hbar}D(t_1-t_0)}\int e^{\frac{i}{\hbar}\left\{S(x_2,t_2;x_1,t_1)+S(x_1,t_1;x_0,t_0)\right\}}\,dx_1$$

$$=\sqrt{\tfrac{1}{2\pi}\beta\tfrac{1}{t_2-t_1}}\sqrt{\tfrac{1}{2\pi}\beta\tfrac{1}{t_1-t_0}}\int \exp\left\{-\tfrac{1}{2}\beta\left[\tfrac{(x_2-x_1)^2}{t_2-t_1}+\tfrac{(x_1-x_0)^2}{t_1-t_0}\right]\right\}dx_1$$

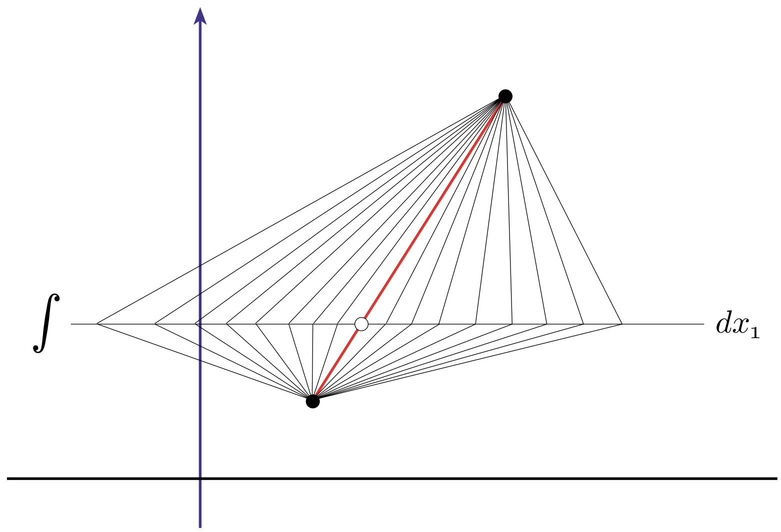



FIGURE 6: *A particle moves freely from $\{x_0, t_0\}$ to $\{x_1, t_1\}$, then freely again from $\{x_1, t_1\}$ to $\{x_2, t_2\}$. Paths within this simple population are identified by specifying the location of the solitary nodal point $x_1$. Analogs of the figure could be drawn for any system: just change "freely" to "dynamically."*

with $\beta \equiv m/i\hbar$. Ask *Mathematica* to perform the gaussian integral and obtain

$$= \sqrt{\tfrac{1}{2\pi}\beta\tfrac{1}{t_2-t_1}}\sqrt{\tfrac{1}{2\pi}\beta\tfrac{1}{t_1-t_0}}\sqrt{2\pi\tfrac{1}{\beta}\tfrac{(t_2-t_1)(t_1-t_2)}{t_2-t_0}}\exp\Big\{-\tfrac{1}{2}\beta\Big[\tfrac{(x_2-x_0)^2}{t_2-t_0}\Big]\Big\}$$
$$= \sqrt{\tfrac{1}{i\hbar}D(t_2-t_0)}\,e^{\frac{i}{\hbar}S(x_2,t_2;x_1,t_1)}$$

after obvious simplifications and a notational adjustment.[51] The integral that led to this result can be regarded as a path integral over the simple class of paths shown above—a path integral to which the only classical path (shown in red) makes a net contribution. We observe in this connection that

$$\frac{\partial}{\partial x_1}\Big\{\frac{m}{2}\Big[\frac{(x_2-x_1)^2}{t_2-t_1}+\frac{(x_1-x_0)^2}{t_1-t_0}\Big]\Big\} = 0 \quad\Rightarrow\quad x_1 = x_0 + \frac{x_2-x_0}{t_2-t_0}(t_1-t_0)$$

[51] Seen in this light, the composition rule (80) expresses a wonderful property of Gaussians, and an even more wonderful property of propagators in general . . . though when extracted from

$$K(x_1,t_1;x_0,t_0) = \sum_n e^{-\frac{i}{\hbar}E_n(t_1-t_0)}\psi_n(x_1)\psi_n^*(x_0)$$

it follows almost trivially from the orthonormality of the eigenfunctions and an elementary property of the exponential function.

In short (and consistently with Hamilton's principle), $S$-minimization places $x_1$ right on the classical path $x_{\mathsf{c}}(t)$ that links $\{x_2, t_2\} \leftarrow \{x_0, t_0\}$: $x_1 = x_{\mathsf{c}}(t_1)$. And we verify by quick calculation that

$$\tfrac{m}{2}\Big[\tfrac{(x_2-x_1)^2}{t_2-t_1} + \tfrac{(x_1-x_0)^2}{t_1-t_0}\Big] = \tfrac{m}{2}\Big[\tfrac{(x_2-x_0)^2}{t_2-t_0}\Big] \quad \text{at } x_1 = x_{\mathsf{c}}(t_1)$$

Notice that $S$-extremization can be rendered as a "smooth splice condition"

$$\tfrac{\partial}{\partial x_1}S(x_1,t_1;x_0,t_0) = -\tfrac{\partial}{\partial x_1}S(x_2,t_2;x_1,t_1)$$
$$\Updownarrow$$
$$\text{final momentum of first leg} = \text{initial momentum of second leg}$$

which is intuitively quite satisfying.

Look on the basis of this experience to

$$\lim_{\hbar\downarrow 0}\int\sqrt{(\tfrac{1}{ih})^n D(2,1)}\sqrt{(\tfrac{1}{ih})^n D(1,0)}\,e^{\frac{i}{\hbar}\{S(2,1)+S(1,0)\}}\,d\boldsymbol{x}_1$$

where we elect to work now in $n$ dimensions, and adopt the abbreviations $S(1,0) \equiv S(\boldsymbol{x}_1, t_1; \boldsymbol{x}_0, t_0)$, etc. The $n$-dimensional method of stationary phase[52] supplies

$$\lim_{\lambda\uparrow\infty}\int f(\boldsymbol{x})\,e^{i\lambda g(\boldsymbol{x})}dx_1\cdots dx_n \sim f(\boldsymbol{x})\,e^{i\lambda g(\boldsymbol{x})}\sqrt{(\tfrac{2\pi i}{\lambda})^n\tfrac{1}{\det\mathbb{G}(\boldsymbol{x})}} \tag{82}$$

where $\boldsymbol{\nabla}g(\boldsymbol{x})$ vanishes at $\boldsymbol{x} = \boldsymbol{x}$, and $\mathbb{G}(\boldsymbol{x}) \equiv \|\partial^2 g(\boldsymbol{x})/\partial x_i\partial x_j\|$. So we have

$$\lim_{\hbar\downarrow 0}\int\{\text{etc.}\}d\boldsymbol{x}_1 \sim (\tfrac{1}{ih})^n\sqrt{D(2,1)D(1,0)}\,e^{\frac{i}{\hbar}\{S(2,1)+S(1,0)\}}\sqrt{\tfrac{(ih)^n}{\det\mathbb{S}}}$$

which—because Hamilton's principle serves to place $\boldsymbol{x}_1$ on the classical path that links $\boldsymbol{x}_0$ to $\boldsymbol{x}_2$, and thus to achieve $S(2,1) + S(1,0) = S(2,0)$—becomes

$$= (\tfrac{1}{ih})^{\frac{n}{2}}\sqrt{\tfrac{D(2,1)D(1,0)}{\det\mathbb{S}}}\,e^{\frac{i}{\hbar}S(2,0)}$$

I will show in a moment that $D(2,1)D(1,0) = D(2,0)\cdot\det\mathbb{S}$, giving finally

$$= \sqrt{(\tfrac{1}{ih})^n D(2,0)}\,e^{\frac{i}{\hbar}S(2,0)} \tag{82}$$

We are brought thus to the quite general conclusion that, though Pauli's WKB approximation to the propagator fails except in special cases to satisfy the composition law (80), it invariably does so *in the classical limit* $\hbar\downarrow 0$.

[52] The one-dimensional formula—encountered most recently at (43), and sometimes attributed to Lord Kelvin (1887)—generalizes straightforwardly with the aid of the $n$-dimensional Gaussian integral formula (71); *i.e.*, by rotating to the coordinate system that renders $\mathbb{G}$ diagonal.

I turn now, as promised, to the proof of

$$D(2,1)D(1,0) = D(2,0)\cdot\det\mathbb{S} \tag{83}$$

which, though it has been seen to lie close to the heart of the mechanism by which classical paths come to make the predominant contributions to Feynman's path integrals, is itself a proposition that lives (obscurely!) within *classical* mechanics. The argument hinges on the fact that since $\boldsymbol{x}$ sits on the dynamical curve that links $\{\boldsymbol{x}_2,t_2\}\leftarrow\{\boldsymbol{x}_0,t_0\}$ it must be a function of the endpoints that serve to define that curve:

$$x_1^p = x_1^p(t_1;\boldsymbol{x}_2,t_2,\boldsymbol{x}_0,t_0)$$

Therefore

$$\begin{aligned}\frac{\partial^2}{\partial x_2^p\partial x_0^q}S(2,0) &= \frac{\partial^2}{\partial x_2^p\partial x_0^q}\big\{S(2,1)+S(1,0)\big\}\\ &= \frac{\partial}{\partial x_2^p}\bigg\{\frac{\partial x_1^r}{\partial x_0^q}\underbrace{\frac{\partial}{\partial x_1^r}\big\{S(2,1)+S(1,0)\big\}}_{0}+\frac{\partial S(1,0)}{\partial x_0^q}\bigg\} &&(\star)\\ &= \frac{\partial x_1^r}{\partial x_2^p}\cdot\frac{\partial^2 S(1,0)}{\partial x_1^r\partial x_0^q}\quad\text{on one hand} &&(84.1)\end{aligned}$$

and, by a similar argument,

$$= \frac{\partial^2 S(2,1)}{\partial x_2^p\partial x_1^r}\cdot\frac{\partial x_1^r}{\partial x_0^q}\quad\text{on the other} \tag{84.2}$$

But if we *work out* the right side of $(\star)$—including the abandoned term—we obtain

$$\begin{aligned}&= \frac{\partial^2 S(2,1)}{\partial x_2^p\partial x_1^r}\cdot\frac{\partial x_1^r}{\partial x_0^q}+\frac{\partial x_1^r}{\partial x_2^p}\cdot\frac{\partial^2 S(1,0)}{\partial x_1^r\partial x_0^q}\\ &\qquad+\bigg\{\frac{\partial^2 S(2,1)}{\partial x_1^r\partial x_1^s}+\frac{\partial^2 S(1,0)}{\partial x_1^r\partial x_1^s}\bigg\}\frac{\partial x_1^r}{\partial x_2^p}\frac{\partial x_1^s}{\partial x_0^q}\\ &\qquad+\frac{\partial^2 x_1^r}{\partial x_2^p\partial x_0^q}\cdot\underbrace{\frac{\partial}{\partial x_1^r}\big\{S(2,1)+S(1,0)\big\}}_{0}\end{aligned}$$

The first two terms on the right are *individually* equal to the expression on the left, so we have

$$= -\bigg\{\frac{\partial^2 S(2,1)}{\partial x_1^r\partial x_1^s}+\frac{\partial^2 S(1,0)}{\partial x_1^r\partial x_1^s}\bigg\}\cdot\frac{\partial x_1^r}{\partial x_2^p}\frac{\partial x_1^s}{\partial x_0^q} \tag{84.3}$$

In (84) we have three different descriptions of the same thing. Recalling from (25) the definition of the Van Vleck determinant (note particularly the presence of the $(-)^n$ factor) we obtain

$$D(2,0) = D(2,1)\cdot\left|\tfrac{\partial \boldsymbol{x}_1}{\partial \boldsymbol{x}_0}\right| = D(1,0)\cdot\left|\tfrac{\partial \boldsymbol{x}_1}{\partial \boldsymbol{x}_2}\right| = \det \mathbb{S}\cdot\left|\tfrac{\partial \boldsymbol{x}_1}{\partial \boldsymbol{x}_2}\right|\left|\tfrac{\partial \boldsymbol{x}_1}{\partial \boldsymbol{x}_0}\right|$$

whence

$$D(2,0) = \frac{D(2,1)D(1,0)}{\det \mathbb{S}}$$

which is the result we sought to establish.[53]

**Multiple classical paths: the particle-on-a-ring problem.** Let a particle $m$ be confined to the (let us say convex) interior of a domain bounded by a reflective barrier. Such a particle can proceed from point $\boldsymbol{x}_0$ to point $\boldsymbol{x}$ by a *direct* path, but can do so also by any number of indirect or *reflective paths*—paths that visit the barrier one or more times. Typically it is not possible to ennumerate the paths in any useful way,[54] but in favorable cases—*i.e.*, within domains of sufficiently regular design—it *is* possible to effect such an ennumeration: in such cases the Feynman formalism gives rise to a powerful "quantum mechanical method of images," the most characteristic rudiments of which I undertake now to describe.

A simple example is provided by the "particle-in-a-box problem" (what Einstein, in a dispute with Born, called the "ball-between-walls problem"). A particle is confined to the interval $0 \leqslant x \leqslant a$. It can proceed directly from $x_0$ to $x$, but will arrive there also if it heads off toward any of the reflective images of $x$. If it heads toward $2na + x$ $(n = 0, \pm1, \pm2, \ldots)$ it will arrive at $x$ after an even number of bounces, while if it heads toward $2na - x$ is will arrive after an odd number of bounces. The action functional $S[\text{path}]$ has acquired *multiple local extrema*, with which the various dynamical paths are associated: at $x$ the familiar 2-point action function has become *multi-valued*

$$S(x,t;x_0,t_0) = \frac{m}{2}\frac{(x-x_0)^2}{t-t_0} \quad\longrightarrow\quad S_n^{(\pm)}(x,t;x_0,t_0) = \frac{m}{2}\frac{(2na \pm x - x_0)^2}{t-t_0}$$

and the Hamilton-Jacobi equation $\frac{1}{2m}S_x^2 + S_t = 0$ is satisfied on each of its branches. Pauli was apparently the first to apply path integral methods to the particle-in-a-box problem,[55] though in 1971 Richard Crandall and I thought we were.

---

[53] The preceding argument was taken from TRANSFORMATIONAL PHYSICS & PHYSICAL GEOMETRY (1971–1983): "Semi-classical quantum theory," page 77, which was itself taken from some research notes written prior to 1958.

[54] A classic example is provided by the so-called "stadium problem," wherein a particle is allowed to bounce around inside a 2-dimensional "stadium" of roughly elliptical shape. Most trajectories are aperiodic, and trajectories that begin at $\boldsymbol{x}_0$ with slightly different velocities soon become wildly divergent. The model has become a favorite laboratory for studying classical/quantum *chaos*.

[55] See pages 170–172 in the 1950/51 lecture notes.[6] Pauli's objective was to demonstrate the accuracy of his technique for managing potentials (in this case, infinite wall potentials), but considered the example "interesting in its own right."

Even simpler than the one-dimensional particle-in-a-box problem is the "particle-on-a-ring problem," discussed below.

A mass $m$ moves freely on a ring—more generally: on a not-necessarily-circular loop—of circumference $a$. It encounters no obstacle, experiences no bouncing, becomes periodic and exhibits path multiplicity not because of the boundedness of the domain on which it moves, but because of its topology: the points $x$ and $x + na$ $(n = 0, \pm 1, \pm 2, \ldots)$ are physically identical. Quantum mechanically one has

$$-\tfrac{\hbar^2}{2m}\psi_{xx} = E\psi$$

and requires $\psi$ to be periodic: $\psi(x + a) = \psi(x)$.[56] Normalized solutions are $\psi(x) = \exp\{\tfrac{i}{\hbar}px\}$ with $p = \sqrt{2mE}$, and the periodicity condition enforces $pa = 2\pi n\hbar$. So we have

$$\psi_n(x) = \tfrac{1}{\sqrt{a}}e^{\frac{i}{\hbar}p_n x} \quad \text{with} \quad p_n \equiv nh/a$$

whence $E_n = \mathcal{E}n^2$ where $\mathcal{E} \equiv h^2/2ma^2$.[57] Notice that the ground state has become flat

$$\psi_0(x) = \tfrac{1}{\sqrt{a}} \quad \text{with} \quad E_0 = 0$$

and that the excited states are 2-fold degenerate:

$$E_{-n} = E_n \quad : \quad n = 1, 2, 3, \ldots$$

In those respects the ring problem differs markedly from the box problem.

The spectral representation of the propagator becomes

$$\begin{aligned} K(x, t; y, 0) &= \sum_{-\infty}^{+\infty} e^{-\frac{i}{\hbar}\mathcal{E}n^2 t}\tfrac{1}{a}e^{\frac{i}{\hbar}p_n(x-y)} \\ &= \tfrac{1}{a}\Big\{1 + 2\sum_{n=1}^{\infty} e^{-\frac{i}{\hbar}\mathcal{E}n^2 t}\cos\Big[2n\pi\frac{x-y}{a}\Big]\Big\} \end{aligned} \tag{85}$$

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

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

---

[56] In the particle-in-a-box problem one, on the other hand, requires

$$\psi(0) = \psi(a) = 0$$

[57] Angular momentum, by the way, is conserved *only if the constraining loop is circular*, and is given then by $(a/2\pi)p_n = n\hbar$.

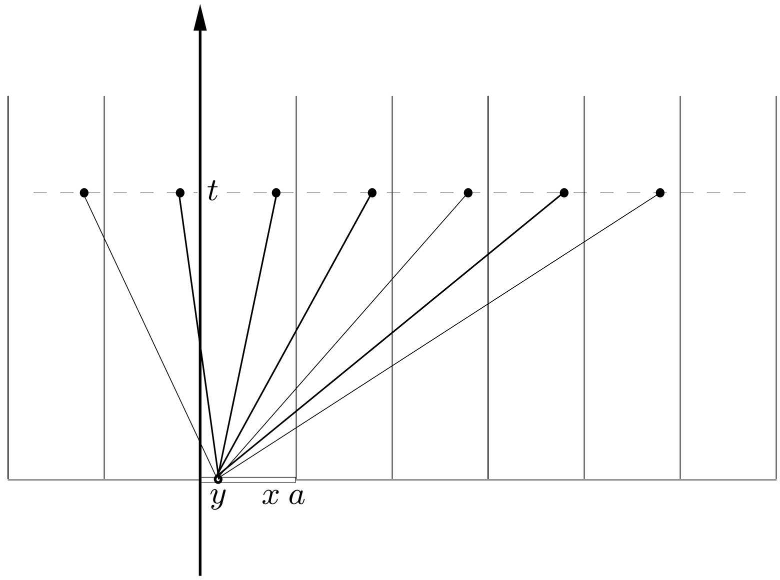



FIGURE 7: *On a loop of circumference $a$, with $x$ taken to mean arc length, the points $x$ and $x+na$ are physically identical. There are therefore an infinite number of distinct dynamical paths linking any specified pair of endpoints. The figure appears also as Figure 2 in* "Applied theta functions of one or several variables" (1997).

and its wonderful properties fill books, most accessibly Richard Bellman's *A Brief Introduction to Theta Functions* (1961). In this notation (85) reads

$$K(x,t;y,0) = \tfrac{1}{a}\vartheta_3(z,\tau) \tag{86}$$

with

$$z = \pi\tfrac{x-y}{a} \quad \text{and} \quad \tau = -\tfrac{\mathcal{E}t}{\pi\hbar} = -\tfrac{2\pi\hbar t}{ma^2}$$

On the other hand, Feynman's path integral method—interpreted in this instance (since the particle moves freely) to mean not summation over all conceivable paths but summation over all classical paths (free motion directed from $y$ to all the images of $x$, as displayed in the figure)—immediately supplies

$$\begin{aligned} K(x,t;y,0) &= \sqrt{\tfrac{m}{iht}}\sum_{-\infty}^{+\infty}\exp\Big\{\tfrac{im}{\hbar 2t}(x+na-y)^2\Big\} \\ &= \sqrt{\tfrac{m}{iht}}\,\exp\Big\{\tfrac{i}{\hbar}\tfrac{m}{2t}(x-y)^2\Big\}\cdot\sum_{-\infty}^{+\infty}e^{i(\pi\tau n^2-2nz)} \\ &= \text{———— ditto ————}\cdot\vartheta_3(z,\tau) \\ &= \tfrac{1}{a}\sqrt{\tau/i}\,e^{iz^2/\pi\tau}\vartheta_3(z,\tau) \end{aligned} \tag{87}$$

with

$$\tau = \tfrac{ma^2}{2\pi\hbar t} = -\tfrac{1}{\tau} \quad \text{and} \quad z = -\tfrac{ma}{2\hbar t}(x-y) = \tfrac{z}{\tau}$$

The expressions on the right sides of (86) and (87) do not much resemble each other. BUT the theory of theta functions supplies a zillion wonderful identities[58] . . . among them this one

$$\vartheta_3(z,\tau) = \sqrt{i/\tau}\, e^{z^2/\pi i\tau}\vartheta_3(\tfrac{z}{\tau}, -\tfrac{1}{\tau}) \tag{88}$$

which is called the "Jacobi theta transformation" (or "Jacobi's identity," when confusion with other wonders of that name is unlikely), concerning which Bellman remarks that

> *". . . it has amazing ramifications in the fields of algebra, number theory, geometry and other parts of mathematics. In fact, it is not easy to find another identity of comparable significance."*

In the present application (to the free-particle-on-a-ring problem)

> Jacobi's identity asserts the identity of the spectral representation (86) and the path integral representation (87)—of what Born calls the "wave representation" and the "particle representation"—of the propagator.

Note the sense in which $\tau$ lives upstairs on the left side, but downstairs on the right side of (88) . . . with consequences that can be quite dramatic. It is, for example, an implication of (88) that (set $\tau = it/\pi$ and $z = 0$)

$$f(t) \equiv \sum_{-\infty}^{+\infty} e^{-tn^2} = \sqrt{\tfrac{\pi}{t}} \sum_{-\infty}^{+\infty} e^{\pi^2 n^2/t}$$

While the left and right sides of the preceding equation are identically equal, they are *not computationally identical!* For suppose we wanted to evaluate $f(.01)$: working from the sum on the left, it follows from

$$e^{-25} \approx 10^{-10.8}$$

that we would (since $n^2/100 = 25$ entails $n = 50$) have to keep about 50 terms to achieve 10-place accuracy. Working, on the other hand, from the sum on the right, we have

$$f(.01) = \sqrt{100\pi}(1 + 2\underbrace{e^{-100\pi^2}}_{\approx 10^{-434}} + \cdots)$$

[58] See, for example, Chapter 21 in E. T. Whittaker & G. N. Watson, *Modern Analysis* (4th edition, 1952), especially §**21.51**. The proof of (88) elegantly simple: see §4 in "2-dimensional 'particle-in-a-box' problems in quantum mechanics" (1997), where I sketch also a theory of *theta functions of several variables* and the corresponding generalization of (88).

and have achieved accuracy to better than 400 places with only two terms! The situatiuon would be reversed if we were to evaluate $f(100)$. Physically, $t$ enters upstairs (through terms of the form $\exp\{-\frac{i}{\hbar}E_n t\}$) into the design of the spectral representation of the propagator, but downstairs (through the Van Vleck determinant and terms of the form $\exp\{\frac{i}{\hbar}(x-y)^2/t\}$) into the design of the path integral representation. We might therefore

- expect the part integral representation to be more useful when $t$ is small;
- expect the spectral representation to be more useful when $t$ is large

but the presence of the $i$'s clouds the issue, since $e^{-\omega t}$ and $e^{i\omega t}$ have entirely different asymptotic properties.

The essentials of the preceding discussion are by no means special to the ring problem. For the oscillator one has

$$K_{\text{spectral}} = \sum_0^\infty e^{-i\omega(n+\frac{1}{2})t}\psi_n(x)\psi_n(y)$$
$$\psi_n(x) = \left(\tfrac{2m\omega}{h}\right)^{\frac{1}{4}}\tfrac{1}{\sqrt{2^n n!}}e^{-\frac{1}{2}(m\omega/\hbar)x^2}H_n\left(\sqrt{\tfrac{m\omega}{\hbar}}x\right)$$
$$K_{\text{path integral}} = \sqrt{\tfrac{m\omega}{ih\sin\omega t}}\exp\left\{\tfrac{i}{\hbar}\tfrac{m\omega}{2\sin\omega t}\left[(x^2+y^2)\cos\omega t - 2xy\right]\right\}$$

and the equivalence follows[59] from an obscure but pretty identity known as "Mehler's formula:"

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

To the extent that Feynman has succeeded in constructing an alternative to the standard formulation of quantum mechanics—to the extent, that is to say, that we are presented with alternative but equivalent formulations of the propagator

$$\boxed{K_{\text{spectral}}\,,\text{ with } t \text{ upstairs}} \Longleftrightarrow \boxed{K_{\text{path integral}}\,,\text{ with } t \text{ downstairs}}$$

—to that extent we expect there to exist a "super identity" that serves to link one formulation with the other, a super identity of which Jacobi's and Mehler's identities are particular instances. And, though here motivated by physics, we expect the super identity to issue from pure mathematics, from (it would appear) the theory of partial differential equations.

[59] For the details, see "Jacobi's theta transformation & Mehler's formula: their interrelation, and their role in the quantum theory of angular momentum" (2000). F. G. Mehler (1835–1895) published his result in 1866—sixty years before it acquired quantum mechanical work to do. In QUANTUM MECHANICS (1967) I describe how Mehler's formula can be used to *prove the completeness* of the oscillator eigenfunctions (see Chapter 2, pages 64–65)—something that, David Griffiths has remarked, is more often talked about (assumed) than done.

The methods described here as they relate to the ring problem were—as already remarked—applied to the one-dimensional particle-in-a-box problem by Pauli and by Born & Ludwig[60] in the early 1950's, and by me to a number of exceptionally tractable two-dimensional particle-in-a-box problems in the early 1970's.[61] Suppose, for example, that a particle moves freely within the triangular box shown below:

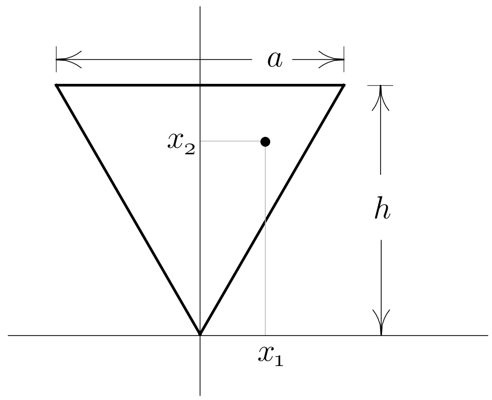



FIGURE 8: *An otherwise free particle is confined to the interior of an equilateral triangular box. The problem is to solve*

$$-\tfrac{\hbar^2}{2m}\nabla^2\psi = E\psi$$

*subject to the conditions* $\psi(\partial\triangledown) = 0$ *and* $\iint_\triangledown |\psi(x_1,x_2)|^2\,dx_1dx_2 = 1$.

Working from Figure 9, one is able to

- ennumerate,
- assign a classical action to, and
- sum over

the image paths to construct an exact description of $K_\text{path}(\boldsymbol{x},t;\boldsymbol{y},0)$. One can then use a 2-dimensional generalization of Jacobi's identity to construct

$$K_\text{spectral}(\boldsymbol{x},t;\boldsymbol{y},0) = \sum_{\boldsymbol{n}} e^{-\frac{i}{\hbar}E(\boldsymbol{n})t}\psi_{\boldsymbol{n}}(\boldsymbol{x})\psi^*_{\boldsymbol{n}}(\boldsymbol{y})$$

from which the eigenvalues and eigenfunctions can then be read off. The energy

---

[60] M. Born & W. Ludwig, "Zur quantenmechanik des kräftfreien Teilchens," Z. Physik **150**, 106 (1958).

[61] That old work has been revisited and expanded in two lengthy recent essays: "2-dimensional 'particle-in-a-box' problems in quantum mechanics. Part I: Propagator & eigenfunctions by the method of images" (1997) and "Applied theta functions of one or several variables" (1997).

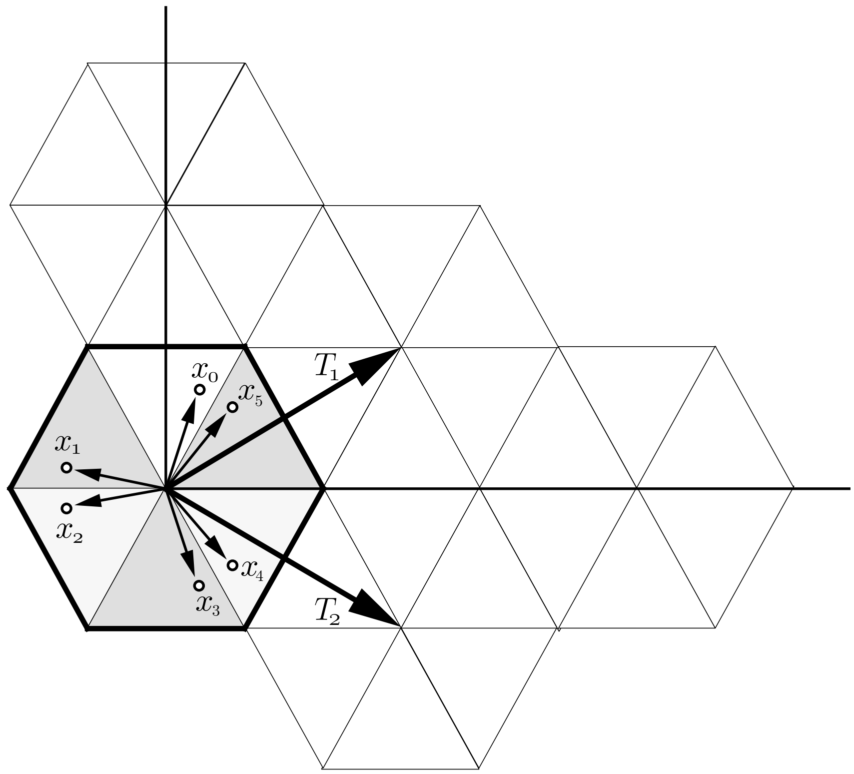



FIGURE 9: *Identification of the six elements that are taken to comprise the "fundamental unit," names assigned to the associated "fundamental images" of the physical target point $\boldsymbol{x} \equiv \boldsymbol{x}_0$, and the (non-orthogonal) translation vectors $\boldsymbol{T}_1$ and $\boldsymbol{T}_2$ that serve to replicate the fundamental unit. White cells are even (in the sense that they give rise to paths with an even number of reflection points), and shaded cells are odd.*

eigenvalues out to be describable

$$E_{\hat{\boldsymbol{n}}} = \hat{\mathcal{E}}(\hat{n}_1^2 + 3\hat{n}_2^2) \quad \text{with} \quad \hat{\mathcal{E}} \equiv \tfrac{1}{18}\tfrac{h^2}{ma^2}$$

where $\hat{n}_1$ and $\hat{n}_2$ are integers (either both even or both odd) drawn from the shaded sector in Figure 10. The associated eigenfunction is

$$\psi_{\hat{\boldsymbol{n}}}(\boldsymbol{x}) = \sqrt{\tfrac{8}{6\,\text{area}}}\Big\{G_{\hat{\boldsymbol{n}}}(\boldsymbol{x}) + iF_{\hat{\boldsymbol{n}}}(\boldsymbol{x})\Big\}$$

where

$$\begin{aligned} G_{\hat{\boldsymbol{n}}}(\xi_1,\xi_2) &= \cos[2\hat{n}_1\xi_1]\sin[2\hat{n}_2\xi_2] + \cos[2\tfrac{-\hat{n}_1+3\hat{n}_2}{2}\xi_1]\sin[2\tfrac{-\hat{n}_1-\hat{n}_2}{2}\xi_2] \\ &\qquad + \cos[2\tfrac{-\hat{n}_1-3\hat{n}_2}{2}\xi_1]\sin[2\tfrac{+\hat{n}_1-\hat{n}_2}{2}\xi_2] \\ F_{\hat{\boldsymbol{n}}}(\xi_1,\xi_2) &= \sin[2\hat{n}_1\xi_1]\sin[2\hat{n}_2\xi_2] + \sin[2\tfrac{-\hat{n}_1+3\hat{n}_2}{2}\xi_1]\sin[2\tfrac{-\hat{n}_1-\hat{n}_2}{2}\xi_2] \\ &\qquad + \sin[2\tfrac{-\hat{n}_1-3\hat{n}_2}{2}\xi_1]\sin[2\tfrac{+\hat{n}_1-\hat{n}_2}{2}\xi_2] \end{aligned}$$

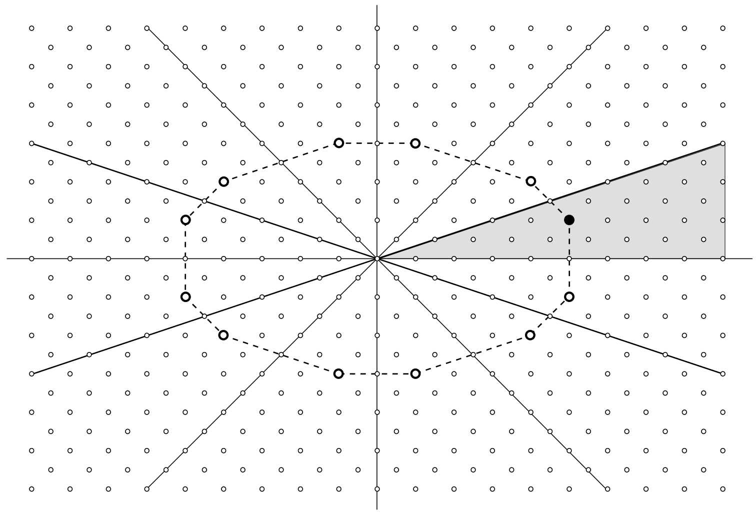


FIGURE 10: *The integers that label states in the equilateral box problem are drawn from the shaded sector of the equi-parity lattice. The significance of the polyhedral trace is explained in the essay*[61] *from which the figure was taken.*

with $\xi_1 \equiv \frac{\pi}{3a}x_1$ and $\xi_2 \equiv \frac{\pi}{3a}\sqrt{3}x_2$.

The interesting point is that while the eigenfunctions reported above demonstrably do satisfy the Schrödinger equation, they appear in this instance to be obtainable *only* by the method just described; *i.e.*, by $\sum_{\text{paths}}$, for the eigenfunctions are not of a form which could be achieved by separation of variables.

**Management of operators within the Feynman formalism.** Fundamental to the formal apparatus of quantum mechanics are self-adjoint operators $\mathbf{A}$, $\mathbf{B}$, ... (representative of "observables") and the algebraic relations among them (for example: $[\mathbf{x}, \mathbf{p}] = i\hbar\mathbf{I}$). In his "space-time approach to quantum mechanics" Feynman assigns a preferred role to the position operator(s) $\mathbf{x}$. Our assignment is to describe the placement of operators-in-general within such a biased formalism.

It is not by their naked selves but indirectly, through constructions of the form $(\alpha|\mathbf{A}|\beta)$, that operators engage the world of physical experience. Occupying a special place within that population of complex numbers are the real numbers $(\psi|\mathbf{A}|\psi)$ that we call "expectation values." To describe the dynamical time-dependence of such numbers we write

$$(\alpha|\mathbf{A}|\beta) \longrightarrow (\alpha|\mathbf{U}^{-1}(t)\,\mathbf{A}\,\mathbf{U}(t)|\beta) \quad \text{with} \quad \mathbf{U}^{-1}(t) = \mathbf{U}^{+}(t)$$

which in the Schrödinger picture we attribute to motion of state vectors

$$|\beta)_0 \longrightarrow |\beta)_t = \mathbf{U}(t)|\beta)_0 \quad \text{while } \mathbf{A} \text{ just sits there}$$

Feynman, however, finds (though he nowhere says so in plain words) that his formal needs are best served by the conceptual and notational resources provided by the Heisenberg picture, where

$$\mathbf{A}(0) \longrightarrow \mathbf{A}(t) = \mathbf{U}^{-1}(t)\,\mathbf{A}(0)\,\mathbf{U}(t) \quad \text{while } |\beta) \text{ sits there}$$

*Moving operators drag their eigenbases with them*: suppose $\mathbf{A}|a) = a|a)$, which in the more explicit notation that has been forced upon us reads

$$\mathbf{A}(0)|a,0) = a|a,0)$$

This can be written $\mathbf{U}(t)\mathbf{A}(t)\mathbf{U}^{-1}|a,0) = a|a,0)$, which on multiplication by $\mathbf{U}^{-1}(t)$ becomes

$$\mathbf{A}(t)|a,t) = a|a,t) \quad \text{with} \quad |a,t) = \mathbf{U}^{-1}(t)|a,0)$$

*The motion of the dragged eigenbasis is retrograde.* Particularize $\mathbf{A} \mapsto \mathbf{x}$, writing

$$\mathbf{x}(t)|x,t) = x|x,t)$$
$$|x,t) = \mathbf{U}^{-1}(t)|x,0) \quad \text{and} \quad (x,t| = (x,0|\mathbf{U}(t)$$

In that notation the equation $K(x,t;y,0) \equiv (x|\mathbf{U}(t)|y)$ that heretofore has served to define the propagator becomes

$$K(x,t;y,0) = (x,t|y,0)$$

It is a familar proposition, fundamental to the quantum theory, that

$$(a|\psi) = \begin{cases} \text{probability amplitude that an } \mathbf{A}\text{-measurement, performed} \\ \text{on a system in state } |\psi), \text{ will return the result ``}a\text{''} \end{cases}$$

If we wait a time $t$ this complex number will have changed, from $(a,0|\psi)$ to $(a,t|\psi)$. Feynman is motivated to introduce some process-oriented terminology:

$$\begin{aligned}(a,t|\psi) &= \begin{cases} \text{``transition amplitude'' that the system will go} \\ \text{in time } t \text{ from state } |\psi) \text{ to state } |a) \end{cases} \\ &= (a|\mathbf{U}(t)|\psi) \\ &= \iint (a|x)dx\,\underbrace{(x,t|y,0)}_{\text{propagator}}\,dy(y|\psi)\end{aligned}$$

The "process-oriented" bias built into Feynman's preferred language becomes plain with the realization that one could equally well say

$$= \begin{cases} \text{probability amplitude that an } \mathbf{A}\text{-measurement, performed} \\ \underline{\text{on the evolved system}}, \text{ will return the result ``}a\text{''} \end{cases}$$

or (in the Schrödinger picture)

$$= \text{weight of } |a)\text{-component acquired by } |\psi)_t$$

To describe the moving expectation value $\langle \mathbf{A} \rangle_t = (\psi|\mathbf{A}(t)|\psi)$ Feynman makes devious use of the "mixed representation trick": he picks a pair of times $t_1$ and $t_0$ that straddle $t$ and writes

$$\langle \mathbf{A} \rangle_t = \iint (\psi|x,t_1)\,dx(x,t_1|\mathbf{A}(t)|y,t_0)\,dy(y,t_0|\psi) \quad : \quad t_1 > t > t_0$$

and, more generally,

$$(\alpha|\mathsf{A}(t)|\beta) = \iint (\alpha|x,t_1)\,dx(x,t_1|\mathsf{A}(t)|y,t_0)\,dy(y,t_0|\beta) \quad : \quad t_1 > t > t_0$$

Insert

$$(x,t_1|\mathsf{A}(t)|y,t_0) = \iint (x,t_1|x,t)\,dx\,(x,t|\mathsf{A}(t)|y,t)\,dy\,(y,t|y,t_0)$$

and obtain

$$(\alpha|\mathsf{A}(t)|\beta) \tag{89}$$
$$= \iiiint \underbrace{(\alpha|x,t_1)\,dx(x,t_1|x,t)\,dx\,(x,t|}_{\text{state that evolves to }(\alpha|\text{ as }t_1\leftarrow t}\mathsf{A}(t)\underbrace{|y,t)\,dy\,(y,t|y,t_0)\,dy(y,t_0|\beta)}_{\text{state that evolves from }|\beta)\text{ as }t\leftarrow t_0}$$

This construction provides the platform upon which Feynman proceeds to build (and provides evidence of how literally/seriously understood his objective, which was to devise a "space-time formulation of ... quantum mechanics").

Specialize: $\mathsf{A}(t) \longmapsto \mathsf{F}(t) \equiv F(\mathsf{x}(t))$. Use

$$(x,t|F(\mathsf{x}(t))|y,t) = F(x)\,\delta(x-y)$$

in (89) to obtain the matrix element

$$(\alpha|F(\mathsf{x}(t))|\beta) \tag{90}$$
$$= \iint (\alpha|x,t_1)\,dx\Big\{\int (x,t_1|x,t)F(x)dx(x,t|y,t_0)\Big\}\,dy(y,t_0|\beta)$$

which Feynman calls the "transition element" between the daughter of $|\beta)$ and the mother of $(\alpha|$. In orthodox notation the meaning of the transition element is not at all obscure:

$$= \int \alpha^*(x,t)F(x)\beta(x,t)\,dx$$

But by placing "breathing room" on both sides of $t$ Feynman has introduced an expression $\{\int \text{etc.}\}$ to which he is able to assign a very interesting interpretation (see the following figure):

$$\Big\{\int (\text{etc.})\,dx\Big\} = \int \Big\{\int_{x\leftarrow x} e^{\frac{i}{\hbar}S[\text{path}]}\mathcal{D}[\text{paths}]\Big\}F(x)\Big\{\int_{x\leftarrow y} e^{\frac{i}{\hbar}S[\text{path}]}\mathcal{D}[\text{paths}]\Big\}dx$$
$$= \int_{x\leftarrow y}\mathcal{F}[\text{path}]e^{\frac{i}{\hbar}S[\text{path}]}\mathcal{D}[\text{paths}] \tag{91}$$
$$\mathcal{F}[\text{path}]\begin{cases}\text{looks at the path } x(u) : t_0 < u < t_1\\ \text{and announces } F(x(t))\end{cases}$$

where a normalization factor has been absorbed into the meaning of $\mathcal{D}$[paths].

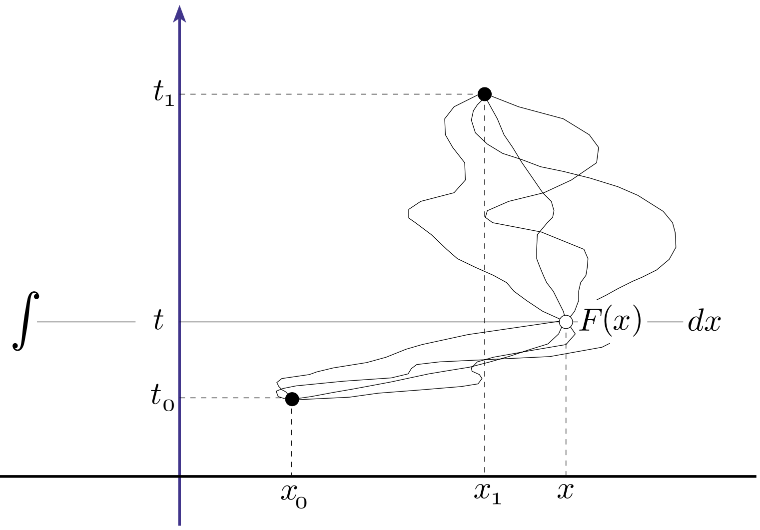



FIGURE 11: *Diagramatic interpretation of (90) that gives rise to the functional integral (91).*

Equation (91) captures the bracketed essence of (90); it renders what Feynman calls the "transition element" $(x|\mathsf{U}(t_1,t)\mathsf{F}(t)\mathsf{U}(t,t_0)|y)$—abbreviated $\langle\mathsf{F}\rangle_S$ —as a "functional integral"[62]

$$\langle\mathsf{F}\rangle_S = \int_{x\leftarrow y} \mathcal{F}[\text{path}]e^{\frac{i}{\hbar}S[\text{path}]}\mathcal{D}[\text{paths}] \tag{92}$$

and gives back Feynman's fundamental postulate/insight in the case $\mathsf{F}(t) = \mathsf{I}$.

The more general construction $(x|\mathsf{U}(t_1,t'')\mathsf{F}(t'')\mathsf{U}(t'',t')\mathsf{G}(t')\mathsf{U}(t',t_0)|y)$ makes quantum mechanical good sense only if $t_1 \geqslant t'' \geqslant t' \geqslant t_0$. Define the

$$\text{“chronological product” } \mathcal{P}\big\{\mathsf{F}(t'')\mathsf{G}(t')\big\} \equiv \begin{cases} \mathsf{F}(t'')\mathsf{G}(t') & \text{if } t'' > t' \\ \mathsf{G}(t')\mathsf{F}(t'') & \text{if } t'' < t' \end{cases}$$

The argument that gave (91) then gives

$$\langle\mathcal{P}\big\{\mathsf{F}\,\mathsf{G}\big\}\rangle_S = \int_{x\leftarrow y} \mathcal{F}[\text{path}]e^{\frac{i}{\hbar}S[\text{path}]}\mathcal{D}[\text{paths}]$$

where now $\mathcal{F}[\text{path}]$ looks at the path $x(u)$ and announces $F(x(t''))G(x(t'))$. This demonstrates the rebustness of the functional integral concept, but...

*Why should we have interest in chronological products?* Motivation comes from the simplest aspects of the theory of coupled differential equations. Look to

$$\tfrac{d}{dt}\boldsymbol{x} = \mathbb{M}\,\boldsymbol{x} \quad \text{with} \quad \boldsymbol{x}_0 \equiv \boldsymbol{x}(0) \text{ specified}$$

---

[62] See Chapter 7 of Feynman & Hibbs.

or—which is the same—to

$$\boldsymbol{x}(t) = \boldsymbol{x}_0 + \int_0^t \mathbb{M}\,\boldsymbol{x}(u)\,du$$

The immediate solution is

$$\boldsymbol{x}(t) = e^{\mathbb{M}\,t}\boldsymbol{x}_0$$

which, however, fails if $\mathbb{M}$ is allowed to be itself variable. The system

$$\boldsymbol{x}(t) = \boldsymbol{x}_0 + \int_0^t \mathbb{M}(u)\,\boldsymbol{x}(u)\,du$$

can, however, be solved by iteration: one is led to

$$\boldsymbol{x}(t) = \left\{\mathbb{I} + \int_0^t \mathbb{M}(u)\,du + \int_0^t\!\!\int_0^u \underbrace{\mathbb{M}(u)\mathbb{M}(v)}_{\text{NOTE the spontaneous time-ordering}}\,dudv + \cdots\right\}\boldsymbol{x}_0$$

which by a little trickery becomes

$$= \left\{\mathbb{I} + \int_0^t \mathbb{M}(u)\,du + \tfrac{1}{2!}\underbrace{\int_0^t\!\!\int_0^t}_{\text{NOTE the identical upper limits}} \mathcal{P}\big\{\mathbb{M}(u)\mathbb{M}(v)\big\}\,dudv + \cdots\right\}\boldsymbol{x}_0$$

$$\equiv \mathcal{P}\left\{\exp\int_0^t \mathbb{M}(u)\,du\right\}\boldsymbol{x}_0$$

and gives back $e^{\mathbb{M}\,t}\boldsymbol{x}_0$ when $\mathbb{M}(u)$ is constant. We are, on the basis of these remarks, not surprised to discover that the chronological product is a tool that sees heavy service in time-dependent perturbation theory

$$i\hbar\tfrac{\partial}{\partial t}|\psi) = \big\{\mathbf{H}_0 + \lambda\mathbf{V}(t)\big\}|\psi)$$

And that, in fact, is the application which ostensibly motivates the discussion in Chapter 7 of Feynman & Hibbs.

But Feynman's interest in time-ordered operators is motivated also by a second, more fundamental consideration: he finds it natural to read the operator product $\mathbf{AB}$ as "first $\mathbf{B}$ then $\mathbf{A}$;" *i.e.*, as symbolic of measurements performed in *temporal sequence*, however brisk. I turn now to discussion of how that idea is implemented in illustrative cases.

We look first to $\mathbf{x}\,\mathbf{p} - \mathbf{p}\,\mathbf{x} = i\hbar\mathbf{I}$, which in most formulations of quantum mechanics is postulated, but in Feynman's formulation has the status of a *deduced consequence* of postulates invested elsewhere. The statement

$$\mathbf{x}\,\mathbf{p} - \mathbf{p}\,\mathbf{x} = i\hbar\mathbf{I} \quad\Longleftrightarrow\quad F(\mathbf{x})\,\mathbf{p} - \mathbf{p}\,F(\mathbf{x}) = i\hbar F'(\mathbf{x})$$

if $F(x)$ is sufficiently nice (developable as a power series). In the Heisenberg picture we therefore have

$$F(\mathbf{x}(t))\,\mathbf{p}(t) - \mathbf{p}(t)F(\mathbf{x}(t)) = i\hbar F'(\mathbf{x}(t)) \tag{93}$$

and it is actually *this* statement that Feynman undertakes to reproduce. To that end he examines

$$\big\langle \tfrac{\partial F(\mathsf{x})}{\partial \mathsf{x}} \big\rangle_S \equiv \int (x_1,t_1|x,t)\tfrac{\partial F(x)}{\partial x}(x,t|x_0,t_0)\,dx$$

Integration by parts gives

$$\big\langle \tfrac{\partial F(\mathsf{x})}{\partial \mathsf{x}} \big\rangle_S = (\text{boundary term}) - \int F(x)\tfrac{\partial}{\partial x}\big\{(x_1,t_1|x,t)(x,t|x_0,t_0)\big\}\,dx$$

Discard the boundary term on grounds that it can make no contribution to $\iint \alpha(x_1,t_1)\,dx_1(\text{boundary term})\,dx_0\beta(x_0,t_0)$ if the states $|\alpha)$ and $|\beta)$ satisfy typical boundary conditions. Then

$$\begin{aligned} &= -\int \tfrac{\partial}{\partial x}(x_1,t_1|x,t)\cdot F(x)(x,t|x_0,t_0)\,dx \\ &\qquad - \int (x_1,t_1|x,t)F(x)\cdot\tfrac{\partial}{\partial x}(x,t|x_0,t_0)\,dx \end{aligned}$$

Write

$$\begin{aligned} \tfrac{\partial}{\partial x}(x,t|x_0,t_0) &= \int \tfrac{\partial}{\partial x}\underbrace{(x,t|y,t-\tau)}_{=\,\frac{1}{A(\tau)}\exp\left\{\frac{i}{\hbar}\left[\frac{m}{2\tau}(x-y)^2-\tau V(x)\right]\right\}}dy(y,t-\tau|x_0,t_0) \\ &= \tfrac{i}{\hbar}\int \Big\{\tfrac{m}{\tau}(x-y)-\tau V'(x)\Big\}(x,t|y,t-\tau)\,dy(y,t-\tau|x_0,t_0) \\ &= \tfrac{i}{\hbar}\Big\{\tfrac{m}{\tau}x-\tau V'(x)\Big\}(x,t|x_0,t_0) \\ &\qquad - \tfrac{i}{\hbar}\int (x,t|y,t-\tau)\Big\{\tfrac{m}{\tau}y\Big\}\,dy(y,t-\tau|x_0,t_0) \\ \tfrac{\partial}{\partial x}(x_1,t_1|x,t) &= \int (x_1,t_1|y,t+\tau)\,dy\tfrac{\partial}{\partial x}\underbrace{(y,t+\tau|x,t)}_{\frac{1}{A(\tau)}\exp\left\{\frac{i}{\hbar}\left[\frac{m}{2\tau}(y-x)^2-\tau V(y)\right]\right\}} \\ &= -\tfrac{i}{\hbar}\int \Big\{\tfrac{m}{\tau}(y-x)\Big\}(x_1,t_1|y,t+\tau)\,dy(y,t+\tau|x,t) \\ &= +\tfrac{i}{\hbar}(x_1,t_1|x,t)\Big\{\tfrac{m}{\tau}x\Big\} \\ &\qquad - \tfrac{i}{\hbar}\int (x_1,t_1|y,t+\tau)\Big\{\tfrac{m}{\tau}y\Big\}\,dy(y,t+\tau|x,t) \end{aligned}$$

and note the use made here of Feynman's fundamental postulate (also that it has entered in a system-specific way: $L = \frac{1}{2}m\dot{x}^2 - V(x)$). We now have

$$
\begin{aligned}
i\hbar\big\langle\tfrac{\partial F(\mathsf{x})}{\partial\mathsf{x}}\big\rangle_S = &-\iint(x_1,t_1|y,t+\tau)\Big\{\tfrac{m}{\tau}y\Big\}\,dy(y,t+\tau|x,t)F(x)\,dx(x,t|x_0,t_0)\\
&+\int(x_1,t_1|x,t)\Big\{\tfrac{m}{\tau}x\Big\}F(x)\,dx(x,t|x_0,t_0)\\
&+\int(x_1,t_1|x,t)F(x)\Big\{\tfrac{m}{\tau}x\Big\}\,dx(x,t|x_0,t_0)\\
&-\iint(x_1,t_1|x,t)F(x)\,dx(x,t|y,t-\tau)\Big\{\tfrac{m}{\tau}y\Big\}\,dy(y,t-\tau|x_0,t_0)\\
&-\int(x_1,t_1|x,t)\Big\{\tau V'(x)\Big\}F(x)\,dx(x,t|x_0,t_0)\\
=&-(x_1,t_1|\underbrace{m\frac{\mathsf{x}(t+\tau)-\mathsf{x}(t)}{\tau}\cdot F(\mathsf{x}(t))-F(\mathsf{x}(t))\cdot m\frac{\mathsf{x}(t)-\mathsf{x}(t-\tau)}{\tau}}_{\text{NOTE the chronological order}}\\
&+\tau V'(\mathsf{x}(t))F(\mathsf{x}(t))|x_0,t_0)
\end{aligned}
\tag{94}
$$

Which brings us to a characteristic feature of the "space-time formulation of ... quantum mechanics:" Feynman considers momentum to be a derived concept, and its meaning to be system-dependent:[63]

$$\mathsf{p}(t)\equiv\lim_{\tau\downarrow 0} m\frac{\mathsf{x}(t+\tau)-\mathsf{x}(t)}{\tau}$$

By this interpretation (94) becomes, as $\tau\downarrow 0$,

$$
\begin{aligned}
i\hbar\big\langle\tfrac{\partial F(\mathsf{x})}{\partial\mathsf{x}}\big\rangle_S &= (x_1,t_1|\big\{F(\mathsf{x})\mathsf{p}-\mathsf{p}F(\mathsf{x})\big\}|x_0,t_0)\\
&=\big\langle\big\{F(\mathsf{x})\mathsf{p}-\mathsf{p}F(\mathsf{x})\big\}\big\rangle_S
\end{aligned}
$$

which, since valid for all endstates, entails

$$F(\mathsf{x})\mathsf{p}-\mathsf{p}F(\mathsf{x})=i\hbar F'(\mathsf{x})$$

and in the case $F(x)=x$ becomes (if we restore the notationally surpressed $t$)

$$\mathsf{x}(t)\mathsf{p}(t)-\mathsf{p}(t)\mathsf{x}(t)=i\hbar\mathbf{I}\tag{95}$$

The argument shows clearly the mechanism by which *operator order arises as an expression of temporal order*.

Back up to (94) and set $F(x)=1$. Divide by $\tau$ and obtain

$$0=-(x_1,t_1|m\frac{\frac{\mathsf{x}(t+\tau)-\mathsf{x}(t)}{\tau}-\frac{\mathsf{x}(t)-\mathsf{x}(t-\tau)}{\tau}}{\tau}+V'(\mathsf{x}(t))|x_0,t_0)$$

---

[63] As, indeed, so also does Lagrange: $p\equiv\partial L/\partial\dot{x}$.

which in the limit $\tau \downarrow 0$ becomes

$$\ddot{\mathsf{x}}(t) = -V'(\mathsf{x}(t)) \tag{96}$$

This is, in Feynman's phrase, "the matrix expression of Newton's law." Since (96), when spelled out, means

$$(\alpha|\,\ddot{\mathsf{x}}(t)|\beta) = (\alpha| - V'(\mathsf{x}(t))|\beta) \quad : \quad \text{all } |\alpha) \text{ and } \beta)$$

it might better be called "Ehrenfest's theorem in the Heisenberg picture."

We have recently been discussing properties of path functionals of a form first encountered at (92)

$$\langle \mathcal{F} \rangle_S = \int \mathcal{F}[\text{path}]\, e^{\frac{i}{\hbar}S[\text{path}]}\mathcal{D}[\text{paths}]$$

where it is now understood that the paths in question link $(x_1, t_1) \leftarrow (x_0, t_0)$. We consider such objects now from a somewhat more general point of view (and will formally disregard the fact that the functionals $\mathcal{F}[\text{path}]$ encountered in preceding arguments were of specialized design). Clearly the *set* of paths is *invariant* under

$$\text{path} \;\longrightarrow\; \text{path} + \epsilon\eta(t) \quad : \quad \eta(t_0) = \eta(t_1) = 0$$

Therefore

$$\langle \mathcal{F} \rangle_S = \int \mathcal{F}[\text{path} + \epsilon\eta(t)]\, e^{\frac{i}{\hbar}S[\text{path}+\epsilon\eta(t)]}\mathcal{D}[\text{paths}]$$

which, when the concept of "functional derivative" is brought into play,[64] becomes

$$\begin{aligned} &= \int \left\{ \mathcal{F}[\text{path}] + \epsilon \int \frac{\delta\mathcal{F}[\text{path}]}{\delta\eta(t)}\eta(t)\,dt + \cdots \right\} \\ &\qquad\qquad \cdot\, e^{\frac{i}{\hbar}\left\{S[\text{path}] + \epsilon\int\frac{\delta S\,[\text{path}]}{\delta\eta(t)}\eta(t)\,dt + \cdots\right\}}\mathcal{D}[\text{paths}] \\ &= \langle \mathcal{F} \rangle_S + \epsilon \int \left\{ \int \left[ \frac{\delta\mathcal{F}[\text{path}]}{\delta\eta(t)} + \tfrac{i}{\hbar}\mathcal{F}[\text{path}]\frac{\delta S\,[\text{path}]}{\delta\eta(t)} \right] \right. \\ &\qquad\qquad\qquad\qquad \left. \cdot\, e^{\frac{i}{\hbar}S[\text{path}]}\mathcal{D}[\text{paths}] \right\} \eta(t)\,dt + \cdots \end{aligned}$$

and this, since valid for all tickle functions $\eta(t)$, entails

$$\int \left\{ \frac{\delta\mathcal{F}[\text{path}]}{\delta\eta(t)} + \tfrac{i}{\hbar}\mathcal{F}[\text{path}]\frac{\delta S\,[\text{path}]}{\delta\eta(t)} \right\} e^{\frac{i}{\hbar}S[\text{path}]}\mathcal{D}[\text{paths}] = 0$$

---

[64] For a fairly detailed introduction to this subject, see Chapter 5: "Calculus of Functionals" in CLASSICAL FIELD THEORY (1999).

which can be written

$$\left\langle \frac{\delta \mathcal{F}[\text{path}]}{\delta \eta(t)} \right\rangle_S = -\tfrac{i}{\hbar} \left\langle \mathcal{F}[\text{path}] \frac{\delta S[\text{path}]}{\delta \eta(t)} \right\rangle_S \tag{97}$$

Classical mechanics supplies

$$S[\text{path} + \epsilon \eta(t)] = S[\text{path}] + \epsilon \int \left\{ \frac{\partial L}{\partial x(t)} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}(t)} \right\} \eta(t)\, dt + \cdots$$

so we have

$$\left\langle \frac{\delta \mathcal{F}[\text{path}]}{\delta \eta(t)} \right\rangle_S = -\tfrac{i}{\hbar} \left\langle \mathcal{F}[\text{path}] \left\{ \frac{\partial L}{\partial x(t)} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}(t)} \right\} \right\rangle_S$$

In the case $\mathcal{F}[\text{path}] \equiv 1$ we obtain

$$\left\langle \frac{\partial L}{\partial x(t)} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}(t)} \right\rangle_S = 0 \tag{98}$$

which is a pretty variant of (96).

*Noether's theorem*, as it is encountered in classical mechanics, describes the first-order response $\delta_\omega S[\text{path}]$ of the classical action to parameterized maps of various kinds (rotations, space and time translations, gauge transformations, *etc.*) and the conservation laws that result when in fact those maps describe symmetries of the action: $\delta_\omega S[\text{path}] = 0$. Feynman has placed us in position to translate that theory directly into quantum mechanics.[65]

Feynman and Hibbs, writing in 1965, remark (at page 173) that "Julian Schwinger has been investigating the formulation of quantum mechanics suggested by" (98). They refer to an elaborate theoretical ediface which I survey elsewhere in these notes, and which had, in fact, been essentially complete since the early 1950's.[66]

**Diffractive failure of the method of images.** Exact descriptions of the propagator are available in only a relatively few "textbook" cases. It is a remarkable fact that in the cases considered thus far $K(x_1, t_1; x_0, t_0)$, when rendered in the form suggested by the Feynman formalism, was found to involve summation not over "all conceivable paths," as the formalism contemplates, but only over the *classical* path or paths $(x_1, t_1) \leftarrow (x_0, t_0)$: in those cases

$$\sum_{\text{“all” paths}} \text{ “collapses” to } \sum_{\underline{\text{classical}} \text{ paths}}$$

---

[65] See QUANTUM MECHANICS (1967), Chapter 3 page 95 for details. The present discussion has drawn heavily on material presented there.

[66] See Julian Schwinger, "The theory of quantized fields. I," Phys Rev. **82**, 914 (1951) and "The theory of quantized fields. II," Phys Rev. **91**, 913 (1953)—both of which are reproduced in the Schwinger Collection[3]—and related papers reprinted in Schwinger's *Quantum Kinematics & Dynamics* (1991).

Early students of the Feynman formalism were led by such experience to inquire whether "collapse" might be a *general/universal* phenomenon. The following discussion, based on a paper by Richard Crandall,[67] will demonstrate that collapse—far from being the rule—is the rare exception.

Let a mass $m$ be confined by reflecting barriers to the interior of an open sector or "wedge" with vertex angle $\alpha$. We look first to the ordinary quantum mechanics of such a system: we solve the time-independent Schödinger equation, and use the information thus gained to assemble $K_{\text{spectral}}$. This we then compare to the $K_{\text{paths}}$ supplied by the method of images.

The Schrödinger equation reads

$$\nabla^2\psi + \tfrac{2mE}{\hbar^2}\psi = 0$$

$$\nabla^2 = \tfrac{1}{r}\tfrac{\partial}{\partial r}r\tfrac{\partial}{\partial r} + \tfrac{1}{r^2}\tfrac{\partial^2}{\partial\theta^2} \quad \text{in polar coordinates: } \begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$$

and upon separation $\psi = R(r)\cdot Y(\theta)$ becomes

$$\left\{\tfrac{d^2}{dr^2} + \tfrac{1}{r}\tfrac{d}{dr} + \tfrac{2mE}{\hbar^2} - \tfrac{\mu^2}{r^2}\right\}R(r) = 0 \tag{99.1}$$

$$\tfrac{d^2}{d\theta^2}Y(\theta) = -\mu^2 Y(\theta) \tag{99.2}$$

where the separation constant is called $-\mu^2$ in order to emphasize that, in view of the boundary conditions $Y(0) = Y(\alpha) = 0$, we have interest only in the oscillatory solutions of (99.2). Immediately

$$\begin{aligned} Y(\theta) &= (\text{constant})\cdot\sin\mu\theta \\ \mu &= n\tfrac{\pi}{\alpha} \quad : \quad n = 0, 1, 2, \ldots \end{aligned} \tag{100}$$

Return with this information to (99.1). Multiply by $r^2$ and obtain

$$\left\{\rho^2\tfrac{d^2}{d\rho^2} + \rho\tfrac{d}{d\rho} + \rho^2 - \mu^2\right\}\mathcal{R}(\rho) = 0 \quad \text{with} \quad \begin{cases} \rho \equiv r\sqrt{2mE/\hbar^2} \\ \mathcal{R}(\rho) \equiv R(\rho/\sqrt{2mE/\hbar^2}) \end{cases}$$

This is Bessel's equation, of which the interesting solutions[68] are $\mathcal{R}(\rho) = J_\mu(\rho)$.

[67] "Exact propagator for motion confined to a sector," J. Phys. A: Math. Gen. **16**, 513 (1982). During the late 1960's and early 1970's Richard (then a graduate student at MIT) and I cultivated (as best we could in those snail-mail days) a mutual interest in the Feynman formalism. During the late 1970's and early 1980's, after Richard had joined the Reed College faculty, we were able to resume that collaboration . . . working in adjoining offices, in the dead of night. One such night the "wedge problem" entered into our conversation . . . sent us to the men's room lugging a mirror, which we held against the mirrors there to gain a more vivid sense of how things would appear in a triangular barber shop. The lesson of that adventure was quickly/brilliantly worked out by Richard, and is reported in the paper cited above.

[68] See J. Spanier & K. Oldham, *Atlas of Functions* (1987), page 523.

So we have

$$R(r) = J_\mu(r\sqrt{2mE/\hbar^2})$$

where $\mu = n(\pi/\alpha)$ is discrete but—for the free particle on the open wedge as for the free particle on the unbounded plane—$E$ assumes *continuous* values, subject only to the constraint $E \geqslant 0$. We are led thus to "wedge eigenfunctions" of the form

$$\Psi_{E,\,\mu}(r,\theta) = (\text{constant})\cdot J_\mu(r\sqrt{2mE/\hbar^2})\sin(\mu\theta) \tag{101}$$

but must resolve several ticklish issues before we can make practical use of this information:

TICKLISH POINT #1 The eigenfunctions (101) are not normalizable on the wedge

$$\int_0^\infty\int_0^\alpha |\Psi_{E,\,\mu}(r,\theta)|^2\, rd\theta dr \text{ is undefined}$$

and therefore do not describe possible *states* of the system. They must be looked upon as the raw material from which normalized states (wavepackets) are assembled by superposition. This is not an uncommon situation. Were we studying free motion on the line we would at this point have constructed

$$\Psi_p(x) = (\text{constant})\cdot e^{\frac{i}{\hbar}px} \quad : \quad p = \pm\sqrt{2mE}$$

which are neither normalizable nor orthogonal an any literal sense:

$$\int_{-\infty}^{+\infty}\Psi_p^*(x)\Psi_q(x)\,dx \text{ is undefined for all } p \text{ and } q$$

We are, however, able to establish formal completeness ... by trickery: we construct

$$\int e^{-\lambda p^2}\Psi_p(x)\Psi_p^*(y)\,dp = |\text{constant}|^2\sqrt{\pi/\lambda}\exp\Big\{-\tfrac{(x-y)^2}{4\hbar^2\lambda}\Big\}$$

The expression on the right becomes a normalized Gaussian if we set

$$|\text{constant}|^2 = \tfrac{1}{2\pi\hbar}$$

That done, we have

$$\begin{aligned}&\downarrow\\ &= \delta(x-y) \quad \text{in the limit } \lambda\downarrow 0\end{aligned}$$

giving

$$\begin{aligned}\psi(x) &= \int\delta(x-y)\psi(y)\,dy\\ &= \int\Big\{\int\Psi_p(x)\Psi_p^*(y)\,dp\Big\}\psi(y)\,dy\\ &= \int\Psi_p(x)\Big\{\int\Psi_p^*(y)\psi(y)\,dy\Big\}dp\end{aligned}$$

from which formal orthonormality follows as a corollary:

$$\int \Psi_q^*(x)\psi(x)\,dx = \int \underbrace{\left\{\int \Psi_q^*(x)\Psi_p(x)\,dx\right\}}_{\delta(p-q)} \left\{\int \Psi_p^*(y)\psi(y)\,dy\right\}dp$$

Notice that it is completeness—not orthonormality—that lies nearer the heart of the matter. The wedge problem poses an identical formal difficulty, which can be resolved by identical means:

Crandall snooped through (my copy) of what he calls "the Russians"[69] and extracted this identity:

$$\int_0^\infty e^{-\lambda p^2} J_\mu(ap)J_\mu(bp)\,p\,dp = \tfrac{1}{2\lambda}\exp\left\{-\tfrac{a^2+b^2}{4\lambda}\right\}I_\mu\big(\tfrac{ab}{2\lambda}\big)$$

where the presumption is that $\Re[\mu] > -1$ and $|\arg p| < \frac{1}{4}\pi$, and $I_\mu(x)$ is the modified (or hyperbolic) Bessel function of order $\mu$.[70] He on this basis obtains

$$\begin{aligned} G_\mu &\equiv \int_0^\infty e^{-\lambda E} J_\mu(r\sqrt{2mE/\hbar^2})J_\mu(r_0\sqrt{2mE/\hbar^2})\,dE \\ &= \tfrac{1}{m}\int_0^\infty e^{-\lambda p^2/2m} J_\mu(rp/\hbar)J_\mu(r_0 p/\hbar)\,p\,dp \\ &= \tfrac{1}{\lambda}\exp\Big\{-\tfrac{m}{2\lambda\hbar^2}(r^2+r_0^2)\Big\}I_\mu\big(\tfrac{m}{2\lambda\hbar^2}2rr_0\big) \end{aligned}$$

We are informed that $I_\mu(x) \sim \frac{1}{\sqrt{2\pi x}}e^x$ as $x \to \infty$, so have

$$\begin{aligned} &\downarrow \\ &= \tfrac{\hbar}{\sqrt{2\pi m r r_0 \lambda}}\exp\Big\{-\tfrac{m}{2\lambda\hbar^2}(r^2-r_0^2)\Big\} \quad \text{as } \lambda \downarrow 0 \end{aligned}$$

It now follows (on the tentative assumption that the normalization constant is the same for all eigenfunctions) that

$$\begin{aligned} &\int_0^\infty e^{-\lambda E}\Big\{\sum_\mu \Psi_{E,\mu}(r,\theta)\Psi_{E,\mu}(r_0,\theta_0)\Big\}dE \\ &\qquad \sim |\text{constant}|^2\tfrac{\hbar}{\sqrt{2\pi m r r_0\lambda}}\exp\Big\{-\tfrac{m}{2\lambda\hbar^2}(r^2-r_0^2)\Big\}\cdot\sum_{\nu=0}^\infty \sin(\mu\theta)\sin(\mu\theta_0) \end{aligned}$$

---

[69] I. S. Gradshteyn & I. M. Ryzhik, *Table of Integrals, Series & Products* (1965). The identity in question appears as item **6.615**. G. N. Watson devotes §**13.31** in *Treatise on the Theory of Bessel Functions* (1966) to discussion—under the head "Weber's second exponential integral"— of this identity, which he attributes to H. Weber (1868). He reproduces a proof due to L. Gegenbauer (1876), and remarks that the identity occurs in A. Sommerfeld's dissertation: "Mathematische theorie der diffraction," Math. Ann. **47**, 317 (1896).

[70] See Spanier & Oldham,[68] Chapter 50.

which gives

$$\downarrow$$

$$= \tfrac{\hbar^2}{m}|\text{constant}|^2 \cdot \tfrac{1}{r}\delta(r - r_0) \cdot \sum_{\mu=0}^{\infty} \sin(\mu\theta)\sin(\mu\theta_0)$$

in the limit $\lambda \downarrow 0$. Notice that the $\mu$'s have detached themselves from the Bessel factors, and reside now only in the angular factors ...which Crandall subjects to this clever analysis: We have

$$\sin(\mu\theta)\sin(\mu\theta_0) = \tfrac{1}{2}\cos\mu(\theta - \theta_0) - \tfrac{1}{2}\cos\mu(\theta + \theta_0)$$

so

$$\sum_{\mu=0}^{\infty} \sin(\mu\theta)\sin(\mu\theta_0) = \tfrac{1}{4}\sum_{n=-\infty}^{\infty} e^{in(\pi/\alpha)(\theta-\theta_0)} - \tfrac{1}{4}\sum_{n=-\infty}^{\infty} e^{in(\pi/\alpha)(\theta+\theta_0)}$$

But the Poisson summation formula[71] supplies

$$\begin{aligned}\sum_{n=-\infty}^{\infty} e^{in(\pi/\alpha)\vartheta} &= \sum_{n=-\infty}^{\infty} \int_{-\infty}^{+\infty} e^{i[(\pi/\alpha)\vartheta - 2\pi n]y}\,dy \\ &= \sum_{n=-\infty}^{\infty} 2\pi\delta\big(\tfrac{\pi}{\alpha}[\vartheta - 2n\alpha]\big) \\ &= 2\alpha \sum_{n=-\infty}^{\infty} \delta(\vartheta - 2n\alpha)\end{aligned}$$

giving

$$\sin(\mu\theta)\sin(\mu\theta_0) = \tfrac{1}{2}\alpha\sum_{n=-\infty}^{\infty}\Big\{\delta(\theta - \theta_0 - 2n\alpha) - \delta(\theta + \theta_0 - 2n\alpha)\Big\} \tag{102}$$

Integrals of the form

$$\iint_{\text{wedge}} f(r,\theta)\, r\, dr\theta$$

see only a single one of those $\delta$-spikes: namely $\delta(\theta - \theta_0)$. Motivated by the detailed results now in hand, we assign

$$|\text{constant}| = \sqrt{\tfrac{2m}{\alpha\hbar^2}}$$

---

[71] See page 21 in "2-dimensional 'particle-in-a-box' problems in quantum mechanics: Part I" (1997) for discussion of this elegant formula, which in the general case reads

$$\sum_{n=-\infty}^{\infty} g(n) = \sum_{n=-\infty}^{\infty} \int_{-\infty}^{+\infty} g(y)e^{-2\pi i n y}\,dy$$

We are brought thus to the conclusion that if, on the wedge, we set

$$\Psi_{E,\mu}(r,\theta) = \sqrt{\tfrac{2m}{\alpha\hbar^2}}\,J_\mu(r\sqrt{2mE/\hbar^2}\,)\sin(\mu\theta) \tag{103}$$

then $\downarrow$

$$\lim_{\lambda\downarrow 0}\int_0^\infty e^{-\lambda E}\Big\{\sum_\mu \Psi_{E,\mu}(r,\theta)\Psi_{E,\mu}(r_0,\theta_0)\Big\}dE = \tfrac{1}{r}\delta(r-r_0)\delta(\theta-\theta_0) \tag{104}$$

Notice that the espressions on both left and right have dimension $(\text{length})^{-2}$.

TICKLISH POINT #2 A moment ago, when we looked to the free particle on a line, we found that we needed both of the eigenfunctions $\exp\{\pm i\sqrt{2mE/\hbar^2}\}$. Why had we no need of the functions $J_\mu(-r\sqrt{2mE/\hbar^2}\,)$ when discussing free motion on a wedge? Why were the functions (103) sufficient in themselves to permit assembly of the delta function; *i.e.*, to establish completeness? Because

$$\begin{aligned} J_\mu(-x) &= (-)^\mu J_\mu(x)\\ (-)^\mu &= \cos\mu\pi + i\sin\mu\pi \end{aligned}$$

The functions $\exp(\pm ix)$ are linearly independent, but the functions $J_\mu(\pm x)$ are not.

We have now only to make the replacement $\lambda \to \frac{i}{\hbar}(t-t_0)$ to obtain

$$K(r,\theta,t;r_0,\theta_0,t_0) = \int_0^\infty e^{-\frac{i}{\hbar}E\,(t-t_0)}\Big\{\sum_{\substack{\mu\\ \mu\equiv n(\pi/\alpha)}} \Psi_{E,\mu}(r,\theta)\Psi_{E,\mu}(r_0,\theta_0)\Big\}dE \tag{105}$$

which—since it satisfies the Schrödinger equation

$$-\tfrac{\hbar^2}{2m}\big\{\tfrac{1}{r}\tfrac{\partial}{\partial r}r\tfrac{\partial}{\partial r} + \tfrac{1}{r^2}\tfrac{\partial^2}{\partial\theta^2}\big\}K = i\hbar\tfrac{\partial}{\partial t}K$$

and possesses the property that

$$\lim_{t\downarrow t_0} K(r,\theta,t;r_0,\theta_0,t_0) = \tfrac{1}{r}\delta(r-r_0)\delta(\theta-\theta_0)$$

—must provide the *spectral description of the exact propagator* for free motion on a wedge.[72]

We are in position now to engage the "collapse problem." Let $m$ be confined to the upper half-plane; *i.e.*, to the interior of the wedge $\alpha = \pi$ (see Figure 12). Equation (105) then supplies

$$K(r,\theta,t;r_0,\theta_0,0) = \int_0^\infty e^{-\frac{i}{\hbar}E\,t}\Big\{\sum_{n=0}^\infty \Psi_{E,n}(r,\theta)\Psi_{E,n}(r_0,\theta_0)\Big\}dE \tag{106}$$

$$\Psi_{E,n}(r,\theta) = \sqrt{\tfrac{2m}{\pi\hbar^2}}\,J_n(r\sqrt{2mE/\hbar^2}\,)\sin(n\theta)$$

[72] Notice that we did all the work when we established completeness. *Proofs of completeness are rare for exactly the same reason that exact propagators are rare*: each supplies the other.

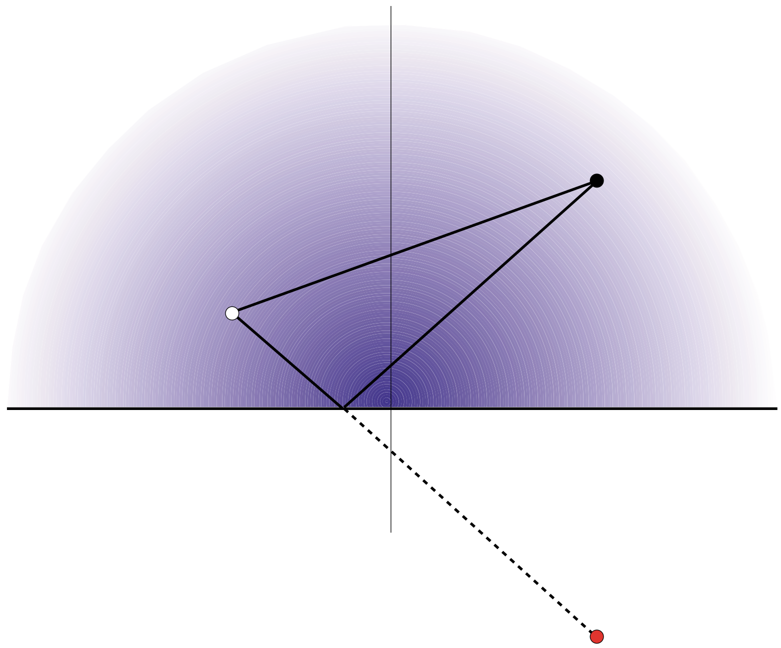


FIGURE 12: *The simplest instance of the wedge problem arises when $\alpha = \pi$, which entails confinement to the half-plane. Two paths link the source-point* $\circ$ *to the target-point* $\bullet$. *The direct path has length*

$$\ell_{\text{direct}} = (x - x_0)^2 + (y - y_0)^2 = r^2 + r_0^2 - 2\,rr_0\cos(\theta - \theta_0)$$

*while the indirect/reflected path has length*

$$\ell_{\text{reflected}} = (x - x_0)^2 + (y + y_0)^2 = r^2 + r_0^2 - 2\,rr_0\cos(\theta + \theta_0)$$

Summation over classical paths (collapsed Feynman formalism), on the other hand, supplies

$$\begin{aligned} K(r,\theta,t;r_0,\theta_0,0) = \;& \tfrac{m}{iht}\exp\Big\{\tfrac{i}{\hbar}\tfrac{m}{2t}\big[(x-x_0)^2+(y-y_0)^2\big]\Big\} \qquad (107.1)\\ & -\tfrac{m}{iht}\exp\Big\{\tfrac{i}{\hbar}\tfrac{m}{2t}\big[(x-x_0)^2+(y+y_0)^2\big]\Big\}\\ = \;& \tfrac{m}{iht}\exp\Big\{\tfrac{i}{\hbar}\tfrac{m}{2t}\big[r^2+r_0^2-2rr_0\cos(\theta-\theta_0)\big]\Big\} \qquad (107.2)\\ & -\tfrac{m}{iht}\exp\Big\{\tfrac{i}{\hbar}\tfrac{m}{2t}\big[r^2+r_0^2-2rr_0\cos(\theta+\theta_0)\big]\Big\} \end{aligned}$$

where the minus sign—introduced to achieve compliance with the boundary condition

$$K(\text{edge of the wedge}, t; \bullet, \bullet, \bullet) = 0 \qquad (108)$$

—can be understood to arise from the conjectured circumstance[73] that

$$\left.\begin{array}{c}\text{classical action displays a } \textit{jump discontinuity}\\ \Delta S = \tfrac{1}{2}h \text{ at reflection points}\end{array}\right\} \tag{108}$$

The $K$ described at (107) satisfies the free Schrödinger equation and the wedge boundary condition; moreover

$$\begin{aligned}\lim_{t\downarrow 0} K(x,y,t;x_0,y_0,0) &= \delta(x-x_0)\Big\{\delta(y-y_0)-\delta(y+y_0)\Big\}\\ &\downarrow\\ &= \delta(x-x_0)\delta(y-y_0) \quad \text{on upper half-plane}\end{aligned}$$

Those same three properties are claimed by the $K$ described at (106). Evidently we have in hand an instance of "collapse"—two distinct descriptions of the same propagator.

How to establish the point analytically? Borrowing notation from page 76 and a trigonometric identity from page 77, we observe that (106) can be written

$$\begin{aligned}K &= \tfrac{2m}{\pi\hbar^2}\sum_n\Big\{\tfrac{1}{2}G_n\cos n(\theta-\theta_0)-\tfrac{1}{2}G_n\cos n(\theta+\theta_0)\Big\}\\ &= \tfrac{m}{\pi\hbar^2}\tfrac{1}{\lambda}\exp\Big\{-\tfrac{m}{2\lambda\hbar^2}(r^2+r_0^2)\Big\}\\ &\qquad\cdot\sum_n I_n\big(\tfrac{m}{2\lambda\hbar^2}2rr_0\big)\Big\{\cos n(\theta-\theta_0)-\cos n(\theta+\theta_0)\Big\}\end{aligned} \tag{109}$$

with $\lambda = \frac{i}{\hbar}t$. It is management of the surviving $\sum_n$ that inspires the following

MATHEMATICAL DIGRESSION

We are informed by "the other Russians"[74] that

$$\sum_{n=0}^{\infty} I_{n\nu}(z)\cos n\varphi = \tfrac{1}{2\nu}\sum_{k=k_-}^{k_+} A_k \exp\Big\{z\cos\tfrac{2k\pi+\varphi}{\nu}\Big\}+\tfrac{1}{2}I_0(z) \tag{110}$$

where

$$\begin{aligned}k_\pm &\equiv \pm\big[\tfrac{\nu\pi\mp\varphi}{2\pi}\big] \quad &&: \quad \text{here the square bracket means "integral part of"}\\ A_k &= 1 \quad &&\text{for } k\neq k_\pm\\ A_{k_\pm} &= \begin{cases}\tfrac{1}{2} & \text{if } \tfrac{\nu\pi\mp\varphi}{2\pi} = \ldots,-2,-1,0,1,2,\ldots\\ 1 & \text{otherwise}\end{cases}\end{aligned}$$

---

[73] Write $-1 = e^{i\pi} = e^{\frac{i}{\hbar}\Delta S}$. For discussion, see page 8 in an essay cited previously.[58] Recall that when the methods of ray optics are used to describe reflection-induced interference effects one encounters a similar phenomenon.

[74] A. P. Prudnikov, Yu. A. Brychkov & O. I. Marichev, *Integrals & Series: Volume II* (1986), entry **5.8.5.4**, which appears at the bottom of page 695.

But Prudnikov *et al* cite no source. So on the afternoon of 16 May 2001 I laid the problem before Ray Mayer, a mathematical colleague whose command of classical analysis is locally unrivaled, and who by the next morning had produced the following argument: Let $\nu = 1, 2, 3, \ldots$ and let $z$ be complex. Define

$$F_{\nu,z}(\varphi) \equiv \sum_{k=0}^{\nu-1} e^{z\cos\frac{\varphi+2\pi k}{\nu}}$$

and notice that $F_{\nu,z}(\varphi)$ is an <u>even function with period $2\pi$</u>,[75] which admits therefore of Fourier development

$$\begin{aligned}
F_{\nu,z}(\varphi) = \tfrac{1}{2}a_0 + \sum_{m=1}^{\infty} a_m\cos(m\varphi)\\
a_m &= \tfrac{1}{\pi}\int_0^{2\pi} F_{\nu,z}(\varphi)\cos(m\varphi)\,d\varphi\\
&= \tfrac{1}{\pi}\sum_{k=0}^{\nu-1}\int_0^{2\pi} e^{z\cos\frac{\varphi+2\pi k}{\nu}}\cos(m\varphi)\,d\varphi\\
&= \tfrac{1}{\pi}\sum_{k=0}^{\nu-1}\int_{\frac{2\pi k}{\nu}}^{\frac{2\pi(k+1)}{\nu}} e^{z\cos\theta}\cos(m[\nu\theta-2\pi k])\,\nu d\theta\\
&= \tfrac{1}{\pi}\nu\int_0^{2\pi} e^{z\cos\theta}\cos(m\nu\theta)\,d\theta\\
&= 2\nu\cdot\tfrac{1}{\pi}\int_0^{\pi} e^{z\cos\theta}\cos(m\nu\theta)\,d\theta
\end{aligned}$$

---

[75] To see how this comes about, look for example to the case $\nu = 3$: appealing to familiar properties of the cosine, we find that the operation $\varphi \rightarrow -\varphi$ sends the set

$$\Big\{\cos\left(\tfrac{1}{3}\varphi\right), \cos\left(\tfrac{1}{3}\varphi + \tfrac{1}{3}2\pi\right), \cos\left(\tfrac{1}{3}\varphi + \tfrac{2}{3}2\pi\right)\Big\}$$

into

$$\begin{aligned}
&\Big\{\cos\left(\tfrac{1}{3}\varphi\right), \cos\left(\tfrac{1}{3}\varphi - \tfrac{1}{3}2\pi\right), \cos\left(\tfrac{1}{3}\varphi - \tfrac{2}{3}2\pi\right)\Big\}\\
&\qquad = \Big\{\cos\left(\tfrac{1}{3}\varphi\right), \cos\left(\tfrac{1}{3}\varphi + \tfrac{2}{3}2\pi\right), \cos\left(\tfrac{1}{3}\varphi + \tfrac{1}{3}2\pi\right)\Big\}\\
&\qquad = \text{permutation of the original set}
\end{aligned}$$

while $\varphi \rightarrow \varphi + 2\pi$ produces

$$\begin{aligned}
&\Big\{\cos\left(\tfrac{1}{3}\varphi + \tfrac{1}{3}2\pi\right), \cos\left(\tfrac{1}{3}\varphi + \tfrac{1}{3}2\pi + \tfrac{1}{3}2\pi\right), \cos\left(\tfrac{1}{3}\varphi + \tfrac{1}{3}2\pi + \tfrac{2}{3}2\pi\right)\Big\}\\
&\qquad = \Big\{\cos\left(\tfrac{1}{3}\varphi + \tfrac{1}{3}2\pi\right), \cos\left(\tfrac{1}{3}\varphi + \tfrac{2}{3}2\pi\right), \cos\left(\tfrac{1}{3}\varphi\right)\Big\}\\
&\qquad = \text{permutation of the original set}
\end{aligned}$$

But Abramowitz & Stegun report (at **9.6.20**) and Watson proves (in §**6.22**) that

$$\tfrac{1}{\pi}\int_0^\pi e^{z\cos\theta}\cos(m\nu\theta)\,d\theta = I_{m\nu}(z) + \tfrac{1}{\pi}\sin(m\nu\pi)\int_0^\infty e^{-z\cosh t - m\nu t}\,dt$$

provided $|\arg(z)| < \frac{1}{2}\pi$. In the intended physical application $z = -\frac{i}{\hbar}\frac{m}{2t}2rr_0$ places us—characteristically—right at the edge of the allowed region. But Mayer remarks that if $m\nu$ is an integer—which in the present context is certainly the case—then the second term on the right drops away; we are left with entire functions on left and right, so the condition $|\arg(z)| < \frac{1}{2}\pi$ can be disregarded. The implication is that

$$a_m = 2\nu I_{m\nu}(z) \quad \text{for all complex numbers } z$$

which gives

$$\begin{aligned} F_{\nu,z}(\varphi) \equiv \sum_{k=0}^{\nu-1} e^{z\cos\frac{\varphi+2\pi k}{\nu}} &= \nu I_0(z) + 2\nu\sum_{m=1}^\infty I_{m\nu}(z)\cos(m\varphi) \\ &= 2\nu\sum_{n=0}^\infty I_{n\nu}(z)\cos(n\varphi) - \nu I_0(z) \end{aligned}$$

or

$$\sum_{n=0}^\infty I_{n\nu}(z)\cos(n\varphi) = \tfrac{1}{2\nu}\sum_{k=0}^{\nu-1} e^{z\cos\frac{\varphi+2\pi k}{\nu}} + \tfrac{1}{2}I_0(z) \tag{111}$$

The sum on the right can, by the way, be written

$$\sum_{k=p}^{\nu-1+p} \quad \text{for all } \pm \text{ integers } p$$

Mayer's (111) is a special instance of (110), and his argument owes much of its elegance to his decision to ignore the possibility that $\nu$ might *not* be an integer. Here ends the mathematical digression.

Returning with (111)—which in the case $\nu = 1$ reads

$$\sum_{n=0}^\infty I_n(z)\cos(n\varphi) = \tfrac{1}{2}e^{z\cos\varphi} + \tfrac{1}{2}I_0(z) \tag{111.1}$$

—to (109), we observe that the $\frac{1}{2}I_0(z)$ terms (which enter with opposite signs) cancel, and that we are left with

$$\begin{aligned} K = \tfrac{m}{iht}\exp\Big\{\tfrac{i}{\hbar}\tfrac{m}{2t}(r^2 + r_0^2)\Big\} \\ \cdot\Big\{\exp\big[-\tfrac{i}{\hbar}\tfrac{m}{2t}2rr_0\cos(\theta-\theta_0)\big] - \exp\big[-\tfrac{i}{\hbar}\tfrac{m}{2t}2rr_0\cos(\theta+\theta_0)\big]\Big\} \end{aligned}$$

which precisely reproduces the collapsed Feynman sum (107.2).

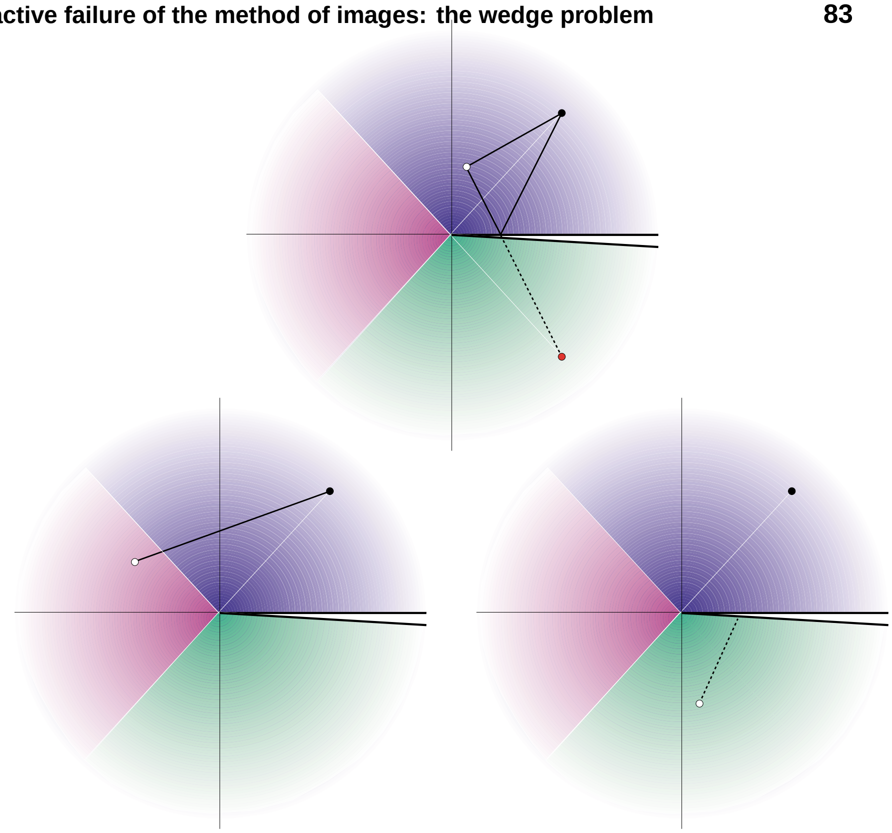


FIGURE 13: *Construction of the 10 reflective paths* $\circ \rightarrow \bullet$ *in a wedge with* $\nu \equiv \pi/\alpha = 5$. *Red paths are of even order (involve an even number of reflections); blue paths are of odd order.*

Suppose, more generally, that the wedge angle $\alpha$ divides $\pi$ an integral number of times; *i.e.*, that

$$\mu = n\nu \quad \text{where} \quad \nu \equiv \frac{\pi}{\alpha} = \underset{\uparrow\text{case just considered}}{1}, 2, 3, \ldots$$

We then have

$$K = \nu \tfrac{m}{\pi\hbar^2}\tfrac{1}{\lambda}\exp\Big\{-\tfrac{m}{2\lambda\hbar^2}(r^2+r_0^2)\Big\} \cdot \sum_n I_{n\nu}\big(\tfrac{m}{2\lambda\hbar^2}2rr_0\big)\Big\{\cos n\nu(\theta-\theta_0) - \cos n\nu(\theta+\theta_0)\Big\}$$

which gives back (109) in the case $\nu = 1$. Drawing upon (111) we get

$$= \tfrac{m}{iht}e^{\frac{i}{\hbar}\frac{m}{2t}(r^2+r_0^2)} \cdot \sum_{k=0}^{\nu-1}\Big\{e^{-\frac{i}{\hbar}\frac{m}{2t}rr_0\cos(\theta-\theta_0+2\alpha k)} - e^{-\frac{i}{\hbar}\frac{m}{2t}rr_0\cos(\theta+\theta_0+2\alpha k)}\Big\}$$

$$= \Big\{\sum_{\text{images of even order}} - \sum_{\text{images of odd order}}\Big\}\tfrac{m}{iht}e^{\frac{i}{\hbar}\frac{m}{2t}(\text{path length})^2}$$

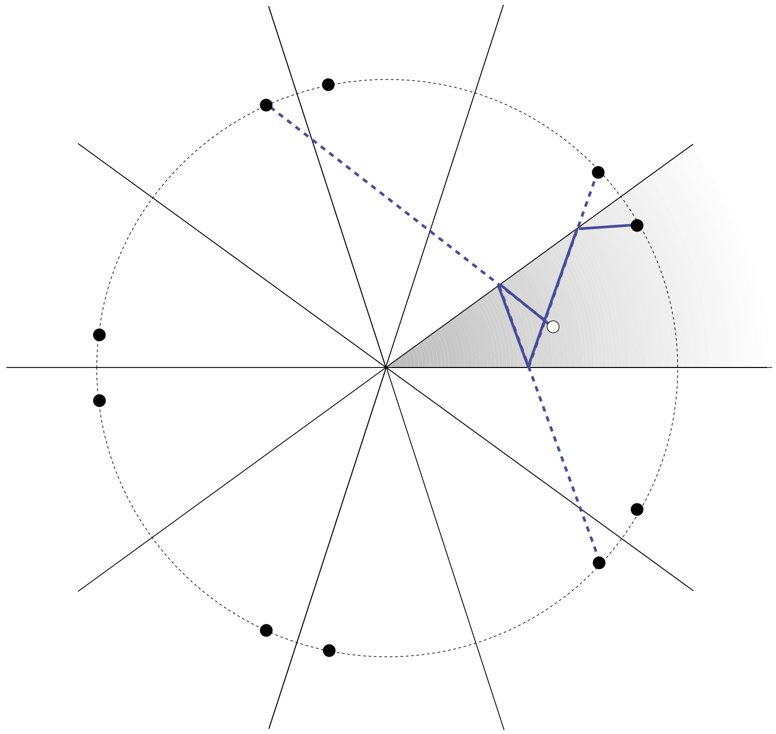


FIGURE 14: *Illustration of how the preceding diagram is used to deduce the design of a reflected path—here a path of odd order 3.*

In all such cases the method of images (collapsed Feynman formalism) works to perfection, and has a secure analytical base. But the method of images encounters grave difficulties when

$$\nu \equiv \frac{\pi}{\alpha} \quad \text{is } \textit{not} \text{ an integer}$$

Assume the wedge to be in "standard position:" $\theta = 0$ on right edge; $\theta = \alpha$ on left edge. Reflection in either edge is $r$-preserving. It is geometrically evident that

$$\begin{aligned} \mathsf{R} &\equiv \text{reflection in right edge sends } \theta \rightarrow \phantom{2\alpha} - \theta \\ \mathsf{L} &\equiv \text{reflection in left edge sends } \theta \rightarrow 2\alpha - \theta \end{aligned}$$

and $\mathsf{R}^2 = \mathsf{L}^2 = \mathsf{I}$, so reflective images of $k^{\text{th}}$ order are produced by the first else second of these of these operations

$$\underbrace{\cdots \mathsf{RLRLR}}_{k \text{ factors}} \quad : \quad \text{terminates} \begin{cases} \mathsf{L} \text{ if } k \text{ is even} \\ \mathsf{R} \text{ if } k \text{ is odd} \end{cases}$$

$$\underbrace{\cdots \mathsf{LRLRL}}_{k \text{ factors}} \quad : \quad \text{terminates} \begin{cases} \mathsf{R} \text{ if } k \text{ is even} \\ \mathsf{L} \text{ if } k \text{ is odd} \end{cases}$$

which produce

$$
\left.\begin{array}{r}
\theta\\
-\,\theta\\
2\alpha+\theta\\
-2\alpha-\theta\\
4\alpha+\theta\\
-4\alpha-\theta\\
\vdots\\
\pm\big[k\tfrac{2\pi}{\nu}+\theta\,\big]
\end{array}\right.
\quad\text{else}\quad
\begin{array}{r}
\theta\\
2\alpha-\theta\\
-2\alpha+\theta\\
4\alpha-\theta\\
-4\alpha+\theta\\
6\alpha-\theta\\
\vdots\\
\pm\big[k\tfrac{2\pi}{\nu}-\theta\,\big]
\end{array}
\quad : \quad
\begin{array}{l}
0^{\text{th}}\text{ order}\\
1^{\text{st}}\text{ order}\\
2^{\text{nd}}\text{ order}\\
3^{\text{rd}}\text{ order}\\
4^{\text{th}}\text{ order}\\
5^{\text{th}}\text{ order}\\
\vdots\\
k=0,1,2,\ldots
\end{array}
$$

with the understanding that $+\big[0\frac{2\pi}{\nu}-\theta\,\big]$ is to be omitted from the list on the right. That same merged pointset can also—and more tranparently—be produced by merging the flip-flop-flip-flop results (see the following figure) of $\circlearrowleft/\circlearrowright$ "reflective tesselation"

$$
\begin{array}{rcr}
\circlearrowleft\text{-tesselation} & & \circlearrowright\text{-tesselation}\\
\theta & & \theta\\
-\,\theta & \text{flip} & -2\alpha-\theta\\
2\alpha+\theta & \text{flop} & -2\alpha+\theta\\
2\alpha-\theta & \text{flip} & -4\alpha-\theta\\
4\alpha+\theta & \text{flop} & -4\alpha+\theta\\
4\alpha-\theta & \text{flip} & -6\alpha-\theta\\
\vdots & & \vdots\\
+\big[k\tfrac{2\pi}{\nu}\mp\theta\,\big] & \text{flip/flop} & -\big[k\tfrac{2\pi}{\nu}\mp\theta\,\big]
\end{array}
$$

Flip images are odd, flop images are even. It is clear (especially from the figure) that

- if $\nu$ = integer then $(\text{flip-flop})^\nu$ restores the wedge face-up to its original position: $\theta$ has become $\theta+2\pi\equiv\theta \bmod 2\pi$ but the wedge lies now "on the next higher sheet." Progress in the reverse sense $\circlearrowright$ yields the same point set, but deposits the wedge "on the next lower sheet;"

- if $\nu=\frac{1}{2}\cdot$integer then $(\text{flip-flop})^{\frac{1}{2}\nu}$ restores the wedge face-*down* to its original position: one must complete a second tour to restore the wedge to its original state, but it lies then "on the second higher sheet." Progress in the reverse sense $\circlearrowright$ places the wedge "on the second lower sheet;"

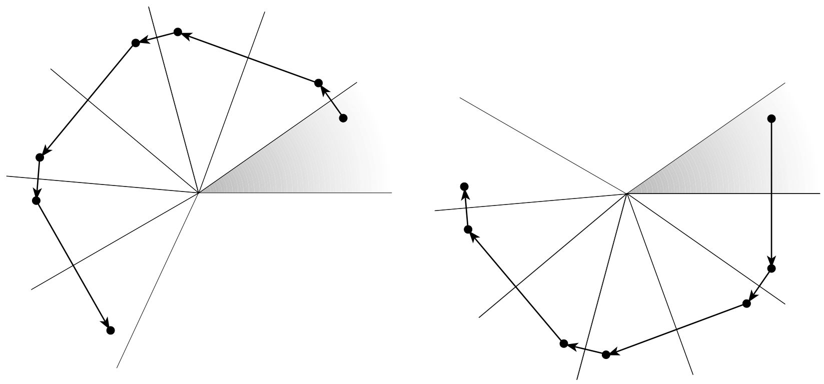


FIGURE 15: *Reflective tesselation of type ↺ on the left, of type ↻ on the left. The two operations are, in an obvious sense, inverses of one another. Here $\alpha = 35°$, so $\nu = \frac{36}{7}$ is non-integral.*

- if $\nu = \frac{a}{b}$ is rational then (flip-flop)$^a$ restores the wedge face-up to its original position, but places it on "on the upper $b^{\text{th}}$ sheet." Progress in the reverse sense ↻ places the wedge "on the lower $b^{\text{th}}$ sheet;"

- if $\nu$ is irrational then (flip-flop)$^{\text{power}}$ *never* restores the wedge its original position: the initial point has a *continuum* of reflective images.

But if we place ourselves at ∘ inside a mirrored $\alpha$-wedge with the intention of shooting elastic pellets at the reflective images of a target • we find that the number of visible target images depends jointly upon where the target has been placed and where we stand. And that, as we fire in all directions, the

$$\text{greatest possible number of reflections} = \text{least integer} \geqslant \nu \equiv \frac{\pi}{\alpha}$$

but the number of reflections depends not only upon the direction in which fire but where we stand, and (unless $\nu$ is an integer) the greatest achievable number may be less than that.

These points are illustrated in Figure 16. Wedges in the left column were generated by ↺ tesselation: • images lie on the principal sheet, • images lie on the sheet above. Wedges in the right column were generated by ↻ tesselation: black and red images have enchanged places: • images lie on the principal sheet, • images lie on the sheet below. The upper row illustrates a situation in which ∘ sees only one image in the left mirror, but two in the right (only the 2$^{\text{nd}}$-order trajectory is shown). In the middle row, ∘ has been moved to a position where a second image has become visible in the left mirror (two remain visible in the right mirror). In the lower row a third image has become visible in the left mirror (and again, two remain visible in the right mirror). Three is maximal in this instance, since (least integer greater than $\frac{5}{2}$)=3. Reflected

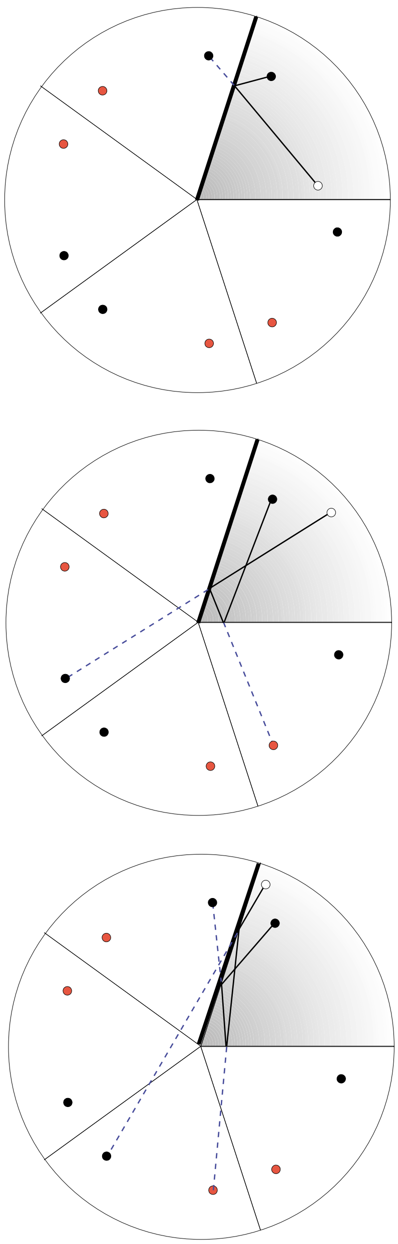


FIGURE 16: *The number of target images visible from a source point depends on the relative placement of target and source. The top, middle and bottom rows illustrate three typical situations, in a wedge with* $\nu = \frac{5}{2}$*. Details are discussed in the text.*

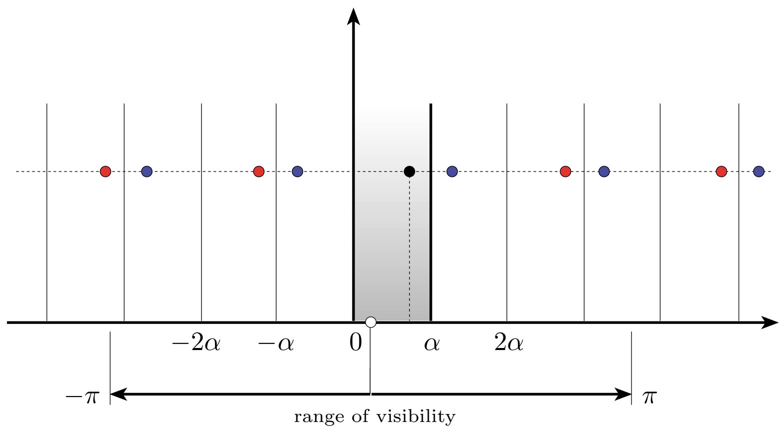



FIGURE 17: *The time axis runs up, the theta axis to the right. The source point* ∘ *has angular coordinate* $\theta_0$*, the target point* • *has angular coordinate* $\theta$*. Even images* • *of the target appear at points* $\theta+k2\alpha = k\frac{2\pi}{\nu}+\theta$*, odd images* • *at points* $-\theta+k2\alpha = k\frac{2\pi}{\nu}-\theta$*, where* $k = \ldots,-2,-1,0,1,2,\ldots$ *In this angular analog of the familiar "barber shop construction" the time axis has no metric significance, but serves only to distinguish "before" from "after."*

rays/particles visit alternately first one mirror/wall then the other, which is why in each figure the successive images that enter into the construction of a trajectory proceed • • • • ··· • •.

Notice also that—by the simplest of geometrical arguments—if • : $(r,\theta)$ is visible/invisible from ∘ : $(r_0,\theta_0)$ then

✓ so is • visible/invisible from $(r_0+a_0,\theta_0)$

✓ so is $(r+a,\theta)$ visible/invisible from ∘

In short: the number of •-images visible from ∘ depends only upon $\theta$ and $\theta_0$. This conclusion will be sharpened in a moment.

The discovered irrelevance of $r$ and $r_0$ means that we can use an angular analog of the familiar "barber shop construction" (Figure 17) to resolve all of the "who sees what" questions presented by the wedge problem. Examination of such figures discloses that

> The number of even images • (including the $0^{\text{th}}$-order image •) visible from ∘ depends on $\theta$ and $\theta_0$ only through their difference $\theta-\theta_0$. The order index $k$ ranges from $k_-^{\text{even}}$ to $k_+^{\text{even}}$ where

$$\begin{aligned} k_+^{\text{even}} &\equiv \text{greatest integer such that } k_+^{\text{even}}2\alpha+(\theta-\theta_0)\leqslant\pi \\ &= \left[\tfrac{\pi-(\theta-\theta_0)}{2\alpha}\right] \\ &= +\left[\tfrac{\nu\pi-\varphi}{2\pi}\right] \quad : \quad \nu\equiv\frac{\pi}{\alpha} \text{ and } \varphi\equiv\nu(\theta-\theta_0) \\ k_-^{\text{even}} &= -\left[\tfrac{\nu\pi+\varphi}{2\pi}\right] \end{aligned}$$

The number of odd images • visible from ∘ depends on $\theta$ and $\theta_0$ only through their sum $\theta+\theta_0$. The order index $k$ ranges from $k_-^{\text{odd}}$ to $k_+^{\text{odd}}$ where

$$
\begin{aligned}
k_+^{\text{odd}} &\equiv \text{greatest integer such that } k_+^{\text{even}} 2\alpha + (\theta+\theta_0) \leqslant \pi \\
&= \left[\tfrac{\pi-(\theta+\theta_0)}{2\alpha}\right] \\
&= +\left[\tfrac{\nu\pi-\varphi}{2\pi}\right] \quad : \quad \nu \equiv \frac{\pi}{\alpha} \text{ and } \varphi \equiv \nu(\theta+\theta_0) \\
k_-^{\text{odd}} &= -\left[\tfrac{\nu\pi+\varphi}{2\pi}\right]
\end{aligned}
$$