Quantization: Path Integrals and Zeta Functions

David Seery
Pedagogical Seminar, 25 February 2003

This lecture is an attempt to present some examples of what quantization can mean. Undergraduate courses in quantum mechanics typically emphasize the problem-solving and computational aspects of the subject (like the Schrödinger representation, perturbation theory, and so on) that are needed for the practical business of calculating physical quantities. Here, we take a more abstract view:

What does it *mean* for a physical system to be quantized?

Why do we bother? One sometimes encounters the point of view that quantum mechanics consists in some mystical re-interpretation of the wave equations of classical physics which has *already been done*, once and for all. In that case, all that remains is to explore the consequences of the theory: it isn't meaningful to ask how the equations of quantum mechanics were obtained; all that matters is that they work. Unfortunately this kind of mysticism sometimes leaks out of the scientific loop into philosophy and the popular culture.

At least one difficulty with this attitude is that the quantization programme is not yet complete: as we all know, there is no satisfactory quantization of gravity. (We are ignoring string theory, which does appear to contain a perturbatively finite, consistent quantum gravity which reduces to "something like" Einstein gravity[1] in the low energy limit, and moreover correctly reproduces and in supersymmetric cases can *explain* semiclassical results like the Bekenstein–Hawking entropy. Nevertheless, whether or not this gravity is *our* gravity, the gravity we see in the world around us, remains to be seen. It is also true that this quantum gravity did not arise from quantizing an existing theory of a spin-2 field.) In simple quantum mechanics without phase space constraints or gauge artefacts, the various methods of mystical reinterpretation usually contrive to give the same result. For gravity, this is no longer the case. There is no problem with mystically re-interpreting the field equations of general relativity in much the same way as any other theory: in the Schrödinger representation, the result is the Wheeler–de Witt equation. If this is done, for simplicity's sake, in three dimensions (where the dynamics is trivial) then one finds that the methods of mystical reinterpretation do not all agree. To decide which quantization is correct (if any) one has to study the quantization methods themselves. One expects that this is the reverse of the real situation in Nature, where the theory is presumably *a priori* quantum mechanical and we are trying to understand the classical limit $\hbar \downarrow 0$.

Studying quantization methods is also valuable for other reasons. It is unfortunate that many historical terms such as 'wave–particle duality', 'second quantization' or the 'Dirac sea', descending from the days of *ad hoc* quantum mechanics, still linger on today with the result that, for example, anyone who asks why non-relativistic quantum mechanics is governed by Schrödinger's equation is liable to come away with the impression that it has something to do with 'matter waves'. These days we do not insist on waves or particles, only fields: to obtain a second-quantized theory one quantizes exactly once: and (as Julian Schwinger said) the Dirac sea is "best regarded as a historical curiosity, and forgotten." Studying *exactly* how the classical and quantum theories are related allows one to rigorously dispense with oddities like 'matter waves' which do not usefully contribute to one's understanding.

In part the momentum for studying different methods of quantization comes from the drive in physics to study increasingly complicated systems where the methods applied to non-relativistic quantum mechanics in the 1920s simply do not work. In the 1930s and 40s the focus was on quantum field theory itself. Non-Abelian gauge theories were introduced by Yang and Mills in 1954 and their quantizations were studied

[1] More precisely, the string sigma-model action gives supergravity with a dilaton, NS–NS antisymmetric tensor, and extra fields from the Ramond–Ramond sector. The bosonic part of this, ignoring the NS–NS antisymmetric tensor and the Ramond–Ramond fields, is essentially Brans–Dicke gravity, and there are strong solar system constraints on the dilatonic coupling in this theory.

by Feynman, by Fadeev and Popov, and by de Witt. Supersymmetry appeared in the early 1970s as an offspring of fermionic string theory, and the first supergravity was written down in the early 1980s. Many of the quantizations successfully applied to these theories were originally invented with an eye on quantum gravity and as such provide sufficient justification for studying theories like loop quantum gravity or strings which appear divorced from experiment, but which contribute essentially to our understanding of theories like QCD of great contemporary experimental relevance. The B-physics seen at B-factories like BaBar and BELLE and the quark–gluon plasma we all hope to see at the LHC represents some of the most important new physics we shall find any time soon. But QCD is not easy to quantize and we still do not have any kind of control over its low-energy perturbation theory. Nevertheless, it is a long-standing conjecture that, for large N, $SU(N)$ Yang–Mills theory is dual to a string theory and it is possible that our best control over QCD may come from an understanding of its holographic dual. (QCD is the case $N = 3$.)

This lecture is split roughly in two, corresponding to Hamiltonian and Lagrangian quantizations. We first spend some time developing quantization using phase space techniques to obtain some strong results (at least in the case of finite-dimensional phase spaces), and then move on to study the Lagrangian path integral, which is probably the most *practical* quantization tool.

1. Hamiltonian and phase space techniques

Hamiltonian quantizations were the first to be invented and today remain the most rigorous, although in practice they are not very convenient to use. The main goal in this section is the Stone–von Neumann theorem which will guarantee the uniqueness of a quantization based on a finite-dimensional phase space. First we briefly revise how to write a classical theory in phase space form.

Consider some classical system, like the Newtonian point particle. It is described by a second-order differential equation supplemented with some boundary conditions:

$$m\frac{\mathrm{d}^2 q_a}{\mathrm{d}t^2} = -\partial_a V(q).$$

In this case the boundary conditions consist in specifying any two of q_a and $\dot{q}_a$ at $t = t_{\text{initial}}$ or $t = t_{\text{final}}$. The quantities appearing in this equation are the *coordinates* q^a, where (by convention) $a = 0, \ldots, D-1$ for a D-dimensional theory; the *mass* m of the particle; and a *potential* $V(q)$. The symbol ∂_a is a convenient abbreviation for $\partial/\partial q^a$. This gives us D second-order differential equations for D quantities, the q_a. On the other hand we can perfectly well rewrite Newton's law using an 'auxiliary' quantity, the momentum, defined by $p_a = m\dot{q}_a$:

$$\frac{\mathrm{d}p_a}{\mathrm{d}t} = -\partial_a V(q).$$

This gives us $2D$ first-order differential equations for $2D$ quantities, the coordinates x_a and the momenta p_a.

The q_a are coordinates on a D-dimensional space, the physical space in which the particle can move, called **configuration space**. The q_a and p_a together are coordinates on a $2D$-dimensional space called **phase space**.

The point particle has a conserved quantity associated with its motion which as always can be found by taking an inner product with the velocity $\dot{q}_a$:

$$m\dot{q}_a\ddot{q}_a = -\dot{q}_a\partial_a V(q) \implies \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m}{2}\dot{q}^2\right) = -\frac{\mathrm{d}}{\mathrm{d}t}V.$$

This can be integrated at once to find $m\dot{q}^2/2 + V = H$ for some constant H. We are using the Einstein summation convention: any index appearing twice on the same side of an equation is to be summed over. This convention is very convenient and is used universally in physics and mathematics. Evidently H is what one conventionally calls the energy, and it can be rewritten in terms of p and q alone as $H = p^2/2m + V(q)$. (It is always possible to write H in terms of p and q without time derivatives of either, although we will not prove it here.)

Whatever path the particle takes on phase space, it must conserve H. Therefore one has

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial q_a}\frac{\mathrm{d}q_a}{\mathrm{d}t} + \frac{\partial H}{\partial p_a}\frac{\mathrm{d}p_a}{\mathrm{d}t} = 0. \tag{1}$$

We can satisfy this by taking the q_a and p_a to satisfy the evolution equations

$$\frac{\mathrm{d}q_a}{\mathrm{d}t} = \frac{\partial H}{\partial p_a} \quad \text{and} \quad \frac{\mathrm{d}p_a}{\mathrm{d}t} = -\frac{\partial H}{\partial q_a}.$$

These are called **Hamilton's equations**, and the quantity H is called the **Hamiltonian**.

There is an important way to rewrite these results. The time derivative of *any* function o along a path in phase space satisfies Eq. (1):

$$\begin{aligned}\frac{\mathrm{d}o}{\mathrm{d}t} &= \frac{\partial o}{\partial q_a}\frac{\mathrm{d}q_a}{\mathrm{d}t} + \frac{\partial o}{\partial p_a}\frac{\mathrm{d}p_a}{\mathrm{d}t} \\ &= \frac{\partial o}{\partial q_a}\frac{\partial H}{\partial p_a} - \frac{\partial o}{\partial p_a}\frac{\partial H}{\partial q_a} \\ &= \{o, H\},\end{aligned}$$

where we have introduced the **Poisson bracket**, defined for any two functions f and g on phase space by

$$\{f, g\} = \frac{\partial f}{\partial q_a}\frac{\partial g}{\partial p_a} - \frac{\partial f}{\partial p_a}\frac{\partial g}{\partial q_a}. \tag{2}$$

The Poisson bracket is one of the most useful structures in classical mechanics and has an important interpretation in the context of Lie algebras, but we will not have space to develop this in detail here. What is important is the time-derivative relation $\dot{o} = \{o, H\}$ derived above, and the Poisson brackets of the elementary functions p_i and q_i on phase space,

$$\{q_i, q_j\} = \{p_i, p_j\} = 0 \quad \{q_i, p_j\} = \delta_{ij}.$$

The observables in the classical theory are *any* functions of the coordinates and momenta, so they are simply the functions on phase space. (This is a large group of observables, so one might try to cut it down by considering only nice functions on phase space, eg., those well-behaved enough to have a Fourier transform, at least in a distributional sense. This is the usual choice; such a set of functions is said to constitute *Schwarz space*.) The Poisson bracket defines a kind of multiplication in which the *classical observables* may not commute even before quantization.

The problem now is to construct a quantum theory from this classical framework—a phase space M coordinatized by the (q_i, p_i) whose points correspond to states of the system, the set of functions on M satisfying suitable niceness requirements and constituting the classical observables, and the noncommutative Poisson bracket. There have been several attempts to attack this problem over the years, but the issue is not yet closed.

The first generalization is to enlarge our concept of a state. We endow the states with the ability to add with each other to produce a new state which in some sense is a *superposition* of the original two. Of course this makes no sense from the classical point of view in which states are points in a manifold M, so at this stage we are required to abandon the classical phase space and introduce some other structure we will call $\mathcal{H}$ to hold the states. We will choose $\mathcal{H}$ to be a complex vector space whose elements are denoted by Dirac's kets $|\psi\rangle$ and worry about whether any of this makes physical sense later.

The next step is to take the classical observables over to $\mathcal{H}$. Since $\mathcal{H}$ is a complex vector space there are not very many choices that make sense; we choose the observables to be (bounded) linear operators whose eigenvalues constitute the allowed measurements. In a sense, this is the quantization step: eigenvalues of bounded linear operators are typically discrete. To this end we invent a map $\hat{}$ which takes a classical observable and assigns to it the corresponding quantum operator. Unfortunately, no-one knows how to find a $\hat{}$ which does this job. Dirac's proposal was to consider $\hat{}$ as a Lie algebra homomorphism taking the algebra of classical observables under the Poisson bracket into the algebra of bounded linear operators under the commutator. Dirac's proposal is then that

$$\widehat{\{f, g\}} = -\frac{i}{\hbar}[\hat{f}, \hat{g}] \tag{3a}$$

$$\widehat{fg} = \frac{1}{2}(\hat{f}\hat{g} + \hat{g}\hat{g}) \tag{3b}$$

(The sign in the first equation depends on the definition of the Poisson bracket. It is chosen to get $[\hat{q},\hat{p}] = i\hbar$. The second equation, describing the action of $\hat{}$ on fg is chosen so that $\widehat{fg}$ is Hermitian if $\hat{f}$ and $\hat{g}$ are.) Note that there is obviously something underhand going on here, because we have had to introduce a new constant of nature $\hbar$ while nobody was looking. In a sense, this is the first appearance of $\hbar$ in physics, although a suspicious constant with the dimensions of $\hbar$ is implicitly present in the classical world (see, eg., Azcárraga & Izquierdo [**3**]). (Experimentalists prefer to say that $\hbar$ appears with the observed quantization of light, $E = \hbar\omega$. But this is an obviously detectable experimental heresy.)

Dirac's quantization does not tell us how to explicitly find $\hat{f}$ and $\hat{g}$ but only describes a constraint they must obey. It is usually possible by trial and error to find appropriate operators, but one can easily show that (3a) is not consistent for arbitrary quartic polynomials in the positions and momenta: one cannot deal with observables of the form $qqpp$ (or higher combinations of p and q). The usual solution is to abandon the prescription for $\widehat{fg}$ and find an operator ordering by hand. As a result, the quantization procedure is not algorithmic, much less automatic. Since all good physicists are lazy, this is far too much work.

The Hamiltonian is usually expressible as a function quadratic in the p together with some arbitrary function of the q, so Dirac's quantization *is* sensible *in this case*. From Dirac's Poisson bracket $\mapsto$ commutator rule, the time evolution of a quantum observable $\hat{o}$ is given by

$$\frac{\mathrm{d}\hat{o}}{\mathrm{d}t} = \frac{1}{i\hbar}[\hat{o},\hat{H}].$$

This is called *Heisenberg's equation of motion*, although as we have seen it is just the quantum version of Hamilton's equations and (as far as I know) it was first written down by Dirac. This equation makes sense: we constructed the quantum state space $\mathcal{H}$ from the classical phase space, which did not depend on time. Thus the states $|\psi\rangle$ in $\mathcal{H}$ do not carry time dependence; this is said to be quantum mechanics in the *Heisenberg picture*. However, although the Heisenberg picture arises most naturally in the quantization process, there are alternatives and one of these will be useful in the derivation of the path integral to come.

To this end, let us call the operators we have been discussing up to this point *Heisenberg picture operators*, and denote this by affixing a subscript H to each operator. We can define a related set of operators, called *Schrödinger picture operators* and denoted by a subscript S, by writing

$$\hat{o}_S = \mathrm{e}^{-i\hat{H}t/\hbar}\hat{o}_H \mathrm{e}^{i\hat{H}t/\hbar}.$$

The time evolution of the Schrödinger picture operators then satisfies

$$\begin{aligned}\frac{\mathrm{d}\hat{o}_S}{\mathrm{d}t} &= -\frac{i}{\hbar}\hat{H}\hat{o}_S + \frac{i}{\hbar}\hat{o}_S\hat{H} + \mathrm{e}^{-i\hat{H}t/\hbar}\frac{\mathrm{d}\hat{o}_H}{\mathrm{d}t}\mathrm{e}^{i\hat{H}t/\hbar}\\ &= \frac{i}{\hbar}[\hat{o}_S,\hat{H}] + \mathrm{e}^{-i\hat{H}t/\hbar}\frac{\mathrm{d}\hat{o}_H}{\mathrm{d}t}\mathrm{e}^{i\hat{H}t/\hbar},\end{aligned}$$

so from substituting Heisenberg's equation of motion for $\mathrm{d}\hat{o}/\mathrm{d}t$ one finds

$$\begin{aligned}\frac{\mathrm{d}\hat{o}_S}{\mathrm{d}t} &= \frac{i}{\hbar}[\hat{o}_S,\hat{H}] - \frac{i}{\hbar}\mathrm{e}^{-i\hat{H}t/\hbar}[\hat{o}_H,\hat{H}]\mathrm{e}^{i\hat{H}t/\hbar}\\ &= \frac{i}{\hbar}[\hat{o}_S,\hat{H}] - \frac{i}{\hbar}[\hat{o}_S,\hat{H}].\end{aligned}$$

Thus the Schrödinger picture operators are time independent. To compensate for the extra time dependence in the $\hat{o}_S$ we must introduce some time dependence in the states $|\psi\rangle$. Denote Heisenberg picture states by a subscript H as before, and let $|\psi_n\rangle_H$ be a collection of eigenstates of the Heisenberg picture operator $\hat{o}_H$ with eigenvalues λ_n. Thus $\hat{o}_H|\psi_n\rangle_H = \lambda_n|\psi_n\rangle_H$, so introducing factors of $\mathrm{e}^{\pm i\hat{H}t/\hbar}$ one obtains

$$\mathrm{e}^{-i\hat{H}t/\hbar}\hat{o}_H\mathrm{e}^{i\hat{H}t/\hbar}\mathrm{e}^{-i\hat{H}t/\hbar}|\psi_n\rangle_H = \lambda_n\mathrm{e}^{-i\hat{H}t/\hbar}|\psi_n\rangle_H \quad \text{or} \quad \hat{o}_S|\psi_n(t)\rangle_S = \lambda_n|\psi_n(t)\rangle_S$$

where $|\psi_n(t)\rangle_S = \mathrm{e}^{-i\hat{H}t/\hbar}|\psi_n\rangle_H$. The Schrödinger states obviously obey the evolution equation

$$\frac{\partial}{\partial t}|\psi(t)\rangle = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle. \tag{4}$$

This is the Schrödinger equation in its conventional form.

We can now build the quantum mechanics that is in all the textbooks. To each state $|\psi\rangle$ in $\mathcal{H}$ we associate the *wavefunction* $\Psi(q) = \langle q|\psi\rangle$ and take Ψ to lie in the space L^2 of square-integrable functions.[2] The operators $\hat{q}_i$ are represented by left-multiplication with q_i and p_i by left-derivation with $-i\hbar\partial/\partial q_i$. Letting $\hat{}$ act on the Hamiltonian gives

$$\hat{H} = \widehat{\left(\frac{p^2}{2m} + V(q)\right)} = -\frac{\hbar^2}{2m}\triangle + V(q)$$

which when combined with the Schrödinger state evolution equation above gives back what we usually consider to be the Schrödinger equation, together with all its ramifications:

$$\left(-\frac{\hbar^2}{2m}\triangle + V(q)\right)\Psi = i\hbar\frac{\partial\Psi}{\partial t}.$$

But there remains something unsatisfactory about the whole procedure. For one thing, it is odd that one apparently cannot quantize arbitrary (well-behaved) observables, which there is no reason to suppose might not one day be required.

One can do better, although not along the lines of Dirac's prescription: Groenewold proved in 1949 that there is *no* quantization of any sensible (large) observable algebra such that $\hat{}$ is a Lie algebra homomorphism in Dirac's sense. Consider instead the Fourier representation of an observable f:

$$f(q,p) = \int \widetilde{f(\lambda,\eta)}\, \mathrm{e}^{i(\lambda q + \eta p)}\, \mathrm{d}\lambda\, \mathrm{d}\eta$$

where $\widetilde{f(\lambda,\eta)}$ is the inverse Fourier transform of $f(q,p)$. Let us promote the q and p to non-commuting coordinates $\hat{q}$ and $\hat{p}$ satisfying Dirac's condition $[\hat{q},\hat{p}] = i\hbar$. This associates to each $f(q,p)$ a symbol $\mathsf{W}(f)$,

$$\mathsf{W}(f) = \int \widetilde{f(\lambda,\eta)}\, \mathrm{e}^{i(\lambda\hat{q} + \eta\hat{p})}\, \mathrm{d}\lambda\, \mathrm{d}\eta.$$

This is *Weyl's quantization.* The composition of operators satisfies

$$\mathsf{W}(f)\,\mathsf{W}(g) = \mathsf{W}(f\star g)$$

where $\star$ is called the *star* or *Moyal* product and satisfies

$$f\star g = \exp\left(\frac{i\hbar}{2}P^{ij}\frac{\partial}{\partial x^i}\frac{\partial}{\partial y^j}\right) f(x)g(y)\bigg|_{y\to x} = 1 + \frac{i\hbar}{2}\{f,g\} + O(\hbar^2). \tag{5}$$

Here we have gathered the q and p together into a single coordinate x (respectively, y) and written $\{x^i,x^j\} = P^{ij}$, so this is the exponentiation of the Poisson bracket. (The y are just relabelled xs, so they share the same Poisson bracket.) The commutator $[\mathsf{W}(f),\mathsf{W}(g)]$ in Weyl's quantization is then

$$[\mathsf{W}(f),\mathsf{W}(g)] = \mathsf{W}(f\star g) - \mathsf{W}(g\star f) = \mathsf{W}[i\hbar\{f,g\} + O(\hbar^2)].$$

This is an entire power series in $\hbar$, and Dirac's Lie algebra homomorphism is recovered as $\hbar \downarrow 0$. Eq. (5) is known as the *Weyl relation* and turns out to be remarkably strong: in fact, it effectively determines the Hilbert space $\mathcal{H}$ and thus the entire quantum mechanics.

THEOREM 1 (Stone–von Neumann). *Let (M,P^{ij}) be a finite dimensional symplectic vector space, and let $(\mathcal{H},\mathsf{W})$ and $(\mathcal{H}',\mathsf{W}')$ be two representations of the Weyl relations. Then $(\mathcal{H},\mathsf{W})$ and $(\mathcal{H}',\mathsf{W}')$ are unitarily equivalent.*

For finite dimensional phase spaces this guarantees that the quantum mechanics we find is unique, provided that we are willing to extend the observable algebra to the full Weyl algebra. (This is usually what physicists, especially quantum field theorists, mean when they talk about observables in quantum mechanics.) Unfortunately we cannot apply this theorem directly to our real interest, quantum field theory, which is the case of an infinite-dimensional phase space. This is disappointing. Nonetheless, there is a remarkable construction based on what are called c^* algebras which can be used to prove uniqueness for

[2]The wavefunction $\Psi(q)$ is a function of the coordinates, whereas the state $|\psi\rangle$ is not. It is easy to see that $\Psi(q)$ has all the other properties we usually expect of a wavefunction.

flat space QFT (see, eg., Wald [**13**]); as far as I know, the full case of a curved background is still an open question. In fact, the c^*-algebra approach involves an interesting role-reversal: the conventional Hilbert space quantization revolves around first constructing states which are vectors in the Hilbert space. One then defines operators, as we have done, as operators on these vectors. The c^*-algebra approach is the reverse: one first constructs observables as elements of some abstract algebra, and defines states as operators on elements of this algebra which associate a real number to each observable. This allows one to treat *all* states, in particular those arising from inequivalent quantum field theory constructions, on an equal footing.

2. The Path Integral

There is another approach to quantization which does not rely on phase space at all, but turns out to provide a remarkably convenient tool for the quantization of many important theories (like gravity, or the gauge theories of contemporary particle physics) which are just about hopeless to try and quantize with the operator methods described in the previous section.

Consider some classical theory described by a Hamiltonian $H = H(p, q)$ where for simplicity we deal only with a single position and momentum. (The extension to the multivariate case is obvious.) The equations of motion are as before,

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \text{and} \quad \dot{p} = -\frac{\partial H}{\partial q}.$$

Now introduce a new quantity L, called the *Lagrangian*, and defined by $L = p\dot{q} - H(p, q)$, where $\dot{q} = \dot{q}(p, q)$ is considered to be a function of p and q. Evidently one has

$$\frac{\partial L}{\partial \dot{q}} = p, \quad \text{and} \quad \frac{\partial L}{\partial q} = -\frac{\partial H}{\partial q},$$

so the Hamilton equation giving $\dot{p}$ becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q}$$

which is known as *Lagrange's equation of motion*, or the *Euler–Lagrange equation*. The momentum p does not appear here, so it can be eliminated in favour of $\dot{q}$ using the other Hamilton equation. For example, for the Newtonian point particle, $H = p^2/2m + V(q)$, so $\dot{q} = p/m$ and the Lagrangian becomes

$$L = \frac{1}{2}m\dot{q}^2 - V(q). \tag{6}$$

The Euler–Lagrange equation then gives back the second law, as we expect.

There is another way to understand how the Euler–Lagrange equation arises. Consider the time integral of the Lagrangian L, which we will denote $S = \int \mathrm{d}t\, L$ and call the *action*. The variation in S induced by a change δq in the particle's trajectory satisfies

$$\begin{aligned}\delta S &= \int \mathrm{d}t\, \left(\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta \dot{q}\right)\\ &= \int \mathrm{d}t\, \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}}\right)\delta q\end{aligned}$$

where we have integrated by parts and discarded a surface term. (This can be justified by setting boundary conditions so that the surface term does actually vanish.) Evidently demanding that the variation δS vanish is equivalent to the Euler–Lagrange equation, so the equations of classical mechanics reduce to the statement that the action is extremized.

This principle of extremal action is an extremely convenient way to study classical systems and (as far as I know) all interesting classical system have such a formulation. Feynman's contribution was to show that the Lagrangian is also important in quantum mechanics. To see this, consider the amplitude for the point particle to propagate from $q = q_a$ at $t = t_a$ to $q = q_b$ at $t = t_b$. This is $K_{ab} = \langle q_b, t_b | q_a, t_b\rangle$ which (especially in QFT) is sometimes called the *propagator*.

The eigenstates of any (bounded, self-adjoint) observable like $\hat{q}$ or $\hat{p}$ form a sufficiently large set that other states in $\mathcal{H}$ can be written as an infinite sum over them:

$$|\psi\rangle = \int \mathrm{d}q\ \psi_q |q\rangle,$$

for some arbitrary state $|\psi\rangle$. One should think of this as a kind of generalized Fourier series. If we normalize the $|q\rangle$ so that $\langle q'|q\rangle = \delta(q'-q)$, then the coefficients ψ_q can be picked out:

$$\begin{aligned}\langle q'|\psi\rangle &= \int \mathrm{d}q\ \psi_q \langle q'|q\rangle \\ &= \int \mathrm{d}q\ \psi_q \delta(q'-q) \\ &= \psi_{q'} \implies |\psi\rangle = \left(\int \mathrm{d}q\ |q\rangle\langle q|\right)|\psi\rangle\end{aligned}$$

A convenient form of this result is to write $1 = \int \mathrm{d}q\ |q\rangle\langle q|$: this is called a *resolution of the identity.*

At any intermediate time t, $t_a < t < t_b$, we can insert such a resolution of the identity:

$$K_{ab} = \langle q_b, t_b|1|q_a, t_a\rangle \implies K_{ab} = \int \mathrm{d}q\ \langle q_b, t_b|q, t\rangle\langle q, t|q_a, t_a\rangle.$$

This process can be repeated for N evenly spaced intermediate times. Define $\varepsilon = (t_b - t_a)/N$, $t_0 = t_a$, $t_N = t_b$ and $t_m = t_a + \varepsilon N$ for $0 < m < N$. Then

$$\begin{aligned}K_{ab} &= \int \mathrm{d}q_1 \cdots \mathrm{d}q_{N-1} \prod_{m=0}^{N-1} \langle q_{m+1}, t_{m+1}|q_m, t_m\rangle \\ &= \int \mathrm{d}q_1 \cdots \mathrm{d}q_{N-1} \prod_{m=0}^{N-1} \langle q_{m+1}|\mathrm{e}^{-i\hat{H}(\hat{p},\hat{q})\varepsilon/\hbar}|q_m\rangle,\end{aligned}$$

where we have used the time evolution equations derived above to evolve $|q_m, t_m\rangle$ into $|q_m, t_{m+1}\rangle$. If we can evaluate the matrix element appearing here, then we are done.

To this end, introduce a further set of partitions of unity, this time summing over intermediate eigenstates of momentum,

$$K_{ab} = \int \mathrm{d}q_1 \cdots \mathrm{d}q_{N-1}\mathrm{d}p_1 \cdots \mathrm{d}p_{N-1} \prod_{m=0}^{N-1} \langle q_{m+1}|p_m\rangle\langle p_m|\mathrm{e}^{-i\hat{H}(\hat{p},\hat{q})\varepsilon/\hbar}|q_m\rangle.$$

We now cheat: if ε is very small so that there are many intermediate times, then $\mathrm{e}^{-i\hat{H}\varepsilon/\hbar}$ can be taken to be at most linear in $\hat{H}$. Therefore, if we choose to order $\hat{H}$ with all the $\hat{p}$ to the left and all the $\hat{q}$ to the right, then we can evaluate $\mathrm{e}^{-i\hat{H}\varepsilon/\hbar}$ just by replacing $\hat{p}$ and $\hat{q}$ with their eigenvalues. Thus,

$$K_{ab} = \int \mathrm{d}q_1 \cdots \mathrm{d}q_{N-1}\mathrm{d}p_1 \cdots \mathrm{d}p_{N-1} \prod_{m=0}^{N-1} \mathrm{e}^{-iH(p_m,q_m)\varepsilon/\hbar}\langle q_{m+1}|p_m\rangle\langle p_m|q_m\rangle. \tag{7}$$

We now only need an expression for the inner product $\langle q|p\rangle$, which can be found by recalling that $\hat{p}$ acts on wavefunctions $\Psi = \langle q|\psi\rangle$ by $-i\hbar\partial/\partial q$. Therefore,

$$-i\hbar\frac{\partial}{\partial q}\langle q|p\rangle = p\langle q|p\rangle \implies \frac{\partial \log\langle q|p\rangle}{\partial q} = \frac{ip}{\hbar}.$$

This equation has the obvious solution $\langle q|p\rangle \propto \mathrm{e}^{ipq/\hbar}$, and the constant of proportionality can be fixed by requring that $\langle q'|q\rangle = \delta(q'-q)$:

$$\begin{aligned}\delta(q'-q) = \langle q'|q\rangle = \int \mathrm{d}p\ \langle q'|p\rangle\langle p|q\rangle &= \int \mathrm{d}p\ |A|^2 \exp\left[\frac{ip}{\hbar}(q'-q)\right] \\ &= 2\pi\hbar|A|^2\delta(q'-q).\end{aligned}$$

If we choose the phase of A to be real, then that gives

$$\langle q|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \mathrm{e}^{ipq/\hbar}.$$

Applying this to the amplitude formula (7) shows that the integrand decomposes into a simple exponential:

$$K_{ab} = \int \mathrm{d}q_1 \cdots \mathrm{d}q_{N-1} \frac{\mathrm{d}p_1}{2\pi\hbar} \cdots \frac{\mathrm{d}p_{N-1}}{2\pi\hbar} \prod_{m=0}^{N-1} \exp\left[\frac{i}{\hbar}\left(q_{m+1}p_m - q_m p_m - H(p_m, q_m)\varepsilon\right)\right].$$

To make progress, let us move the product $\prod_m$ into the argument of the exponential where it becomes a sum, and factor out the term $1/\varepsilon$ to obtain

$$K_{ab} = \left(\int \prod_{m=1}^{N-1} \frac{\mathrm{d}q_m \mathrm{d}p_m}{2\pi\hbar}\right) \exp\left(\frac{i}{\hbar} \sum_{m=0}^{N-1} \varepsilon \left[p_m \frac{q_{m+1} - q_m}{\varepsilon} - H(p_m, q_m)\right]\right).$$

We now take the limit $\varepsilon \downarrow 0$ and agree to denote the infinite product of integrals which appears at the front as $\int [\mathrm{d}q\,\mathrm{d}p]$. This product effectively constitutes an integral over the *functions* $p(t)$ and $q(t)$. In the meantime, the sum in the exponent clearly converges to the Riemann integral of $p_m \dot{q}_m - H$, provided that all the relevant quantities exist:

$$K_{ab} = \int [\mathrm{d}q\,\mathrm{d}p] \exp\left(\frac{i}{\hbar} \int_{t_a}^{t_b} \mathrm{d}t\, (p\dot{q} - H)\right).$$

This is the *functional integral* or *path integral* for the amplitude. In old-fashioned books, the notation $\mathcal{D}p$, $\mathcal{D}q$ is often used for the functional measure.

The integrand in the argument of the exponential looks like it produces the Lagrangian $L = p\dot{q} - H$, except for the fact that p and q are simply variables of integration and are not constrained to obey Hamilton's equations. However there is one important case in which the exponent really does reduce to $i/\hbar$ times the action. To see how this arises, consider the ordinary integral of a quadratic exponential:

$$\begin{aligned} z &= \int \mathrm{d}x\ \mathrm{e}^{-(ax^2+bx+c)} \\ &= \int \mathrm{d}x\ \exp\left[-a(x + \frac{b}{2a})^2 + c - \frac{b^2}{4a}\right] \\ &= \exp\left(c - \frac{b^2}{4a}\right) \int \mathrm{d}x\ \exp\left[-ax^2\right], \end{aligned}$$

where we have completed the square in the argument and in the last line shifted the definition of the integration variable x. The contribution from the remaining integral can be done explicitly, but there is no need: it is just a constant which can be absorbed into the normalization. We will have to consider this constant in more detail later, but for the time being it is enough to note that the exponential term is the original $\mathrm{e}^{-(ax^2+bx+c)}$ evaluated at the extremum of its argument. If this carries over to the path integral, then we expect that for quadratic Hamiltonians, doing the $[\mathrm{d}p]$ integration results in a normalization times the exponential evaluated at the minimum of its argument. This is:

$$\delta_p \int_{t_a}^{t_b} (p\dot{q} - H) = \int_{t_a}^{t_b} \left(\dot{q} - \frac{\partial H}{\partial p}\right) \delta p.$$

Therefore,

$$\frac{\delta}{\delta p} \int_{t_a}^{t_b} (p\dot{q} - H) = 0 \implies \dot{q} - \frac{\partial H}{\partial p} = 0.$$

This is just Hamilton's equation, so evaluating the exponential at this point results in exactly the Lagrangian, and we have

$$\boxed{K_{ab} = \int [\mathrm{d}q]_{q_a}^{q_b}\ \exp\left(\frac{i}{\hbar} \int_{t_a}^{t_b} \mathrm{d}t\ L\right) = \int [\mathrm{d}q]_{q_a}^{q_b}\ \mathrm{e}^{iS[q]/\hbar}.} \tag{8}$$

This is the *Lagrangian path integral formula*. We write the boundary conditions q_a and q_b in explicitly as a reminder that the integration is only over functions q which obey $q = q_a$ and $q = q_b$ at the endpoints.

In fact, one can make an argument that this is formula is valid much more generally than under the rather restrictive conditions with which we have derived it. The relation $\dot{q} = \partial H/\partial q$ is local, so the corrections to the stationary phase approximation are local and can be absorbed into the choice of measure. Also, we were rather cavalier about the limit $\varepsilon \downarrow 0$. The study of this limiting process is called *renormalization*, and (miraculously) is rather well understood. The limit exists provided the Lagrangian satisfies certain conditions (renormalizability) and provided the coefficients in the Lagrangian (bare couplings) are taken to vary appropriately as the limit is taken. The result is independent of the operator ordering choice in the Hamiltonian, the precise definition of the measure and the differences between regulators. One can then adopt the position that, with some appropriate 'good' choice of measure, the path integral is a perfectly satisfactory *definition* of the quantum theory.

There are a couple of points to note about the path integral. The first is that *there are no operators*. Everything which appears in the integration is a classical quantity, and so commutes. (Similarly, there is no Hilbert space.) To understand how this works, consider the functional integration of some quantity q on the range $t = t_a$ to $t = t_b$, obeying the boundary conditions $q(t_a) = q_a$ and $q(t_b) = q_b$, as before. We can decompose this integral into paths on (t_a, t_m), for some intermediate time $t = t_m$, and paths on (t_m, t_b), with some fixed boundary condition $q(t_m) = q_m$ at $t = t_m$. Integrating over all possible boundary conditions should then give back our original integral:

$$\int [\mathrm{d}q]_{q_a,t_a}^{q_b,t_b} = \int \mathrm{d}q_m \int [\mathrm{d}q]_{q_a,t_a}^{q_m,t_m} \int [\mathrm{d}q]_{q_m,t_m}^{q_b,t_b}.$$

Applying this to the path integral representation of the amplitude K_{ab} gives the partition of unity result we have seen before:

$$\begin{aligned}
\int [\mathrm{d}q]_{q_a,t_a}^{q_b,t_b} \exp\left(\frac{i}{\hbar}\int_{t_a}^{t_b} \mathrm{d}t\ L\right) &= \int [\mathrm{d}q]_{q_a,t_a}^{q_m,t_m} \int [\mathrm{d}q]_{q_m,t_m}^{q_b,t_b} \exp\left(\frac{i}{\hbar}\int_{t_a}^{t_m} \mathrm{d}t\ L + \frac{i}{\hbar}\int_{t_m}^{t_b} \mathrm{d}t\ L\right) \\
&= \int \mathrm{d}q_m \int [\mathrm{d}q]_{q_a,t_a}^{q_m,t_m} \exp\left(\frac{i}{\hbar}\int_{t_a}^{t_m} \mathrm{d}t\ L\right) \int [\mathrm{d}q]_{q_m,t_m}^{q_b,t_b} \exp\left(\frac{i}{\hbar}\int_{t_m}^{t_b} \mathrm{d}t\ L\right) \\
&= \int \mathrm{d}q_m\ \langle q_b, t_b | q_m, t_m\rangle \langle q_m, t_m | q_a, t_a\rangle.
\end{aligned}$$

However, we can equally well do this with extra insertions of operators $q(t_1)q(t_2)\cdots q(t_n)$ in the path integral. To evaluate this, slice the integral at times t_1, t_2 and so on, to obtain

$$\begin{aligned}
&\int [\mathrm{d}q]_{q_a}^{q_b} \exp\left(\frac{i}{\hbar}\int_{t_a}^{t_b} \mathrm{d}t\ L\right) q(t_1)q(t_2)\cdots q(t_n) \\
&\qquad = \int \mathrm{d}q_1 \cdots \mathrm{d}q_n \int [\mathrm{d}q]_{q_a}^{q_1} \cdots [\mathrm{d}q]_{q_n}^{q_b} \exp\left[\frac{i}{\hbar}\left(\int_{t_a}^{t_1} + \cdots + \int_{t_n}^{t_b}\right) \mathrm{d}t\ L\right] q(t_1)q(t_2)\cdots q(t_n) \\
&\qquad = \int \mathrm{d}q_a \cdots \mathrm{d}q_n \langle q_b, t_b | q_n, t_n\rangle q_n \cdots q_1 \langle q_1, t_1 | q_a, t_a\rangle \\
&\qquad = \int \mathrm{d}q_a \cdots \mathrm{d}q_n \langle q_b, t_b | \hat{q}(t_n) | q_n, t_n\rangle \cdots \langle q_2, t_2 | \hat{q}(t_1) | q_1, t_1\rangle \langle q_1, t_1 | q_a, t_a\rangle \\
&\qquad = \langle q_b, t_b | \mathrm{T}(\hat{q}(t_1)\cdots \hat{q}(t_2)) | q_a, t_a\rangle.
\end{aligned}$$

Because of the way the path integral slices up into time segments, the operators end up *time ordered*: the time ordering symbol T rewrites its argument with the operator at smallest time on the far right, and operators at successively later times stacked to its left.

From now on, we will usually choose the boundary conditions on the fields q to correspond to the ground state, or state of minimum excitation. We will omit writing these boundary conditions on the functional measure; if particular boundary conditions are required, then we will write them in explicitly. Then for some arbitrary local operator $\mathcal{A}$ composed of the elementary fields

$$\int [\mathrm{d}q]\ \mathrm{e}^{iS/\hbar}\mathcal{A} = \langle 0|\mathcal{A}|0\rangle = \langle\mathcal{A}\rangle \tag{9}$$

is called the *vacuum expectation value* of $\mathcal{A}$.

We are now in a position to derive a lot of the basic results of quantum mechanics. However these will be quite easy using the path integral.

2.1. Ehrenfest's theorem. If the path integral is anything like normal integration, we expect the integral of a total derivative to be zero. Therefore,

$$\begin{aligned} 0 &= \int [\mathrm{d}q]\ \frac{\delta}{\delta q}\mathrm{e}^{iS/\hbar} \\ &= \int [\mathrm{d}q]\ \frac{i}{\hbar}\frac{\delta S}{\delta q}\mathrm{e}^{iS/\hbar} \end{aligned}$$

Thus, $\langle\delta S/\delta q\rangle = 0$. This is Ehrenfest's theorem: the classical equations of motion hold inside expectation values. (Actually there are subtleties when there are insertions in the path integral, but we ignore these for the time being.)

2.2. Noether's theorem. Now suppose that the theory we are studying possesses some symmetry, a set of transformations on the quantities in the theory that leave it invariant. In the case of the point particle, we can consider changing the origin of time, ie., making the transformation $t \mapsto t + \varepsilon$. The fields and Lagrangian change like

$$q \mapsto q + \varepsilon\frac{\mathrm{d}q}{\mathrm{d}t}; \quad L \mapsto q + \varepsilon\frac{\mathrm{d}L}{\mathrm{d}t}.$$

If ε is just a constant, not depending on time, then the change δL is the total time derivative $\mathrm{d}(\varepsilon L)/\mathrm{d}t$ which does not contribute to the action and so may be discarded. Such a transformation, with ε constant, is said to be a *rigid* or *global* symmetry. If one now allows ε to vary with t, then L need not be invariant, but must transform like

$$L \mapsto L + \frac{\mathrm{d}}{\mathrm{d}t}(\varepsilon L) + \frac{\mathrm{d}\varepsilon}{\mathrm{d}t}j$$

for some quantity j, in order that the previous result is recovered when $\dot{\varepsilon} = 0$. Integrating by parts and discarding the surface term gives

$$\delta S = -\int \mathrm{d}t\ \varepsilon\frac{\mathrm{d}j}{\mathrm{d}t}, \tag{10}$$

so when the classical equations $\delta S/\delta q = 0$ are satisfied one must have $\mathrm{d}j/\mathrm{d}t = 0$ and j is a conserved quantity. This is *Noether's theorem*: global symmetries give rise to conserved quantities. One can show by an explicit calculation that j is in fact just the conserved total energy $m\dot{q}^2/2 + V(q)$ we found earlier.

One can see the same effect arising at the quantum level. Discarding total derivatives, the functional integral changes like

$$\int [\mathrm{d}q]\mathrm{e}^{(i/\hbar)\int \mathrm{d}t\ L} \mapsto \int [\mathrm{d}q]\exp\left[\frac{i}{\hbar}\int \mathrm{d}t\ \left(L - \varepsilon\frac{\mathrm{d}j}{\mathrm{d}t}\right)\right].$$

However, $\int[\mathrm{d}q]$ integrates over all paths, whatever their parametrization, so the functional integral cannot change under this transformation. Thus the left and right hand sides of the above transformation must in fact be equivalent:

$$\int [\mathrm{d}q]\mathrm{e}^{iS[q]/\hbar} = \int [\mathrm{d}q]\mathrm{e}^{iS[q]/\hbar}\left(1 - \frac{i}{\hbar}\int \mathrm{d}t\ \varepsilon\frac{\mathrm{d}j}{\mathrm{d}t}\right).$$

Therefore,

$$\int \mathrm{d}t\ \varepsilon\langle\frac{\mathrm{d}j}{\mathrm{d}t}\rangle = 0$$

and $\langle\mathrm{d}j/\mathrm{d}t\rangle = 0$ holds as an operator equation, ie. inside expectation values.

2.3. Calculating functional integrals. Let us return to the Newtonian point particle and at last try to evaluate the functional integral explicitly. For simplicity we pick the free case where the potential $V(q)$ vanishes, so the propagator takes the form

$$K_{ab} = \int [\mathrm{d}q]_{q_a,t_a}^{q_b,t_b} \exp\left(\frac{i}{\hbar}\int_{t_a}^{t_b} \mathrm{d}t\, \frac{1}{2}m\dot{q}^2\right).$$

To see how to evaluate this, consider the more general case where the action is quadratic in the q with coefficient given by some arbitrary second order self-adjoint differential operator which we will choose to call $\triangle$: $K_{ab} = \int [\mathrm{d}q]\,\mathrm{e}^{-(1/2)\int \mathrm{d}t\, q\triangle q}$. (An operator is said to be *self-adjoint* if it satisfies the requirement $\int y_n \triangle y_m = \int y_m \triangle y_n$.) The q appearing in the path integral are constrained by the boundary conditions $q(t_a) = q_a$ and $q(t_b) = q_b$, as before, so we can expand them as a series in the eigenfunctions of $\triangle$:

$$q = \sum_n q_n y_n(t), \quad \text{where} \quad \triangle y_n(t) = \lambda_n y_n(t).$$

Once more, one should understand this as a generalization of the usual Fourier series, which are eigenfunctions of the flat Laplacian. As in the usual proof of Dirichlet's theorem, the resulting eigenfunction representation converges almost everywhere, provided q has only a finite number of finite discontinuities.

To obtain all possible functions constructed this way, one integrates over all possible coefficients q_n, so the functional measure can be taken to be

$$\int [\mathrm{d}q] = \prod_n \mu\, \mathrm{d}q_n$$

where μ is some unknown constant of proportionality. The q_n are pure numbers, so each of the $\int \mathrm{d}q_n$ are just normal integrals. Also, the y_n corresponding to different eigenvalues λ_n are chosen to be orthonormal, so that

$$\int \mathrm{d}t\; y_n y_m = \delta_{nm}.$$

That this is possible follows from self-adjointness of $\triangle$. (To see this, write $\int \mathrm{d}t\; y_n \triangle y_m = \lambda_m \int \mathrm{d}t\; y_n y_m = \lambda_n \int \mathrm{d}t\; y_n y_m$ by using the self-adjoint property; then if $\lambda_n \neq \lambda_m$ the integral must vanish.) The functional integral then becomes

$$\begin{aligned} K_{ab} &= \left(\prod_n \mu\, \mathrm{d}q_n\right) \exp\left(-\frac{1}{2}\int \mathrm{d}t\, \sum_{n,m} q_n q_m \lambda_m y_n(t) y_m(t)\right) \\ &= \left(\prod_n \mu\, \mathrm{d}q_n\right) \exp\left(-\frac{1}{2}\sum_n q_n^2 \lambda_n\right) \\ &= \prod_n \int \mu\, \mathrm{d}q_n\; \mathrm{e}^{-q_n^2\lambda_n/2} \end{aligned}$$

in virtue of the orthogonality. However, this integral is easy to do: the result $I = \int_{-\infty}^{\infty} \mathrm{d}x\; \mathrm{e}^{-x^2} = \sqrt{\pi}$ is standard. Recall that one way to show this is to calculate I^2 instead,

$$\begin{aligned} I^2 &= \int \mathrm{d}x\, \mathrm{d}y\; \mathrm{e}^{-x^2-y^2} = \int r\, \mathrm{d}r\, \mathrm{d}\theta\; \mathrm{e}^{-r^2} \\ &= 2\pi \int \left(-\frac{1}{2}\right) \mathrm{d}(\mathrm{e}^{-r^2}) \\ &= -\pi[\mathrm{e}^{-r^2}]_0^\infty = \pi. \end{aligned}$$

(We are being sloppy about the interchange of limits.) As a result, one obtains an explicit expression for the propagator:

$$K_{ab} = \prod_n \sqrt{\frac{2\pi\mu^2}{\lambda_n}}.$$

We can recast this result in a useful way. The determinant of a matrix is defined to be the product $\prod_n \lambda_n^{\text{matrix}}$ of its eigenvalues, so by analogy we can invent a similar concept for the self-adjoint operator $\triangle$ and declare that $\det \triangle = \prod_n \lambda_n$. Then we have the simple result

$$\boxed{K_{ab} = \left(\det \frac{\triangle}{2\pi\mu^2}\right)^{-1/2}.} \tag{11}$$

Let us apply this general formalism to the free Newtonian point particle. We split q into a classical and quantum part $q = q_{\text{cl}} + q_{\text{q}}$ where q_{cl} obeys the boundary conditions $q_{\text{cl}}(t_a) = q_a$ and $q_{\text{cl}}(t_b) = q_b$, and the classical field equation $\ddot{q}_{\text{cl}} = 0$. Then

$$\exp\left(\frac{i}{\hbar}\int \mathrm{d}q\ \frac{m}{2}\dot{q}^2\right) = \exp\left[\frac{i}{\hbar}\int \mathrm{d}q\ \frac{m}{2}\left(\dot{q}_{\text{cl}}^2 + 2\dot{q}_{\text{cl}}\dot{q}_{\text{q}} + \dot{q}_{\text{q}}^2\right)\right] = \mathrm{e}^{iS_{\text{cl}}/\hbar}\exp\left(\frac{i}{\hbar}\int \mathrm{d}q\ \frac{m}{2}\dot{q}_{\text{q}}^2\right),$$

where we have integrated by parts and used the boundary conditions on q_{q}, and the fact that q_{cl} obeys the classical equation of motion. Here, S_{cl} is the action of the classical path between q_a and q_b. Thus, integrating by parts again, the propagator satisfies

$$K_{ab} = \mathrm{e}^{iS_{\text{cl}}/\hbar}\int [\mathrm{d}q_{\text{q}}]\ \exp\left(-\frac{1}{2}\int \mathrm{d}t\ q_{\text{q}}\left(\frac{im}{\hbar}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\right)q_{\text{q}}\right),$$

where the coefficient of q_{q}^2 is a self-adjoint second order differential operator in virtue of the boundary conditions on q_{q}. (It should be remarked that the property of self-adjointness depends on the operator itself and whatever boundary conditions are employed.) We can now apply our result (11) above. To find the eigenvalue spectrum, we simply solve the relevant eigenvalue equation:

$$\frac{im}{\hbar}\frac{\mathrm{d}^2 y_n}{\mathrm{d}t^2} = \lambda_n y_n \quad \text{subject to } y_n = 0 \text{ at } t = t_a, t_b.$$

The solutions are

$$y_n = A\cos\sqrt{\frac{\hbar\lambda_n}{im}}t + B\sin\sqrt{\frac{\hbar\lambda_n}{im}}t$$

and the boundary conditions then imply

$$\left.\begin{aligned} A\cos\sqrt{\frac{\hbar\lambda_n}{im}}t_a + B\sin\sqrt{\frac{\hbar\lambda_n}{im}}t_a &= 0 \\ A\cos\sqrt{\frac{\hbar\lambda_n}{im}}t_b + B\sin\sqrt{\frac{\hbar\lambda_n}{im}}t_b &= 0 \end{aligned}\right\} \quad \text{so} \quad \tan\sqrt{\frac{\hbar\lambda_n}{im}}t_a = \tan\sqrt{\frac{\hbar\lambda_n}{im}}t_b.$$

Periodicity of $\tan\theta$ then means that

$$\sqrt{\frac{\hbar\lambda_n}{im}}t_b - \sqrt{\frac{\hbar\lambda_n}{im}}t_a = n\pi, \quad n \in \mathbb{Z}$$

$$\text{so,} \quad \lambda_n = \frac{n^2\pi^2}{\Delta t^2}\frac{im}{\hbar} \quad \text{where } \Delta t = t_b - t_a.$$

With this expression for the eigenvalues, the operator determinant we are trying to calculation comes out to be

$$\det \triangle = \prod_n \lambda_n = \prod_n \frac{n^2\pi^2}{\Delta t^2}\frac{im}{\hbar}.$$

Unfortunately, this is product is embarassingly infinite. To fix it, we must find some way to truncate or otherwise sanitize the infinity so that it can be handled in subsequent manipulations. This process is called *regularization*; we will borrow a regularization from number theory introduced by Hawking in the late 1970s for the purpose of studying quantum gravity. Let us define the *zeta function* of a self-adjoint differential operator $\triangle$ by (formally)

$$\zeta_\triangle(s) = \sum_n \lambda_n^{-s} \quad \text{where } s \in \mathbb{C}.$$

This infinite sum converges at least wherever $\mathrm{Re}\, s > 2$,[3] but has a unique analytic continuation to the entire complex plane which defines a meromorphic function of s (that is, the analytic continuation has at worst isolated singularities, but no branch cuts).

Notice that one typically uses the letter s to denote the argument of a zeta function, although z is typical elsewhere in complex analysis. This convention arises from Riemann's original paper on the prototypical zeta function, the Riemann zeta function, defined by $\zeta_R(s) = \sum_{n\geqslant 1} n^{-s}$. The Riemann zeta function remains one of the most important functions in analysis, and probably the most outstanding unproved conjecture in mathematics is the Riemann hypothesis: that all the non-trivial zeroes of the Riemann zeta function lie on the line $\mathrm{Re}\, s = 1/2$. By proving that the Riemann zeta function has no zeros on $\mathrm{Re}\, s = 1$, Hadamard was able to prove the prime number theorem, $\pi(n) \sim \mathrm{li}(n) = \int_0^n \frac{\mathrm{d}t}{\ln t}$. where $\pi(n)$ counts the number of primes less than n and $\mathrm{li}(n)$ is the polylogarithm or logarithmic integral. Later, Wiener showed that the full Riemann hypothesis was enough to obtain the stronger result

$$\pi(n) = \int_0^n \frac{\mathrm{d}t}{\ln t} + O(x^{1/2} \ln x). \tag{12}$$

The Riemann hypothesis has connexions to many other deep results in analysis. Nevertheless, the Riemann hypothesis has resisted all attacks, computational and otherwise, and there is a \$1m prize on offer to the first mathematician to provide a proof.

It is remarkable that such deep results from number theory can provide a profitable input to physics, but quantum mechanics, supersymmetry, and string theory all provide miraculous examples. In some cases, the consistency of the theory depends on startling mathematical "coincidences," like the Green–Schwarz anomaly cancellation or Jacobi's 'abstruse' identity which leads to modular invariance of the one-loop superstring graph. These coincidences lead many workers to believe that even if we are not on the right track, we are at least discovering something interesting.

By differentiating the zeta function with respect to its argument one obtains

$$\zeta'_\triangle(s) = -\sum_n (\ln \lambda_n)\lambda_n^{-s} \implies -\zeta'_\triangle(0) = \sum_n \ln \lambda_n = \ln(\prod_n \lambda_n) = \ln \det \triangle$$

This serves as a definition of the determinant $\det \triangle$. One can include the factor $2\pi\mu^2$ by writing

$$\begin{aligned} \det \frac{\triangle}{2\pi\mu^2} &= \prod_n \frac{\lambda_n}{2\pi\mu^2} \\ \implies \ln \det \frac{\triangle}{2\pi\mu^2} &= -\sum_n \ln 2\pi\mu^2 + \sum_n \ln \lambda_n \\ &= -(\ln 2\pi\mu^2)(\sum_n 1) - \zeta'_\triangle(0) = -\zeta_\triangle(0)(\ln 2\pi\mu^2) - \zeta'_\triangle(0). \end{aligned}$$

The expression we need is thus

$$\left(\det \frac{\triangle}{2\pi\mu^2}\right)^{-1/2} = (2\pi\mu^2)^{\zeta_\triangle(0)/2} \exp\left(\frac{1}{2}\zeta'_\triangle(0)\right).$$

This is unambiguous provided one adopts the convention that all zeta functions are written in terms of the λ_n alone, and not decomposed further. Notice that physics depends on the unknown parameter μ unless $\zeta_\triangle(0) = 0$. If this is true, then the theory is said to be *finite*; otherwise, we will have to prescribe some way to fix the value of μ. This process is known as *renormalization*.

[3] We are being conservative. We have already seen that the eigenvalues of a second-order self-adjoint operator typically go like $\lambda_n \sim n^2$ as $n \uparrow \infty$. In this case, $\zeta(1) \sim \sum_n n^{-2} \sim \pi^2/6$, so the infinite sum defining the zeta function converges. This is perfectly satisfactory in the bosonic sector of the theory, but in realistic cases one will have supersymmetric partner fermionic operators like the Dirac operator $(i\partial\!\!\!/ + m)\Psi$ acting on some fermion Ψ which square to bosonic operators. Such operators are *first* order and their eigenvalues typically have the weaker behaviour $\lambda_n \sim n$ as n becomes large. In this case, one needs the stronger restriction $\mathrm{Re}\, s > 2$ to obtain convergence.

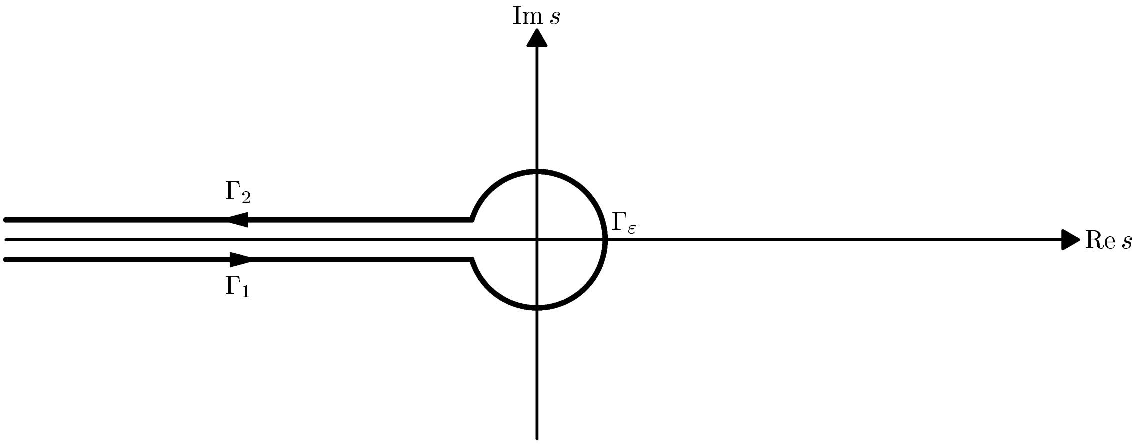
In our case, the zeta function is

$$\zeta_\triangle(s) = \sum_n \left(\frac{n^2\pi^2}{\Delta t^2}\frac{im}{\hbar}\right)^{-s} = \left(\frac{\pi^2}{\Delta t^2}\frac{im}{\hbar}\right)^{-s} \zeta_R(2s)$$

where as before $\zeta_R(s)$ is the Riemann zeta function. However, the expression for the operator determinant instructs us to calculate $\zeta_\triangle(0)$ and $\zeta'_\triangle(0)$, which is a region of the s-plane where the infinite sum representation of $\zeta_R(s)$ does not converge. To do better, we need to carry out the analytic continuation. Consider the function $\Gamma(s)$ defined by

$$\frac{1}{\Gamma(s)} = \frac{1}{2\pi i}\int_{\Gamma_1+\Gamma_\varepsilon+\Gamma_2} \mathrm{e}^t t^{-s}\,\mathrm{d}t = \frac{1}{2\pi i}\int_{-\infty}^{(0+)} \mathrm{e}^t t^{-s}\,\mathrm{d}t$$

where the contours Γ_1, Γ_ε and Γ_2 are shown below; we adopt the convention of notating this contour by $\int_{-\infty}^{(0+)}$. (This is called the *Hankel contour.*)



This is the Hankel representation of the Γ-function, $\Gamma(s) = \int_0^\infty t^{s-1}\mathrm{e}^{-s}\,\mathrm{d}t$. (It is not too hard to show this by writing out the Γ_1 and Γ_2 integrals explicitly. The Γ_ε contour does not contribute. One also needs the identity $\Gamma(s)\Gamma(1-s) = \pi/\sin\pi s$.) We make the change of variable $t \mapsto n\tau$ to find

$$\frac{2\pi i}{\Gamma(1-s)} = \int_{-\infty}^{(0+)} \mathrm{e}^{n\tau}\tau^{s-1}n^{s-1}\,n\,\mathrm{d}\tau = n^s\int_{-\infty}^{(0+)} \mathrm{e}^{n\tau}\tau^{s-1}\,\mathrm{d}\tau,$$

and this lets us write an alternative representation of the zeta function:

$$\zeta_R(s) = \sum_{n\geqslant 1} n^{-s} = \frac{\Gamma(1-s)}{2\pi i}\int_{-\infty}^{(0+)} \tau^{s-1}\sum_{n\geqslant 1}\mathrm{e}^{n\tau}\,\mathrm{d}\tau.$$

It is easy to sum the infinite series appearing here; it is

$$\sum_{n=1}^{N}\mathrm{e}^{n\tau} = \frac{\mathrm{e}^\tau - \mathrm{e}^{(N+1)\tau}}{1-\mathrm{e}^\tau} \implies \sum_{n\geqslant 1}\mathrm{e}^{n\tau} = \frac{1}{\mathrm{e}^{-\tau}-1},$$

provided that $\operatorname{Re}\tau < 0$, which is true on our integration contour. Thus,

$$\zeta_R(s) = \frac{\Gamma(1-s)}{2\pi i}\int_{-\infty}^{(0+)} \frac{\tau^{s-1}}{\mathrm{e}^{-\tau}-1}\,\mathrm{d}\tau.$$

We can now evaluate $\zeta_R(0)$ using Cauchy's theorem, $\oint f(z)\,\mathrm{d}z = 2\pi i\sum_i \operatorname{Res}_{z\to z_i} f(z)$ where the z_i are the singularities of $f(z)$. The singularities of the integrand for $s=0$ evidently occur only at $\tau = 1$, so we expand in a Laurent series to find

$$\begin{aligned}\zeta_R(0) &= \frac{\Gamma(1)}{2\pi i}\cdot 2\pi i \operatorname*{Res}_{\tau\to 0}\left(\frac{1}{\tau}\frac{1}{\mathrm{e}^{-\tau}-1}\right)\\ &= \operatorname*{Res}_{\tau\to 0}\left[\frac{1}{\tau}\left(1-\tau+\frac{1}{2}\tau^2+\cdots-1\right)^{-1}\right] = \operatorname*{Res}_{\tau\to 0}\left[-\frac{1}{\tau^2}\left(1-\frac{1}{2}\tau+\cdots\right)^{-1}\right]\\ &= \operatorname*{Res}_{\tau\to 0}\left[-\frac{1}{\tau^2}\left(1+\frac{1}{2}\tau+\cdots\right)\right] = -\frac{1}{2}\end{aligned}$$

Thus (in some sense), not only is $\sum_{n\geqslant 1} 1$ finite, it is actually negative. One can also work out $\sum_{n\geqslant 1} n = -1/12$ and $\zeta_R'(0) = -(1/2)\ln 2\pi$ in the same way, although we do not give the details.

For our purposes, this means that the $\triangle$ zeta function satisfies

$$\zeta_\triangle(0) = -\frac{1}{2} \quad \text{and} \quad \zeta_\triangle'(0) = \frac{1}{2}\ln\left(\frac{\pi^2}{\Delta t^2}\frac{im}{\hbar}\right) - \ln 2\pi.$$

The fact that $\zeta_\triangle(0)$ is non-vanishing implies that we will have to renormalize, so there is not much point keeping track of the overall constant:

$$\begin{aligned}\left(\det\frac{\triangle}{2\pi\mu^2}\right)^{-1/2} &= (2\pi\mu^2)^{-1/4}\exp\left[\frac{1}{4}\ln\left(\frac{\pi^2}{\Delta t^2}\frac{im}{\hbar}\right) - \frac{1}{2}\ln(2\pi)\right]\\ &\sim \frac{1}{\sqrt{\Delta t}}\end{aligned}$$

The propagator K_{ab} is thus $K_{ab} \sim (\Delta t)^{-1/2}\mathrm{e}^{iS_{\mathrm{cl}}/\hbar}$.

The classical path satisfies $\ddot{q}_{\mathrm{cl}} = 0$, so $q_{\mathrm{cl}} = At + B$. The boundary conditions imply

$$\left.\begin{aligned} q_a &= At_a + B\\ q_b &= At_b + B\end{aligned}\right\} \implies A = \frac{q_b - q_a}{t_b - t_a} = \frac{\Delta q}{\Delta t}, \quad B = q_a t_b - q_b t_a,$$

so the classical action comes out to be

$$\exp\left(\frac{i}{\hbar}\int_{t_a}^{t_b}\frac{1}{2}m\dot{q}^2\,\mathrm{d}t\right) = \exp\left(\frac{i}{\hbar}\frac{m}{2}\frac{\Delta q^2}{\Delta t^2}\Delta t\right) \implies K_{ab} \sim \frac{1}{\sqrt{\Delta t}}\exp\left(\frac{im}{2\hbar}\frac{\Delta q^2}{\Delta t}\right).$$

The overall constant is fixed by requiring that $K_{ab} \to \delta(q_a - q_b)$ as $\Delta t \downarrow 0$ and can be found explicitly by using the representation of the delta function,

$$\lim_{\Delta t\downarrow 0}\sqrt{\frac{m}{2\pi i\hbar\Delta t}}\exp\left(\frac{im\Delta q^2}{2\hbar\Delta t}\right) = \delta(\Delta q).$$

This agrees with the propagator found using more pedestrian methods.

3. References

A readable account of the development of (modern) quantum mechanics in general, and quantum field theory in particular, can be found in the first chapter of Weinberg [**14**].

Phase space quantization techniques are described in many places in the literature. The book by Wald [**13**] introduces the Weyl relations and describes the Stone–von Neumann theorem. There is also considerable material on the c^*-algebra QFT and some applications to curved spacetime. For more details and an overview of the geometric quantization of Kostant and Souriau, there is an article by Deligne available online at

`http://www.cgtp.duke.edu/QFT`,

and some notes by John Baez at

`http://math.ucr.edu/home/baez/quantization.html`.

There are many other useful things at this site.

A short treatment of geometric quantization form the point of view of central extensions of Lie algebras can be found in the book by de Azcárraga and Izquierdo [**3**]. There is a brief introduction to star products and the deformation quantization of Flato, Lichnerowicz and Sternheimer at

`http://idefix.physik.uni-freiburg.de/~stefan/quant_eng.html`;

see also the recent reviews by Gutt [**9**] and Sternheimer [**12**]. There is an article by Weinstein at

`http://math.berkeley.edu/~alanw/`.

See also the deformation quantization homepage at

`http://homepages.uni-regensburg.de/~sij17370/download/quantization1.ps`.

Path integral quantization is treated in any modern field theory text (and even some good quantum mechanics books). The classical reference is Feynman & Hibbs [**8**], but the methods used to evaluate integrals are ugly and old-fashioned; the state of the art is described in a three-volume field theory treatise by Weinberg [**14**]. (Nevertheless, the rather more economical derivation of the path integral formula given here is modelled after Appendix A of Polchinski's modern string theory book [**11**], and the discussion of the operator determinant basically follows D'Eath [**6**].) Additionally there is a good discussion of many technical issues in quantum mechanics in the on-line book by Gerald Teschl at

`http://www.mat.univie.ac.at/~gerald/ftp/book-schroe`

One should not assume that canonical and path integral methods are the only approaches to quantization. The book by Carlip [**5**] discusses at least fifteen different approaches to the quantization of 2+1 dimensional gravity.

Zeta function regularization is discussed in many places, eg., (briefly) in Birrell & Davis [**4**]; the books by Albeverio *et. al.* [**2**] and by Moss [**10**], and the papers of Elizalde (best found on hep-th). There is also a large and expensive textbook by Elizalde *et. al.* [**7**], but I have never seen this. The analytic continuation is discussed in any book on complex analysis, eg., the modern text of Ablowitz & Fokas [**1**].

Riemann's zeta function has a fascinating history and is worth many books alone. A short introduction giving many of the classical results (and some spectacularly unbelieveable identities) can be found at

`http://mathworld.wolfram.com/RiemannZetaFunction.html`;
`http://mathworld.wolfram.com/RiemannHypthesis.html`.

Bibliography

1. Mark. J. Ablowitz and Athanassios S. Fokas, *Complex variables: Introduction and applications*, Cambridge University Press, Cambridge, 1997.
2. Sergio Albeverio, Jürgen Jost, Sylvie Paycha, and Sergio Scarlatti, *A mathematical introduction to string theory*, Cambridge University Press, Cambridge, 1997.
3. José Azcárraga and José de Izquierdo, *Lie groups, lie algebras, cohomology and some applications in physics*, Cambridge University Press, Cambridge, 1996.
4. Neil D. Birrell and Paul C.W. Davies, *Quantum fields in curved space*, Cambridge University Press, Cambridge, 1982.
5. Steven Carlip, *Quantum gravity in 2+1 dimensions*, Cambridge University Press, Cambridge, 1998.
6. Peter D. D'Eath, *Supersymmetric quantum cosmology*, Cambridge University Press, Cambridge, 1994.
7. E Elizalde, S.D. Odintsov, A. Romeo, A.A. Bytsenko, and S. Zerbini, *Zeta regularization techniques with applications*, World Scientific, 1994.
8. Richard P. Feynman and A.R. Hibbs, *Quantum mechanics and path integrals*, McGraw–Hill, 1965.
9. Simon Gutt, *Variations on deformation quantization*, (2000), `math.DG/0003107`.
10. Ian Moss, *Quantum theory, black holes and inflation*, Wiley, 1996.
11. Joseph Polchinski, *String theory*, Cambridge University Press, Cambridge, 1994, two volumes.
12. Daniel Sternheimer, *Deformation quantization: Twenty years after*, (1998), `math.QA/9809056`.
13. Robert M. Wald, *Quantum field theory in curved spacetime and black hole thermodynamics*, University of Chicargo Press, Chicargo, 1994.
14. Steven Weinberg, *The quantum theory of fields*, Cambridge University Press, Cambridge, 1994, three volumes.